European Commission

Trends and Developments in Artificial Intelligence

Challenges to the Intellectual Property Rights Framework

Final report

Prepared by:

IViR

JIIP Joint Institute for Innovation Policy

Communications Networks, Content and Technology

EUROPEAN COMMISSION

Directorate-General for Communications Networks, Content and Technology
Directorate A – Artificial Intelligence and Digital Industry
Unit A.1 – Robotics and Artificial Intelligence
Contact: cnect-a1@ec.europa.eu

European Commission
B-1049 Brussels

EUROPEAN COMMISSION

Trends and Developments in Artificial Intelligence

Challenges to the Intellectual Property Rights Framework

Final report

Authors:

The Joint Institute for Innovation Policy:
Christian HARTMANN
Jacqueline E. M. ALLAN

IViR – University of Amsterdam:
P. Bernt HUGENHOLTZ
João P. QUINTAIS
Daniel GERVAIS

Editor:

The Joint Institute for Innovation Policy:
Jacqueline E. M. ALLAN

LEGAL NOTICE

This document has been prepared for the European Commission however it reflects the views only of the authors, and the European Commission is not liable for any consequence stemming from the reuse of this publication. The Commission does not guarantee the accuracy of the data included in this study.

Manuscript completed in September 2020.

First edition

Luxembourg: Publications Office of the European Union, 2020

Print	ISBN 978-92-76-22448-8	doi:10.2759/683128	KK-02-20-789-EN-C
PDF	ISBN 978-92-76-22447-1	doi:10.2759/458120	KK-02-20-789-EN-N

TABLE OF CONTENTS

ABSTRACT

This Report examines copyright and patent protection in Europe for AI-assisted outputs in general and in three priority domains: science (in particular, meteorology), media (journalism), and pharmaceutical research. It comprises an assessment of the state of the art of uses of AI in the three focus areas, and a legal analysis of how IP laws currently apply to AI-assisted creative and innovative outputs.

The Report concludes that the current state of the art in AI does not require or justify immediate substantive changes in copyright and patent law in Europe. The existing concepts of copyright and patent law are sufficiently abstract and flexible to meet the current challenges from AI. In addition, related rights regimes potentially extend to "authorless" AI productions in a variety of sectors, and the sui generis database right may offer protection to AI-produced databases resulting from substantial investment.

However, taking into account the practical implications of AI technologies, the Report identifies specific avenues for future legal reform (if justified by empirical evidence), offers recommendations for improvements in the application of existing rules (e.g. via guidelines), and highlights the need to study the role of alternative IP regimes to protect AI-assisted outputs, such as trade secret protection, unfair competition and contract law.

EXECUTIVE SUMMARY

Introduction

This Report[1] examines the state of the art of copyright and patent protection in Europe for AI-assisted outputs in general and in three priority domains: science (in particular meteorology), media (journalism), and pharmaceutical research. "AI-assisted outputs" are meant as including productions or applications generated by or with the assistance of AI systems, tools or techniques.

As the state of the art reviewed demonstrates, the use of AI systems in the realms of culture, innovation and science has grown spectacularly in recent years and will continue to do so. AI systems have become almost ubiquitous in meteorology and in pharmaceutical research and are making deep inroads into media and journalism. Outside these distinct domains, AI systems are being used to generate diverse literary and artistic content, including translations, poems, scripts, novels, photos, paintings, etc. Likewise, a wide variety of innovative and inventive activity relies on AI systems for its development and deployment, from facial recognition to autonomous driving.

While AI systems have become – and will become – increasingly sophisticated and autonomous, this Report nonetheless assumes that fully autonomous creation or invention by AI does not exist, and will not exist for the foreseeable future. We therefore view AI systems primarily as tools in the hands of human operators. For this reason, we do not enquire whether AI systems should one day be accorded authorship or inventorship status under future IP Law. We also do not examine the IP protection of AI systems *per se*; legal issues concerning the input of protected subject matter into AI systems (e.g. for text-and-data mining); nor algorithmic moderation or enforcement of IP subject matter, as these topics are beyond the scope of this analysis. Analysis of legal protection regimes beyond copyright and patent law (e.g. trade secrets, unfair competition and contract law) is also outside the terms of reference. An important trend that does emerge from the state of the art review is that more and more AI capability is being offered "as a service" rather than as "bespoke" (tailored) AI systems. Consequently, the emphasis of our analysis is on the users (operators) of AI systems, rather than on their developers.

This Report provided an assessment of the state of the art of uses of AI in the three priority domains and a legal analysis of how IP laws currently apply to AI-assisted creative and innovative outputs. The Report concludes with recommendations regarding possible revision of IP law at the European level.

State of the art in uses of AI

There is no universally accepted definition of AI. At a high level, it can be defined as "computer-based systems that are developed to mimic human behaviour" or a "discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence, with limited or no human intervention."

In **pharmaceutical research**, AI is finding patterns within large data sets and helping to automate the search process. Based mostly on machine learning, AI is assisting in disease diagnoses, predictions of drug efficacy and identification of drug characteristics (e.g. toxicity). Neural networks enable compound discovery, personalised medicine and drug repurposing. AI is being applied in finding molecular drug targets (e.g. proteins, nucleic acids) by searching through libraries of candidates, accelerating the high throughput screening needed to find a candidate substance for further investigation in drug development. It is helping in repurposing of drugs to meet new or different need, in polypharmacology (where a disease is due to multiple malfunctions of the body) and to find and accelerate the development of vaccines (by both gene sequencing and simulations of vaccines). In these processes, some measure of human intervention is usually required, either at the start or throughout the entire process, with human feedback optimising the steps.

(1) Authors of SOTA: Christian HARTMANN, Jacqueline ALLAN (Joint Institute for Innovation Policy (JIIP) / Joanneum Research). Authors of Legal analysis: P. Bernt HUGENHOLTZ, João Pedro QUINTAIS, Daniel GERVAIS (IViR – University of Amsterdam). Editor: Jacqueline ALLAN (JIIP).

In recent years, there has been increasing cooperation between the pharmaceutical industry and AI companies. Some companies pursue an active IP strategy and file patents in both the domains of pharmaceutical and AI technology while others sell services confidentially to pharmaceutic companies.

In the area of science, the Report examines **meteorology** as one of the main application areas where AI is already routinely used. Meteorology predicts the state of the atmosphere, at a certain time in a certain place or over a specific area, based on historical data and knowledge of climate and the atmosphere. Automated tasks include post-processing of weather data; predictive analytics for future forecasts; bias correction of meteorological observations; parameterisation of models to correct for radiation, turbulence, cloud microphysics, etc.; data assimilation; and local downscaling of model outputs to improve predictions.

Weather forecasts rely on vast quantities of data. The ability of machine learning to extract knowledge from complex and extensive databases makes it particularly suitable for numerical weather forecasting. Some companies support media companies in weather reporting and forecasting, providing high-precision, precise weather forecasts in multiple formats daily , including recordings, to suit the various reporting media.

In **journalism**, AI enables automated aggregation, production and distribution of content (data, text, images, audio or video). Assistive technologies support journalists in the creation of media content, including speech recognition, information extraction, clustering, summarising, and machine translation capabilities for multi-lingual access to sources. Generative technologies produce media content with human intervention limited to inputting the data set, defining output specifications, and quality control. Distributing technologies encompass the publication or other forms of communication (e.g. through chat bots) of automatically created content with the help of algorithms.

Several companies currently offer technologies for automated content creation for uses in diverse areas including describing product, summarising patient notes in hospitals, reporting on sports events, reporting share prices and customising local information on property markets. Other companies have in-house capabilities for automated news generation. Know-how is commonly protected through licensing models, rather than asserting ownership of IP. It also relies on the tacit knowledge within a company, with the knowledge on how to develop customer-specific systems acting as a high barrier to competitors looking to enter the market.

Legal analysis under EU copyright and patent laws

The legal analysis examines whether, and to what extent, AI-assisted outputs are protected by European copyright law, related rights or patent law. For copyright, the analysis is concentrated on the so-called EU copyright *acquis* and its interpretation by the Court of Justice of the EU (CJEU). The patent analysis concentrates on the European Patent Convention (EPC).

EU copyright law

As our inquiry into EU copyright law reveals, four interrelated criteria are to be met for an AI-assisted output to qualify as a protected “work”: the output is (1) a “production in the literary, scientific or artistic domain”; (2) the product of human intellectual effort; and (3) the result of creative choices that are (4) “expressed” in the output. Whether the first step is established EU law is however uncertain. Since most AI artefacts belong to the “literary, scientific or artistic domain” anyway, and are the result of at least some “human intellectual effort” (however remote), in practice the focus of the copyright analysis is on steps 3 and 4.

Based on a thorough analysis of the CJEU’s case law, and in light of the findings of two experts workshops, we conclude that the core issue is whether the AI-assisted output is the result of human creative choices that are “expressed” in the output. In line with the CJEU’s reasoning in the *Painer* case, we distinguish three distinct phases of the creative process in AI-assisted production: “conception” (design and specifications), “execution” (producing draft versions) and “redaction” (editing, finalisation). While AI systems play a dominant role at the execution phase, the role of human authors at the conception stage often remains essential. Moreover, in many instances, human beings will also be in charge of the redaction stage. Depending on the facts of the case, this will allow human beings sufficient creative choice. Assuming these choices are expressed in the final AI-assisted output, the output will then qualify as a copyright-protected work. By contrast, if an AI system is programmed to automatically execute content without the output being conceived or redacted by a person exercising creative choices, there will be no work.

Due to the “black box” nature of some AI systems, persons in charge of the conception phase will sometimes not be able to precisely predict or explain the outcome of the execution phase. This however need not present an obstacle to the “work” status of the final output, assuming that such output stays within the ambit of the person’s general authorial intent.

Authorship status will be accorded to the person or persons that have creatively contributed to the output. In most cases, this will be the user of the AI system, not the AI system developer, unless collaboration between the developer and user on a specific AI production indicates co-authorship. If “off-the-shelf” AI systems are used to create content, co-authorship claims by AI developers will also be unlikely for commercial reasons, since AI developers will normally not want to burden customers with downstream copyright claims.

A problem that might arise is the possibility of falsely claiming authorship in respect of AI productions that do not qualify as “works” for lack of human creativity. Producers or publishers might be tempted to falsely attribute authorship in such productions in order to benefit from the authorship presumptions granted under EU law, which allow the person whose name is mentioned as an author to initiate infringement procedures.

British and Irish copyright law accord authorship status to persons undertaking the arrangements necessary for creating computer-generated works in cases where no (human) author can be identified. These provisions have been criticised as being incompatible with EU copyright standards, since “authorless” works do not meet the EU standard of “the author’s own intellectual creation”. Perhaps, they are therefore better understood as a species of related rights.

The related rights harmonised under the EU *acquis* offer various possibilities for protecting AI-assisted outputs that do not qualify for copyright protection. In light of the general absence in related rights’ law of a requirement of human authorship or originality, and its rationale of rewarding economic or entrepreneurial activity, related rights will accommodate AI-assisted output in cases of insufficient human creative input.

While AI-assisted outputs in the form of aural signals (audio data) may benefit from the phonographic right, audio-visual outputs will qualify for protection under the film producer’s right. Moreover, AI-assisted broadcasts may find protection under the related rights of broadcasters. None of these related rights provide for a threshold requirement, making these regimes available for AI-assisted outputs that are generated without any creative human involvement – even absent significant economic investment. In most cases the user, not the developer, of the AI system will be deemed the owner of the related right, since it is the user that triggers the acts that give rise to these rights through his use of the AI system and output generation.

Additionally, databases created using AI will qualify for *sui generis* protection under the EU Database Directive if the databases are the result of substantive investment. This includes investment in AI technology and know-how applied in producing the database. In light of the broad legal notion of “database”, the *sui generis* right potentially offers protection to a wide range of AI-assisted productions. However, it is currently uncertain whether investment in machine-generating data – for example, the generation of weather data with the aid of AI – may be factored in. In any case, the prerequisite of a “database” rules out protection of raw data.

Illustrated via case studies in the Report, it is impossible to make general assessments of the copyright status of AI-assisted outputs in individual cases. In some cases, where the creative role of human beings is evident at various stages of the creative process, such as *The Next Rembrandt* project, the output will most likely be copyright protected. In other cases, where it is difficult or even impossible to identify creative choices, such as automatically-generated sports reports or AI-assisted weather forecasts, copyright protection will be less likely. Note however that this is the same for sports reports and weather forecasts produced without machine assistance. Nevertheless, producers of “authorless” AI-assisted outputs might still find recourse in related (neighbouring) rights.

“Authorless” AI-assisted outputs will remain completely unprotected only in cases where no related right or sui generis right is available. Since such rights attach primarily to aural and audio-visual signals, as well as to databases, such cases are most likely to occur if the AI-assisted output is in alphanumerical form. Whether this absence of IP protection might justify regulatory intervention, is primarily an economic question that cannot be addressed in the context of this Report. Such intervention is justified only if no alternative protection (e.g., under trade secret protection, unfair competition or contract law) is available, and solid empirical economic analysis reveals that the absence of protection harms overall economic welfare in the EU.

Our analysis for EU copyright law and AI leads to the following conclusions and recommendations:

- Current EU copyright rules are generally sufficiently flexible to deal with the challenges posed by AI-assisted outputs.
- The absence of (fully) harmonised rules of authorship and copyright ownership has led to divergent solutions in national law of distinct Member States in respect of AI-assisted works, which might justify a harmonisation initiative.
- Further research into the risks of false authorship attributions by publishers of "work-like" but "authorless" AI productions, seen in the light of the general authorship presumption in art. 5 of the Enforcement Directive, should be considered.
- Related rights regimes in the EU potentially extend to "authorless" AI productions in a variety of sectors: audio recording, broadcasting, audivisual recording, and news. In addition, the sui generis database right may offer protection to AI-producted databases that are the result of substantial investment.

The creation/obtaining distinction in the *sui generis* right is a cause of legal uncertainty regarding the status of machine-generated data that could justify revision or clarification of the EU Database Directive.

EU Patent Law

Our analysis of European patent law – and in particular the EPC – demonstrates that the requirement that an inventor be named on a patent application means that one or several human inventors must be identified. Under the EPC regime, this is essentially a formal requirement. The EPO does not resolve disputes regarding substantive entitlement, which is an issue that is governed by national law. Following this approach, the EPO decided two cases in 2020 (currently under appeal) where it considered that, because AI systems do not have legal personality, they cannot be named inventors on a patent application.

A human inventor typically has the right to be named on the application. Beyond this, inventorship and co-ownership are mostly a matter for national law. It should be noted, however, that as AI technology stands today, the possibility that an AI system would invent in a way that is not causally related to one or more human inventors (e.g. the programmer, the trainer, the user, or a combination thereof) seems remote. As technology stands, no immediate action appears to be required on the issue of inventorship at EPC level.

As regards ownership, there are at least three possible (sets of) claimants to an AI-assisted invention: the *programmer or developer* of the AI system; the *owner* of the system; and the *authorised user* of the system (who provided it with training data or otherwise supervised its training). Neither international law nor the EPC provide clear rules on how ownership of patents may be affected by this new type of AI-assisted inventive activity. It is therefore a matter for national laws. However, that might not require harmonisation as there does not seem to be a problem in establishing a sufficient connection between an AI-assisted invention and a patent applicant.

The granting of a patent requires that, as of the date of filing, *the invention must be new (novel) and involve an inventive step*. While the increasing use of AI systems for inventive purposes does not require material changes to these core concepts, it may have practical consequences for patent offices. AI systems enable qualitatively or quantitively different novelty (prior art) searches, and the practical application of inventiveness may change as certain claimed inventions may be "obvious" to a person of ordinary skill in the art ("POSITA") due to the increasing use of AI systems. Any future changes will likely emerge in legal decisions at European (EPO Boards of Appeal) or national levels where patents will either be upheld or not.

A patent application must also sufficiently *disclose* the invention. The "black box" nature of some AI systems may present challenges to this requirement. In that regard, it has been suggested that a mechanism to deposit

AI algorithms be established, akin to that for microorganisms (the Budapest Treaty). Although it is as yet unclear that a deposit system for AI algorithms would be useful, it seems advisable to at least consider the possibility of requiring applicants to provide this type of information, while maintaining sufficient safeguards to protect all confidential information to the extent it is required under EU or international rules.

Finally, inventions that might otherwise be patentable might be protectable as trade secrets under the 2016 Trade Secrets Directive, a topic for future study as it is outside the scope of the current work.

Our analysis for EU patent law and AI leads to the following conclusions and recommendations:

- The EPC is currently suitable to address the challenges posed by AI technologies in the context of AI-assisted inventions or outputs.

- When assessing *novelty*, IPOs and the EPO should consider investing in maintaining a level of technical capability that matches the technology available to sophisticated patent applicants.

- When assessing the *inventive step*, it may be advisable to update the EPO examination guidelines to adjust the definition of the POSITA and secondary indicia as to track developments in AI-assisted inventions or outputs.

- When assessing sufficiency of *disclosure*, it would be useful to study the feasibility and usefulness of a deposit system for AI algorithms and/or training data and models that would require applicants in appropriate cases to provide information that is relevant to meet this legal requirement.

- For the remaining potential challenges identified in this report arising out of AI-assisted inventions or outputs it may be good policy to wait for cases to emerge to identify actual issues that require a regulatory response, if any.

- Further study on the role of alternative IP regimes to protect AI-assisted outputs, such as trade secret protection, unfair competition and contract law, should be encouraged.

RÉSUMÉ ANALYTIQUE

Introduction

Ce rapport[2] étudie l'état des connaissances actuelles en matière de protection de droit d'auteur et des brevets en Europe pour les productions assistées par IA en général et dans trois domaines prioritaires : la science (en particulier la météorologie), les médias (le journalisme) et la recherche pharmaceutique. Par « productions assistées par IA », il convient d'entendre les productions ou applications générées par ou à l'aide de systèmes, d'outils ou de techniques d'IA.

Comme en témoigne l'état de l'art, l'utilisation de systèmes d'IA s'est répandue de manière spectaculaire ces dernières années dans les domaines de la culture, de l'innovation et de la science, et cette évolution n'est pas prête de s'arrêter. Les systèmes d'IA sont devenus presque omniprésents en météorologie et dans la recherche pharmaceutique et progressent considérablement dans les médias et le journalisme. Outre ces différents domaines, les systèmes d'IA sont utilisés pour générer divers contenus littéraires et artistiques, notamment des traductions, des poèmes, des scripts, des romans, des photos, des peintures, etc. De même, une large variété d'activités innovantes et inventives s'appuient sur les systèmes d'IA dans le cadre de leur développement et de leur déploiement et ce, depuis la reconnaissance faciale jusqu'à la conduite autonome.

Si les systèmes d'IA sont de plus en plus en plus perfectionnés et autonomes (et ne cesseront de le devenir), le présent rapport part néanmoins du principe que l'IA ne génère ni ne générera dans un avenir proche de créations ou d'inventions de manière pleinement autonome. Nous considérons donc que les systèmes d'IA sont avant tout des outils confiés aux mains d'opérateurs humains. C'est pourquoi la question de savoir si les systèmes d'IA devront un jour se voir conférer de droit d'auteur ou d'invention en vertu de la future réglementation en matière de propriété intellectuelle ne se pose pas. Nous ne nous penchons pas non plus sur la protection de la propriété intellectuelle des systèmes d'IA en tant que telle, sur les questions juridiques concernant l'introduction d'objets protégés dans des systèmes d'IA (par exemple pour l'exploration de textes et de données), ni sur la modération algorithmique ou l'application d'objets de propriété intellectuelle, car ces objets dépassent le cadre de la présente analyse. L'analyse des régimes de protection juridique au-delà de droit d'auteur et du droit des brevets (par exemple, les secrets d'affaires, la concurrence déloyale et le droit des contrats) ne relève pas non plus du présent rapport. Une tendance importante se dégage de l'examen de l'état des connaissances : de plus en plus de fonctionnalités liées à l'IA sont proposées « comme un service » plutôt que comme des systèmes d'IA « sur mesure » (personnalisés). Par conséquent, nous mettrons l'accent davantage sur les utilisateurs (opérateurs) des systèmes d'IA, plutôt que sur leurs concepteurs.

Le présent rapport évalue l'état des connaissances actuelles concernant les utilisations de l'IA dans les trois domaines prioritaires et propose une analyse juridique de la manière dont la législation en matière de propriété intellectuelle sont actuellement appliquées aux productions créatives et innovantes assistées par IA. Le rapport se termine par des recommandations concernant une possible révision de la législation en matière de propriété intellectuelle au niveau européen.

État des connaissances actuelles en matière d'utilisation de l'IA

Il n'existe pas de définition universellement acceptée de l'IA. Globalement, il faut entendre par « IA » « des systèmes informatiques développés pour imiter le comportement humain » ou une « discipline de l'informatique consistant à mettre au point des machines et des systèmes capables d'exécuter des tâches considérées comme requérant une intelligence humaine, avec une intervention humaine limitée ou nulle ».

(2) Auteurs: "État des connaissances actuelles en matière d'utilisation de l'IA»: Christian HARTMANN, Jacqueline ALLAN (Joint Institute for Innovation Policy (JIIP) / Joanneum Research). Auteurs de l'analyse juridique : P. Bernt HUGENHOLTZ, João Pedro QUINTAIS, Daniel GERVAIS (IViR – University of Amsterdam). Éditeur : Jacqueline ALLAN (JIIP).

Dans la **recherche pharmaceutique**, l'IA trouve des modèles dans de vastes ensembles de données et contribue à l'automatisation du processus de recherche. Basée principalement sur l'apprentissage automatique, l'IA permet de diagnostiquer des maladies, de prédire l'efficacité de certains médicaments et d'identifier leurs caractéristiques (par exemple leur toxicité). Les réseaux neuronaux permettent la découverte de composés, la médecine personnalisée et l'utilisation de médicaments à d'autres fins. L'IA est appliquée à la recherche de cibles moléculaires de médicaments (par exemple, des protéines, des acides nucléiques) par le biais de bibliothèques de candidats, ce qui accélère le criblage à haut débit nécessaire pour trouver une substance candidate à des fins d'investigation plus poussée dans le développement de médicaments. Elle contribue à la réaffectation de médicaments afin de répondre à des besoins nouveaux ou différents, à la polypharmacologie (lorsqu'une maladie est due à de multiples dysfonctionnements de l'organisme) et à la recherche de vaccins et à l'accélération de leur développement (à la fois par le séquençage génétique et la simulation de vaccins). Dans ces processus, un certain degré d'intervention humaine est généralement nécessaire, soit au début du processus soit tout au long de celui-ci, les retours d'expérience humains permettant d'optimiser les étapes.

Ces dernières années, la coopération entre l'industrie pharmaceutique et les sociétés d'IA s'est accrue. Certaines entreprises poursuivent une stratégie active en matière de propriété intellectuelle et déposent des brevets dans les domaines de la technologie pharmaceutique et de l'IA, tandis que d'autres vendent des services en toute confidentialité aux sociétés pharmaceutiques.

Dans le domaine de la science, le rapport considère la **météorologie** comme l'un des principaux domaines d'application courante de l'IA. La météorologie prédit l'état de l'atmosphère, à un moment donné, en un lieu donné ou au-dessus d'une zone spécifique, sur la base de données historiques et de la connaissance du climat et de l'atmosphère. Les tâches automatisées comprennent le post-traitement de données météorologiques, l'analyse prédictive pour les prévisions futures, la correction des observations météorologiques erronées, le paramétrage des modèles pour corriger les radiations, les turbulences, la microphysique des nuages, etc., l'assimilation des données et la réduction d'échelle locale des sorties des modèles pour améliorer les prévisions.

Les prévisions météorologiques reposent sur d'importantes quantités de données. La capacité de l'apprentissage automatique à extraire des connaissances de vastes bases de données complexes la rend particulièrement adaptée à la prévision météorologique numérique. Certaines entreprises aident les médias à établir des rapports et des prévisions météorologiques en fournissant quotidiennement des prévisions météorologiques précises et de haute qualité dans de nombreux formats, y compris des enregistrements, afin de mieux s'adapter aux différentes couvertures médiatiques.

Au niveau du **journalisme**, l'IA permet l'agrégation, la production et la distribution automatisées de contenus (données, texte, images, audio ou vidéo). Les technologies d'assistance aident les journalistes à créer du contenu médiatique, notamment par la reconnaissance vocale, l'extraction d'informations, le regroupement, le résumé et la traduction automatique, afin d'offrir un accès multilingue aux sources. Les technologies génératives produisent du contenu médiatique grâce à une intervention humaine limitée à la saisie de l'ensemble des données, à la définition des spécifications de sortie et au contrôle de la qualité. Les technologies de distribution englobent la publication ou d'autres formes de communication (par le biais de robots de discussion (*chatbots*), par exemple) ou de contenus créés automatiquement à l'aide d'algorithmes.

Plusieurs entreprises proposent actuellement des technologies de création automatisée de contenus pour des utilisations dans divers domaines, notamment la description de produits, le résumé de dossiers de patients dans les hôpitaux, le compte rendu d'événements sportifs, la communication du cours des actions et la personnalisation d'informations locales sur les marchés immobiliers. D'autres entreprises disposent de capacités internes pour la génération automatisée d'actualités. Le savoir-faire est généralement protégé par des modèles de licence, plutôt que par le fait de faire valoir une appartenance sur les droits de propriété intellectuelle. Il repose également sur la connaissance tacite au sein d'une entreprise, les connaissances relatives à la conception de systèmes spécifiques aux clients constituant un obstacle important pour les concurrents qui cherchent à pénétrer le marché.

Analyse juridique dans le cadre de la législation européenne sur le droit d'auteur et le droit des brevets

L'analyse juridique examine si, et dans quelle mesure, les produits issus de l'IA sont protégés par la législation européenne en matière de droit d'auteur, de droits voisins ou de droit des brevets. Pour le droit d'auteur, l'analyse se concentre sur l'acquis communautaire en matière de droit d'auteur et son interprétation par la Cour de justice de l'UE (CJUE). L'analyse des brevets se concentre sur la Convention sur le brevet européen (CBE).

Législation européenne en matière de droit d'auteur

Comme le révèle notre étude de la législation européenne en matière de droit d'auteur dans le contexte de l'IA, quatre critères interdépendants doivent être remplis pour qu'une production assistée par IA soit considérée comme une « œuvre » protégée : la production est (1) une « production dans le domaine littéraire, scientifique ou artistique » ; (2) le produit d'un effort intellectuel humain ; et (3) le résultat de choix créatifs qui sont (4) « exprimés » dans la production. Il n'est toutefois pas certain que la première étape constitue une jurisprudence établie au niveau européen. Étant donné que la plupart des objets d'IA relèvent du « domaine littéraire, scientifique ou artistique » et sont le résultat d'au moins un certain « effort intellectuel humain » (même lointain), ce sont les étapes 3 et 4 qui nécessitent une analyse plus approfondie dans le cadre de la législation sur le droit d'auteur.

Sur la base d'une analyse approfondie de la jurisprudence de la CJUE et à la lumière des conclusions de deux ateliers d'experts, la question essentielle est de savoir si la production assistée par IA est le résultat de choix créatifs humains qui sont « exprimés » dans la production. Conformément au raisonnement de la CJUE dans l'affaire *Painer*, nous distinguons trois phases du processus créatif dans la production assistée par IA : la « conception » (conception et spécifications), l'« exécution » (production de versions préliminaires) et la « rédaction » (édition, finalisation). Si les systèmes d'IA jouent un rôle dominant dans la phase d'exécution, le rôle des auteurs humains au stade de la conception reste souvent essentiel. En outre, dans de nombreux cas de figure, les êtres humains seront également chargés de la phase de rédaction. En fonction des faits, cela permettra aux êtres humains de disposer d'un choix créatif suffisant. En partant du principe que ces choix sont exprimés dans la production finale assistée par IA, la production sera alors considérée comme une œuvre protégée par le droit d'auteur. En revanche, si un système d'IA est programmé pour exécuter automatiquement du contenu, alors que le produit n'est pas conçu ou rédigé par une personne exerçant des choix créatifs, ce contenu ne sera pas considéré comme une oeuvre.

Compte tenu du caractère « boîte noire » de certains systèmes d'IA, les personnes chargées de la phase de conception ne seront parfois pas en mesure de prédire ou d'expliquer avec précision le résultat de la phase d'exécution. Cela ne doit cependant pas empêcher de conférer le statut d'«œuvre » à la production finale, pour autant que cette production s'inscrive dans le cadre d'une intention générale d'auteur.

Le statut d'auteur (paternité) sera conféré à la personne ou aux personnes qui ont contribué de manière créative à la production. Dans la plupart des cas, il s'agira de l'utilisateur du système d'IA, et non de son concepteur, à moins que la collaboration entre le concepteur et l'utilisateur dans le cadre d'une production spécifique d'IA ne témoigne d'une co-paternité. Si des systèmes d'IA « prêts à l'emploi » sont utilisés pour créer du contenu, il est peu probable que les concepteurs de systèmes d'IA revendiquent la co-paternité pour des raisons commerciales, car ils n'ont en principe pas envie d'ennuyer les clients en revendiquant en aval de droit d'auteur.

Un problème qui pourrait se poser est la possibilité de revendiquer à tort la paternité de productions d'IA n'étant pas considérées comme des « œuvres » faute de créativité humaine. Les producteurs ou éditeurs pourraient être tentés d'attribuer à tort la paternité de ces productions afin de bénéficier des présomptions de paternité accordées en vertu du droit européen, qui permettent à la personne dont le nom est mentionné comme auteur d'engager des procédures d'infraction.

Les législations britannique et irlandaise sur le droit d'auteur confèrent le statut d'auteur aux personnes qui prennent les dispositions nécessaires pour créer des œuvres générées par ordinateur dans les cas où aucun auteur (humain) ne peut être identifié. Ces dispositions ont été critiquées comme étant incompatibles avec les normes européennes en matière de droit d'auteur. En effet, les œuvres « sans auteur » ne répondent pas à la norme européenne de « création intellectuelle propre de l'auteur ». Elles sont donc davantage comprises comme une espèce de droits voisins.

Les droits voisins harmonisés dans le cadre de l'acquis européen offrent diverses possibilités de protection des productions assistées par IA ne pouvant bénéficier de la protection de droit d'auteur. Compte tenu de l'absence générale, dans les droits voisins, d'une exigence de paternité ou d'originalité humaine, et de leur raison d'être qui est de récompenser l'activité économique ou entrepreneuriale, les droits voisins permettront de protéger la production assistée par IA dans les cas où l'apport créatif humain est insuffisant.

Alors que les productions assistées par IA sous forme de signaux sonores (données audio) peuvent bénéficier du droit phonographique, les productions audiovisuelles jouiront de la protection du droit du producteur de films. En outre, les émissions assistées par IA peuvent être protégées par le droit voisin des radiodiffuseurs. Aucun de ces droits voisins ne prévoit de seuil, ce qui rend ces régimes disponibles pour les productions assistées par IA générées sans aucune participation humaine créative – même en l'absence d'investissement économique important. Dans la plupart des cas, c'est l'utilisateur, et non le concepteur, du système d'IA qui sera considéré comme le propriétaire du droit voisin, puisque c'est lui qui déclenche les actes qui donnent lieu à ces droits par son utilisation du système d'IA et génère la production.

En outre, les bases de données générées par IA pourront bénéficier d'une protection *sui generis,* en vertu de la directive européenne sur les bases de données, si les bases de données sont le résultat d'un investissement important. Cela inclut l'investissement dans la technologie d'IA et le savoir-faire appliqué à la production de la base de données. À la lumière de la vaste notion juridique de « base de données », le droit *sui generis* offre potentiellement une protection à un large éventail de productions assistées par IA. Toutefois, il n'est pas certain à l'heure actuelle que les investissements dans les données générées par des machines – par exemple, la génération de données météorologiques à l'aide de l'IA – puissent être pris en considération. En tout état de cause, la condition préalable d'une « base de données » exclut la protection des données brutes.

Comme en témoignent les études de cas dans le présent rapport, il est impossible de faire une évaluation générale du statut de droit d'auteur des productions assistées par IA dans des cas individuels. Dans certains cas, lorsque le rôle créatif d'êtres humains est évident à différents stades du processus créatif, comme le projet *The Next Rembrandt*, la production sera très probablement protégée par le droit d'auteur. Dans d'autres cas, lorsqu'il est difficile, voire impossible, d'identifier les choix créatifs, comme des reportages sportifs générés automatiquement ou des prévisions météorologiques assistées par IA, la protection de droit d'auteur sera moins probable. Notez cependant que ce sera la même chose pour les reportages sportifs et les prévisions météorologiques produits sans assistance de machine. Néanmoins, les producteurs de productions assistées par IA « sans auteur » pourraient toujours recourir aux droits voisins.

Les productions assistées par IA « sans auteur » demeureront intégralement dépourvues de protection uniquement dans les cas où aucun droit voisin ou *sui generis* n'est disponible. Étant donné que ces droits s'appliquent principalement aux signaux sonores et audiovisuels, ainsi qu'aux bases de données, de tels cas sont plus susceptibles de se produire si la production assistée par IA est sous forme alphanumérique. La question de savoir si cette absence de protection de la PI pourrait justifier une intervention réglementaire est avant tout une question économique qui ne peut être traitée dans le cadre du présent rapport. Une telle intervention n'est justifiée que si aucune autre protection (par exemple, au titre de la protection des secrets d'affaires, de la concurrence déloyale ou du droit des contrats) n'est disponible et qu'une analyse économique empirique solide révèle que l'absence de protection nuit au bien-être économique général dans l'UE.

Notre analyse de la législation européenne sur le droit d'auteur et l'IA nous amène aux conclusions et recommandations suivantes :

- Les règles actuelles de l'UE en matière de droit d'auteur sont généralement adaptées aux défis posés par les productions assistées par IA.
- L'absence de règles (totalement) harmonisées en matière de paternité et de propriété de droit d'auteur conduit à des solutions divergentes dans les législations nationales des États membres en ce qui concerne les œuvres assistées par IA, ce qui pourrait justifier une initiative d'harmonisation.
- Les risques d'attribution erronée de paternité par des éditeurs de productions d'IA « de type œuvre » mais « sans auteur », devraient faire l'objet de recherches plus approfondies, à la lumière de la présomption générale de paternité de l'article 5 de la directive relative au respect des droits de propriété intellectuelle.

- Les régimes de droits voisins dans l'UE s'étendent potentiellement aux productions d'IA « sans auteur » dans une variété de secteurs : enregistrement audio, radiodiffusion, enregistrement audiovisuel et actualités. En outre, le droit *sui generis* sur les bases de données peut offrir une protection aux bases de données produites par IA qui sont le résultat d'un investissement considérable.

La distinction création / obtention dans le droit *sui generis* est source d'incertitude juridique concernant le statut des données générées par des machines qui pourrait justifier une révision ou une clarification de la directive européenne sur les bases de données.

Droit européen des brevets

Notre analyse du droit européen des brevets – et en particulier de la CBE – démontre que l'exigence de nommer un inventeur sur une demande de brevet signifie qu'un ou plusieurs inventeurs humains doivent être identifiés. Dans le cadre du régime de la CBE, il s'agit essentiellement d'une exigence formelle. L'OEB ne résout pas les litiges concernant la titularité des droits sur le fond, cette question étant régie par le droit national. Suivant cette approche, l'OEB a tranché deux affaires en 2020 (actuellement en appel) dans lesquelles il a considéré que, puisque les systèmes d'IA n'ont pas de personnalité juridique, ils ne peuvent pas être qualifiés d'inventeurs sur une demande de brevet.

Un inventeur humain a normalement le droit d'être nommé sur la demande. Au-delà de cela, la qualité d'inventeur et la copropriété sont principalement du ressort du droit national. Il convient toutefois de noter qu'en l'état actuel de la technologie de l'IA, la possibilité qu'un système d'IA invente une production sans lien de causalité avec un ou plusieurs inventeurs humains (par exemple le programmeur, le formateur, l'utilisateur ou plusieurs de ces personnes) semble faible. En l'état actuel de la technologie, aucune action immédiate ne semble requise en ce qui concerne la paternité de l'invention au niveau de la CBE.

En ce qui concerne la propriété, il existe au moins trois (groupes de) demandeurs possibles pour une invention assistée par IA : le *programmeur ou le concepteur* du système d'IA ; le *propriétaire* du système ; et *l'utilisateur autorisé* du système (qui lui a fourni des données de formation ou a supervisé sa formation de toute autre manière). Ni le droit international ni la CBE ne prévoient de règles claires sur la manière dont la propriété des brevets peut être affectée par ce nouveau type d'activité inventive assistée par IA. Il s'agit donc d'une question qui relève du droit national qui ne nécessite toutefois peut-être pas d'harmonisation, puisqu'il ne semble pas y avoir de problème pour établir un lien suffisant entre une invention assistée par IA et un demandeur de brevet.

La délivrance d'un brevet exige qu'à la date de dépôt, *l'invention soit nouvelle (inédite) et implique une activité inventive*. Bien que l'utilisation croissante des systèmes d'IA à des fins inventives n'exige pas de modifications importantes de ces concepts fondamentaux, elle peut avoir des conséquences pratiques pour les offices des brevets. Les systèmes d'intelligence artificielle permettent des recherches qualitativement ou quantitativement différentes en matière de nouveauté (art antérieur) et l'application pratique de l'activité inventive peut changer car certaines inventions revendiquées peuvent être « évidentes » pour l'homme du métier en raison de l'utilisation croissante des systèmes d'intelligence artificielle.Tout changement futur apparaîtra probablement dans les décisions juridiques au niveau européen (chambres de recours de l'OEB) ou national où les brevets seront maintenus ou non.

Une demande de brevet doit également suffisamment divulguer l'invention. La nature « boîte noire » de certains systèmes d'IA peut complexifier la tâche consistant à remplir cette exigence. À cet égard, il a été suggéré d'établir un mécanisme de dépôt des algorithmes d'IA, semblable à celui qui existe pour les micro-organismes (le traité de Budapest). Bien que l'utilité d'un système de dépôt des algorithmes d'IA ne soit pas encore clairement établie, il semble souhaitable d'envisager au moins la possibilité d'exiger des candidats qu'ils fournissent ce type d'informations, tout en maintenant des garanties suffisantes pour protéger toutes les informations confidentielles dans la mesure où elles sont requises par les règles européennes ou internationales.

Enfin, des inventions qui, autrement, seraient susceptibles d'être brevetables pourraient être protégées au titre de secrets d'affaires en vertu de la directive sur les secrets d'affaires de 2016 – un sujet qui pourrait faire l'objet d'une prochaine étude puisqu'il ne relève pas du présent rapport.

Notre analyse du droit européen des brevets et de l'IA nous amène aux conclusions et recommandations suivantes :

- La CBE est actuellement apte à relever les défis posés par les technologies d'IA dans le contexte des inventions ou productions assistées par IA.

- Pour évaluer le caractère *novateur*, les OPI et l'OEB devraient envisager d'investir dans le maintien d'un niveau de capacité technique correspondant à la technologie disponible pour les demandeurs de brevets sophistiqués .

- Lors de l'évaluation de l'*activité inventive*, il peut être conseillé de mettre à jour les directives d'examen de l'OEB afin d'ajuster la définition de l'« homme du métier » et les indices secondaires pour suivre l'évolution des inventions ou productions assistées par IA.

- Lors de l'évaluation du caractère suffisant de la *divulgation*, il serait utile d'étudier la faisabilité et l'utilité d'un système de dépôt pour les algorithmes d'IA et/ou les données et modèles de formation qui obligerait les demandeurs, dans les cas appropriés, à fournir des informations pertinentes pour satisfaire à cette exigence légale.

- Pour les autres défis potentiels identifiés dans le présent rapport et découlant des inventions ou productions assistées par IA, il peut être judicieux d'attendre que des cas se présentent pour identifier les problèmes réels qui nécessiteraient ou non une réponse réglementaire.

- Il convient d'encourager une étude plus approfondie sur le rôle des autres régimes de propriété intellectuelle pour protéger les productions assistées par IA, tels que la protection des secrets d'affaires, la concurrence déloyale et le droit des contrats.

INTRODUCTION AND CONTEXT

This Report is the final deliverable of the Study "Trends and Developments in Artificial Intelligence (AI) - Challenges to the IPR Framework" (Call for Tender No SMART 2018/0052) undertaken at the request of DG CNECT. The Study focused on "works/inventions which are generated by AI-based systems and which may be subject to… IPR"[3], that is to say, outputs, productions or applications generated by, or with the assistance of, AI systems, tools or techniques, which are susceptible of protection under Intellectual property (IP) law.

The Report integrates a state of the art review (undertaken by JIIP) and a legal study (under the authorship of IViR) and is structured as follows:

- **Section 1** starts by setting out the objectives of the Study and the scope of the legal analyis. It then advances the relevant definitions of AI and associated terminology. It clarifies the concept of "AI-assisted outputs" as the object of the legal analysis and it provides the necessary institutional and policy context.

- **Section 2** provides an overview of the state of the art of AI algorithms used to generate or produce content and knowledge in three priority domains (pharmaceutical research, science and journalism) and to identify major trends in the market for AI-assisted outputs or products in those three priority domains. It also addresses the need to shed light on the role of human intervention in AI-augmented processes in such priority domains.

- **Section 3** contains the core legal analysis for AI and EU *copyright law*. It first addresses the relevant rules in international copyright law (namely the Berne Convention) and provides an overview of the EU copyright *acquis* as it pertains to copyright-protected "works" and issues of authorship. It then turns to the core question: whether, and under which conditions, AI-assisted outputs might qualify as "works" protected under harmonised copyright standards. Next, it discusses the inextricably linked issue of authorship and ownership of AI-assisted outputs. Since authorship and copyright ownership remain mostly unharmonised, divergent solutions have appeared in national law, such as the rules on computer-generated works in the UK and Ireland. The section continues with an examination of how existing EU rules on the protection of related rights and the *sui generis* database right might adequately protect AI-assisted outputs that lack originality or authorship. This is followed by a brief normative discussion of whether it might be justified to extend current related rights to (better) protect "authorless" outputs. The section closes with case studies on the priority domains of science and journalism, where we apply our analytical framework and findings to AI-assisted outputs in weather forecasting and automated journalism.

- **Section 4** provides the core legal analysis for AI and European *patent law*. It begins by offering a brief overview of the applicable patent law framework, including at international level, before settling on the European Patent Convertion (EPC) as the legal frame of reference. The section then addresses the question of inventorship of AI-assisted outputs, including whether human inventorship is a substantive patentability requirement or merely a formal requirement under the EPC. This is followed by a discussion of patent ownership of AI-assisted outputs. Next, we examine different substantive patentability requirements in relation to such outputs, starting with novelty and then inventiveness – in particular, the inventive step, the interpretation of the concept of "person skilled in the art", and the definition of relevant prior art. Building on this analysis, we examine how the EPC rules of sufficiency of disclosure should be interpreted and applied to AI-assisted outputs. This includes consideration of the suitability of a deposit system for algorithms, modelled on that existing for biologic materials. The section closes with case studies on the priority areas of "science" and "pharmaceutical research", where we apply our analytical framework and findings to AI-assisted outputs in the areas of (patent aspects of) weather forecasting and drug discovery.

- Finally, **Section 5** concludes the legal analysis by providing a synthesis of the legal challenges posed by AI-assisted outputs to European IP law, as well as recommendations on how to address these challenges.

(3) EC, 'Trends and Developments in Artificial Intelligence – Challenges to the Intellectual Property Rights Framework', Shaping Europe's digital future, *SMART 2018/0052*, March 2019 ("Call for Tender").

1. Objectives and Scope of the Study

1.1. Objectives

The objectives of the Study were to provide evidence and information in support of policymaking by feeding into the European Commission's work on artificial intelligence (AI) (as defined below) and thereby to lay the groundwork for policy development regarding intellectual property protection of AI-assisted outputs, applications or productions as works or inventions. The precise scope of the legal analysis is detailed below (at 1.1.1).

This Report provides, firstly, an assessment of the state of the art (SOTA) of uses of AI technology in three sectors that have been selected in consultation with the European Commission (EC, or the Commission) and that rely heavily on intellectual property (IP) rights and, secondly, an analysis of whether European IP law (especially at EU level) applies to such uses and, if so, in what way. Desk research, literature review and scoping interviews with both practitioners of IP and professionals with a deep knowledge of AI informed the analysis of the current status of the AI technologies, and their algorithms, for the three selected sectors. Using qualitative interviews and further desk research on recent trends and developments in IP laws that are AI-related, the specific IP issues raised within the three sectors were analysed, and consideration given to what kind of legal protection, if any, the stakeholders in these areas expect to see implemented in the European IP framework. Legal analysis of that evidence examines how IP laws currently apply to AI-assisted outputs, and lead to recommendations regarding possible adaptation of IP law at EU level, and related best practices.

The specific activities of the Study were undertaken on:

- State of the art of AI algorithms in the priority domains: based on data collected through scoping interviews, desk research and literature review, the Study Team presents an overview of the state of the art of AI algorithms autonomously generating content (especially, major market trends) in the three priority domains: pharmaceutical research, science/meteorology and journalism.

- Analysis of potential challenges to the European intellectual property framework: Using a qualitative survey, interviews and desk research, the Study Team identifies and analyses potential challenges AI algorithms may have on the European IP framework, namely EU copyright and European patent law. The work also helps in identifying the specific legal provisions requiring further analysis in the Study.

- Policy recommendations: possible solutions and mitigation: Identification of which legal solutions should be developed to address the challenges identified in the previous step and the development of recommendations for the EC to mitigate and address the challenges. An expert workshop facilitated discussion and validation of the analysis and findings, informing this Report and its conclusions and recommendations.

Through exploration with reference to the three priority domains (pharmaceutical research, science/meteorology and journalism), the ultimate goals of the Study were:

- To assess whether the current IP framework is suitable for AI-assisted outputs, and

- To support the EC in developing targeted policies, with recommendations for possible adaptations in IP legislation and practices.

This Report provides findings with regard to works/inventions generated by AI-based systems which may be subject to intellectual property rights and associated laws. Challenges and impacts on the European IPR framework are identified, and mitigation measures proposed, with reference to the three priority domains.

1.1.1. Specific scope of the legal analysis

This Report focuses on "works/inventions which are generated by AI systems and which may be subject to… IPR"[4], that is to say, outputs, productions or applications generated by, or with the assistance of, AI systems, tools or techniques, which are susceptible of protection under Intellectual Property (IP) law. For the sake of simplicity, such **outputs, productions or application are referred to hereinafter jointly as "AI-assisted outputs"**, while recognising and explaining below some consequences of this choice (at 1.3). The legal analysis examines the current state of the art of IP protection in European law for AI-assisted outputs both for European copyright and patent law in general and in the three priority domains.

To be sure, this is not a wholly new question. As early as the 1960s, and through to the end of the 1990s, scholars had already grappled with questions related to whether computers could be authors or inventors, and the possibility of IP protection for computer-generated works or inventions.[5] The issue has since gained momentum and there is now an increasingly rich body of scholarship on the topic focusing both on copyright[6] and patent[7] law. Still, recent developments in AI technologies, including the rapid roll-out of (nearly or apparently) autonomous AI systems in the creative and innovative industries, have renewed the interest in the topic.[8]

AI systems are used to generate all manner of literary and artistic content, including translations, news articles, poetry, scripts, novels, photos, paintings, etc.[9] Likewise, all manner of innovative and inventive activity relies on AI systems for its development and deployment, from automated drug discovery to autonomous vehicles, to name

---

(4) EC, 'Trends and Developments in Artificial Intelligence – Challenges to the Intellectual Property Rights Framework', Shaping Europe's digital future, SMART 2018/0052, March 2019 ("Call for Tender").

(5) See e.g. F. Fromm, "Der Apparat Als Geistiger Schopfer," GRUR, 1964; Karl F. Jr. Milde, "Can a Computer Be and Author or an Inventor," Journal of the Patent Office Society 51, no. 6 (1969): 378–406; Stephen Hewitt, "Protection of Works Created by the Use of Computers," New L.J. 133 (1983); Timothy L. Butler, "Can a Computer Be an Author - Copyright Aspects of Artificial Intelligence," Hastings Comm. & Ent.L.J. 4 (1982); Pamela Samuelson, "Allocating Ownership Rights in Computer-Generated Works," in Symposium Cosponsored by University of Pittsburgh Law Review and The Software En on The Future of Software Protection (Pittsburgh, Pennsylvania, USA: University of Pittsburgh Press, 1986), 1185–1228; Daniel Gervais, "The Protection under International Copyright Law of Works Created with or by Computers," IIC - International Review of Intellectual Property and Competition Law 5 (1991): 629–60.

(6) For some leading references on copyright law, see: Bruce E. Boyden, "Emergent Works," *The Columbia Journal of Law & the Arts* 39, no. 3 (June 20, 2016): 377–94; Annemarie Bridy, "The Evolution of Authorship: Work Made by Code," *Columbia Journal of Law & the Arts* 39 (2016): 395–401; Annemarie Bridy, "Coding Creativity: Copyright and the Artificially Intelligent Author," *Stanford Law Review* 5, no. 25 (July 18, 2011): 1–28; Carys J. Craig and Ian R. Kerr, "The Death of the AI Author," *Osgoode Legal Studies Research Paper*, March 25, 2019, https://doi.org/10.2139/ssrn.3374951; Jean-Marc Deltorn and Franck Macrez, "Authorship in the Age of Machine Learning and Artificial Intelligence," in *The Oxford Handbook of Music Law and Policy* (Rochester, NY: OUP, 2018); Robert Denicola, "Ex Machina: Copyright Protection for Computer-Generated Works," *Rutgers University Law Review* 69 (2016): 251–87; Daniel J. Gervais, "The Machine As Author," *Iowa Law Review* 105 (March 25, 2019): 19–35; Jane C. Ginsburg and Luke Ali Budiardjo, "Authors and Machines," *Berkeley Technology Law Journal* 34, no. 2 (October 21, 2019): 343–448; James Grimmelmann, "There's No Such Thing as a Computer-Authored Work — And It's a Good Thing, Too," *Columbia Journal of Law & the Arts* 39 (May 4, 2017): 403–16; Andrés Guadamuz, "Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works," *Intellectual Property Quarterly* 2 (June 5, 2017): 169–86; Ana Ramalho, "Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems," *Journal of Internet Law* 21, no. 1 (June 13, 2017): 12–25, h; Sam Ricketson, "The 1992 Horace S. Manges Lecture - People or Machines: The Bern Convention and the Changing Concept of Authorship," *Columbia-VLA Journal of Law & the Arts* 16, no. 1 (1992 1991): 1–38.

(7) For some leading references on patent law, see: Ben Hattenbach and Gavin Snyder, "Rethinking the Mental Steps Doctrine and Other Barriers to Patentability of Artificial Intelligence," *The Columbia Science and Technology Law Review* 19, no. 2 (May 15, 2018): 313–39; Noam Shemtov, "A Study on Inventorship in Inventions Involving AI Activity," Reader in Intellectual Property and Technology Law, Centre for Commercial Law Studies (Queen Mary University of London: EPO, February 2019); Michael McLaughlin, "Computer-Generated Inventions," *American University Washington College of Law*, January 7, 2018, 1–32; Ryan Abbott, "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law," *Boston College Law Review* 57, no. 4 (2016): 1079–1126; Peter Hendrik Blok, "The Inventor's New Tool : Artificial Intelligence : How Does It Fit in the European Patent System?," *European Intellectual Property Review* 39, no. 2 (2017): 69–73; Ben Hattenbach and Joshua Glucoft, "Patens in an Era of Infinite Monekys and Artificial Intelligence," *Stanford Technology Law Review* 19, no. 2 (October 2015): 32–51; Daria Kim, "'AI-Generated Inventions': Time to Get the Record Straight?," *GRUR International* 69, no. 5 (May 1, 2020): 443–56; Ana Ramalho, "Patentability of AI-Generated Inventions: Is a Reform of the Patent System Needed?," SSRN Scholarly Paper (February 15, 2018), https://doi.org/10.2139/ssrn.3168703; Liza Vertinsky and Todd M. Rice, "Thinking about Thinking Machines: Implications of Machine Inventors for Patent Law," *Boston University Journal of Science & Technology Law* 8, no. 2 (2002): 574–613; Shlomit Yanisky-Ravid and Xiaoqiong (Jackie) Liu, "When Artificial Intelligence Systems Produce Inventions: The 3A Era and an Alternative Model for Patent Law," *Cardoza Law Review* 39 (March 1, 2017): 2215–63.

(8) *See e.g.* in this respect recent editorials in specialised IP reviews: Daniel Gervais, "Is Intellectual Property Law Ready for Artificial Intelligence?," *GRUR International* 69, no. 2 (February 1, 2020): 117–18; Jane C. Ginsburg, "People Not Machines: Authorship and What It Means in the Berne Convention," *IIC - International Review of Intellectual Property and Competition Law* 49, no. 2 (January 29, 2018): 131–35; Gerald Spindler, "Copyright Law and Artificial Intelligence," *IIC - International Review of Intellectual Property and Competition Law* 50, no. 9 (November 1, 2019): 1049–51.

(9) *See, e.g.,* Google Translate, https://translate.google.com/ (translation), Deep L, https://www.deepl.com/en/translator (translation); OpenAI, Musenet, https://openai.com/blog/musenet/ (music generation), OpenAI, Jukebox, https://openai.com/blog/jukebox/ (music generation), Talk to Transformer, https://talktotransformer.com (text generation) The Next Rembrandt, https://www.nextrembrandt.com/ (artistic portrait), Christies, "Is artificial intelligence set to become art's next medium?" (12 December 2018) https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx, OpenAI, Image GPT (17 June 2020), https://openai.com/blog/image-gpt/ (image generation), and Annalee Newitz, "An AI Wrote All of David Hasselhoff's Lines in This Bizarre Short Film," *Ars Technica*, April 25, 2017, https://arstechnica.com/gaming/2017/04/an-ai-wrote-all-of-david-hasselhoffs-lines-in-this-demented-short-film/. (movie script generation). Additional examples of AI outputs susceptive of copyright protection can be found throughout this report

but a few prominent examples.[10] This makes it urgent to find appropriate legal solutions in the field of IP that are consistent with EU policy initiatives and the emerging regulatory framework for AI.[11]

The focus of this Report on AI-assisted outputs excludes from its legal analysis certain aspects at the intersection of AI and IP.[12] In particular, it does not examine issues concerning the protection of AI systems *per se* as copyright or patentable subject matter, the use of protected subject matter as *input* for an AI system (e.g. text-and-data mining and machine-generated data), or the algorithmic moderation or enforcement of IP subject matter.[13]

As regards the relevant legal framework, this Report examines whether, and to what extent, AI outputs are protected by EU copyright law or European patent law, interpreted in light of international law. For copyright law, the analysis is concentrated on the so-called EU copyright *acquis* – predominantly formed by Directives – and its interpretation by the Court of Justice of the EU (CJEU). The most relevant instruments in this respect are the InfoSoc Directive (2001/29/EC), the Database Directive (2001/29/EC) and the Copyright in the Digital Single Market (CDSM) Directive (2019/790).[14] The relevant international legal framework, which mostly impacts our analysis *indirectly*, includes the Berne Convention, the Rome Convention, the TRIPS Agreement, the WIPO Copyright Treaty (WCT), and the WIPO Performances and Phonograms Treaty (WPPT).[15] For patent law, the legal analysis focuses on the European Patent Convention (EPC).[16] The relevant international legal framework includes the Paris Convention, [17] the TRIPS Agreement and the Patent Law Treaty (PLT).[18]

In addition to the general assessment of AI-assisted outputs under European copyright and patent law, the legal analysis zooms in on the aforementioned **three priority domains where the development of AI might pose special challenges to IP**.

The first priority domain is "**pharmaceutical research**", later examined under patent law. As seen in the state of the art review, machine learning applications are increasingly being used in all stages of drug discovery leading to drug development. As these applications become more automated and less "supervised", the question arises whether pharmaceutical products or processes thus discovered might meet the requirements for patentability, whether procedural (e.g. regarding inventorship) or substantive (e.g. relating to inventive step and sufficiency of disclosure), under the rules of the EPC.

The second priority domain is "**science**", with a focus on **meteorological services**, including weather and climate predictions. These services may have implications for copyright law (including the *sui generis* database right) and possibly patent law. For example, as the state of the art review (Section 2) suggests, as deep learning algorithms become capable of producing weather forecasts with limited human intervention, this might have immediate ramifications for copyright protection of the ensuing weather forecasts, since copyright pre-supposes human authorship. Conversely, machine-driven weather forecasts might still be eligible for *sui generis* database protection, if and when "substantial investment" in a "database" can be demonstrated.

(10) See e.g. Nic Fleming, "How Artificial Intelligence Is Changing Drug Discovery," Nature 557, no. 7707 (May 30, 2018): S55–57, https://doi.org/10.1038/d41586-018-05267-x.; European Patent Office et al., "Patents and Self-Driving Vehicles" (EPO, 2018), https://www.epo.org/service-support/publications.html?pubid=177#tab3.

(11) *See later in this Report* and, in particular ,the recent EC, 'White paper On Artificial Intelligence – A European approach to excellence and trust', *COM (2020) 65 final,* 19 February 2020 (summarizing policy initiatives an discussing a future EU regulatory framework).

(12) Although these exclusions are mentioned in the state of the art review in this Report in general terms, they are further specified here to clearly delineate the scope of the legal analysis.

(13) For a literature review covering many of these topics, *see* Maria Iglesias Portela, Sheron Shamuilia, and Amenda Anderberg, "Intellectual Property and Artificial Intelligence - A Literature Review" (Luxembourg: European Commission, December 17, 2019), https://publications.jrc.ec.europa.eu/repository/bitstream/JRC119102/intellectual_property_and_artificial_intelligence_jrc_template_final.pdf.

(14) Respectively: Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (InfoSoc Directive); Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (Database Directive); Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (CDSM Directive).

(15) Berne Convention for the Protection of Literary and Artistic Works, opened for signature Sept. 9, 1886, 828 U.N.T.S. 221 (Berne Convention); International Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organisations (Rome Convention) Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, 1869 U.N.T.S.299 (TRIPS Agreement); WIPO Copyright Treaty, Dec. 20, 1996, 2186 U.N.T.S. 121 (WCT); WIPO Performances and Phonograms Treaty, 1996 O.J. (L 89) (WPPT).

(16) Convention on the Grant of European Patents (European Patent Convention) of 5 October 1973 as revised by the Act revising Article 63 EPC of 17 December 1991 and the Act revising the EPC of 29 November 2000European Patent Convention (EPC).

(17) Paris Convention for the Protection of Industrial Property (as amended on September 28, 1979) (Paris Convention).

(18) Patent Law Treaty, adopted at Geneva on June 1, 2000 (PLT).

The third priority domain is "**media**", with a focus on automated **journalism** and its copyright law implications. As shown in this Report, the role of "automated news" in the newsrooms of today is rapidly increasing, especially in the areas of sports reporting and financial information. This raises copyright issues similar to those triggered by AI-driven weather forecasting. In both cases, copyright protection may depend on the remaining role of human actors, if any. Consider, however, that the diminishing role of human authorship in content production may be compensated by an increasing role for human authors at the levels of content conception and content redaction (both pre- and post-production).

As to **methodology**, the legal analysis is predominantly legal doctrinal and descriptive, as this is a *scoping study*. The work was informed by data collection and extensive desk research, case law analysis and literature review for the relevant legal areas and fields covered therein. This includes, in particular, the drawing on studies, analysis, reports and research on AI, with a focus on interrelations with IP. Where relevant for the purposes of the analysis, non-legal literature on AI in the fields of science, media and pharmaceutical research was examined. In addition to the legal doctrinal analysis, the legal analysis includes empirical elements. In particular, both scoping and in-depth interviews were carried out with experts, and are incorporated in the legal analysis.

1.2. Definitions of Artificial Intelligence

1.2.1. Broad definitions of Artificial Intelligence

For the purposes of the state of the art and legal analysis, the definition of AI provides a starting point to understand the main features of the technology and the degree of human intervention in – or contribution to – AI-assisted outputs.

There is no universally accepted definition of AI.[19] At a high level, AI can be defined as "computer-based systems that are developed to mimic human behaviour"[20] or a "discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence, with limited or no human intervention."[21]

Beyond this, numerous definitions of AI exist. These result not only from the development of the technology over time but also from the different approaches and perspectives taken to define AI across multiple research and policy areas.

Although existing definitions vary significantly, there is consensus that the current state of the art and its (at least) medium-term future are characterised by "narrow AI" – meaning "systems that focus on solving concrete application problems" – as opposed to "general AI" or "super intelligence". The Study Team shares this view.

In the absence of a universally-agreed definition of AI, the Study Team has taken as its reference the definition used by the EC in its 2018 Communication on 'Artificial Intelligence for Europe'[22].

(19) Ryan Calo, "Artificial Intelligence Policy: A Primer and Roadmap" 51 (2017): 399–435. p. 404. See also WIPO, "WIPO Technology Trends 2019: Artificial Intelligence." (World Intellectual Property Organization, 2019), https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5982426.. For a reference work on AI, see S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach", Prentice Hall, 3rd edition (2009). For introductions to AI and Machine learning, see Wolfgang Ertel, Introduction to Artificial Intelligence, Undergraduate Topics in Computer Science (Cham: Springer International Publishing, 2017), https://doi.org/10.1007/978-3-319-58487-4. ; and Miroslav Kubat, An Introduction to Machine Learning (Springer International Publishing, 2015), https://doi.org/10.1007/978-3-319-20010-1.

(20) Josef Drexl et al., "Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, October 8, 2019), https://doi.org/10.2139/ssrn.3465577.

(21) WIPO, "Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence" (WIPO, May 21, 2020), https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=499504. p. 3.

(22) Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe. Brussels, 25.4.2018, COM(2018) 237 final, p.1. https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-237-F1-EN-MAIN-PART-1.PDF

> "Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications). We are using AI on a daily basis, e.g. to translate languages, generate subtitles in videos or to block email spam. Many AI technologies require data to improve their performance. Once they perform well, they can help improve and automate decision making in the same domain. For example, an AI system will be trained and then used to spot cyber-attacks on the basis of data from the concerned network or system".

The Study Team has also taken into consideration the ongoing discussions on AI definitions. For example, subsequent to that 2018 Communication, a High-Level Group on AI[23] (AI-HLEG) was established that proposed an update to the Commission definition as follows[24]:

> Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.
>
> As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimisation), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems)."

The AI-HLEG definition is supported by a relevant statement to the effect that "[h]umans design AI systems directly, but they may also use AI techniques to optimise their design".[25] Beyond emphasising the central role of humans in designing and optimising an AI system, the AI-HLEG definition makes the important distinction between **AI as *a system*, on the one hand, and AI as *a scientific discipline***, on the other. The distinction is useful for its explanatory power, as detailed below.

In 2020, the EC's Joint Research Centre (JRC) proposed a definition in its Technical Report[26] 'AI Watch Defining Artificial Intelligence: Towards an operational definition and taxonomy of artificial intelligence'. The starting point for the JRC's operational definition was the above definition proposed by AI-HLEG. The JRC definition consists of a taxonomy and a list of keywords that characterise the core domains of the AI research field, and covers horizontal topics such as the applications of AI research, as well as ethical and philosophical considerations. The JRC definition was to be adopted in the context of AI Watch, the Commission knowledge service monitoring the development, uptake and impact of AI for Europe.

Also in 2020, the Commission released a Communication[27] 'On Artificial Intelligence - A European approach to excellence and trust' (see text box below). That Communication does not define AI in detail, describing it as 'a collection of technologies that combine data, algorithms and computing power'. Instead, it focusses on setting out policy options on how to promote the uptake of AI while addressing the risks associated with some of its applications.

Finally, the 2020 EC White Paper on AI[28], after advancing a generic definition of AI as "a collection of technologies that combine data, algorithms and computing power" [29], adds some consequential comments to be considered in future AI definitions used at the legislative level. First, it emphasises the need for the right balance between

(23) https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence
(24) https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341
(25) AI-HLEG, 6 (& fn.3).
(26) http://dx.doi.org/10.2760/382730; AI watch: Defining Artificial Intelligence : towards an operational definition and taxonomy of artificial intelligence
(27) https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
(28) EC, 'White paper On Artificial Intelligence – A European approach to excellence and trust', COM (2020) 65 final, 19 February 2020
(29) Id. p.2.

flexibility and clarity of future definitions, so as meet the twin aims of accommodating technological development and ensuring legal certainty. Second, it identifies "data" and "algorithms" as the two main elements that characterise AI for the purposes of "future discussions on policy initiatives". In explaining these elements *vis-à-vis* machine learning, the White Paper makes an important reference to human intervention in AI systems:

> *In case of machine learning techniques, which constitute a subset of AI, algorithms are trained to infer certain patterns based on a set of data in order to determine the actions needed to achieve a given goal. Algorithms may continue to learn when in use.* ***While AI-based products can act autonomously by perceiving their environment and without following a pre-determined set of instructions, their behaviour is largely defined and constrained by its developers. Humans determine and programme the goals, which an AI system should optimise for****."* [30]

Building on these definitions, in particular the AI-HLEG distinction between AI as a system and as a scientific discipline, we briefly explain below the concept of AI, clarify the associated terminology relevant for this report, and identify the mains points of human intervention or contribution in the context of an AI system. We build upon these explanations in the subsequent legal analysis of the IP protection of AI-assisted outputs.

Excerpts from the European Commission Communication[31] 'On Artificial Intelligence - A European approach to excellence and trust' (2020)

The Communication also discusses that "a key issue for the future specific regulatory framework on AI ..... is to determine the scope of its application. The working assumption is that the regulatory framework would apply to products and services relying on AI. AI should therefore be clearly defined for the purposes of this White Paper, as well as any possible future policy-making initiative"..... "In any new legal instrument, the definition of AI will need to be sufficiently flexible to accommodate technical progress while being precise enough to provide the necessary legal certainty". In terms of applications, the Communication notes that, for policy, it is "important to clarify the main elements that compose AI, which are "data" and "algorithms". AI can be integrated in hardware. In case of machine learning techniques, which constitute a subset of AI, algorithms are trained to infer certain patterns based on a set of data in order to determine the actions needed to achieve a given goal. Algorithms may continue to learn when in use. While AI-based products can act autonomously by perceiving their environment and without following a pre-determined set of instructions, their behaviour is largely defined and constrained by its developers. Humans determine and programme the goals, which an AI system should optimise for."

1.2.2. AI as a system

As a *system*, AI is mostly software, which sometimes may be embedded in or combined with hardware, as in the case of robotics (or "embodied AI").[32] In simple terms and at a conceptual level, an AI system is comprised by inputs, the operational logic (with a model or models), and outputs. At the current state of technology, AI systems are "goal-directed". This means that humans specify the objective to be achieved and determine the techniques to do so. Still, as we note below, different techniques have different degrees of autonomy in determining the paths to achieve their goal(s).[33]

(30) Id. p.16 (our emphasis).

(31) https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

(32) See OECD, Artificial Intelligence in Society (Paris: OECD, 2019), chap. 1, https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm., defining "AI System": "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. It does so by using machine and/or human-based inputs to: i) perceive real and/or virtual environments; ii) abstract such perceptions into models through analysis in an automated manner (e.g. with ML, or manually); and iii) use model inference to formulate options for information or action. AI systems are designed to operate with varying levels of autonomy." See also AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 1, 4., noting also that "usually AI systems are embedded as components of larger systems, rather than stand-alone systems" (cit. p.1).

(33) AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines."

At the core of an AI system is the "operational logic". This consists of a "model", i.e. a representation of all or part of the system's external environment that describes the environment's structure and/or dynamics.[34] In the AI-HLEG definition, this is called the "reasoning/information processing module".[35]

Depending on the approach followed in developing the model, its different components will contain a higher or lesser degree of human intervention, for instance in setting objectives and performance measures.[36] For example, the architecture or structure of a model is "usually established by a programmer prior to the training process"[37] and the underlying algorithms are also (at least partially) written by humans.[38]

The AI operational logic is developed on the basis of input in the form of data collected by "sensors" and processed by a machine and/or a human.[39] This input data can be structured or unstructured.[40] It includes any type of information and materials, including copyright-protected works or related subject matter, such as "texts, photographs, musical compositions and the like".[41]

The input data is then processed by the operational logic (or processing module) in accordance with the defined objectives or goals, resulting in a certain output, such as "predictions, recommendations and decisions", which may then interact with virtual and real environments through "actuators".[42]

Many outputs resulting from the most popular type of AI systems today – based on neural networks and deep learning techniques – are of such complexity that is not possible for humans to explain how they are generated.[43] Such scenarios for which AI systems lack the property of "explainability" are often referred to as "black box AI".[44]

1.2.3. Artificial Intelligence – technologies and specific terms

An AI system can be of different types or follow different approaches.[45] One common distinction is between *symbolic AI*, on the one hand, and *statistical or sub-symbolic AI*, on the other[46]. The figure below provides an Euler diagram of relevant AI approaches and techniques, placing them in the context of statistical and symbolic concepts, strands or approaches.[47] Hybrid combinations of both approaches in one AI system also exist.

(34) OECD, Artificial Intelligence in Society, chap. 1. See also Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019., defining a model as "an algorithm based upon a (nonlinear) mathematical function that generates output based on the patterns learned from the training data in the training process."

(35) AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 2.

(36) OECD, Artificial Intelligence in Society, chap. 1.

(37) Josef Drexl et al., "Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective," Max Planck Institute for Innovation & Competition 11, no. 3 (October 8, 2019), https://doi.org/10.2139/ssrn.3465577.

(38) Drexl et al., defining algorithm as: "a step-by-step instruction. In the machine learning context, an algorithm is an instruction coded as software and directed at a computer."

(39) OECD, Artificial Intelligence in Society, chap. 1. See also AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 2., noting that in practice sensors "could be cameras, microphones, a keyboard, a website, or other input devices, as well as sensors of physical quantities (e.g. temperature, pressure, distance, force/torque, tactile sensors)".

(40) See AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines." :"Structured data is data that is organised according to pre-defined models (such as in a relational database), while unstructured data does not have a known organisation (such as in an image or a piece of text)." On the training process in a machine learning system, see, e.g., Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019, 7.

(41) Theodoros Chiou, "Copyright Lessons on Machine Learning: What Impact on Algorithmic Art?," JIPITEC 10, no. 3 (February 21, 2020), https://www.jipitec.eu/issues/jipitec-10-3-2019/5025. (para 3) discussing the topic in the context of machine learning and identifying these "training works" as corresponding "to the data set used as training material."

(42) OECD, Artificial Intelligence in Society, chap. 1. See also AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 3. (on "actuation" and "actuators").

(43) Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019, 9.

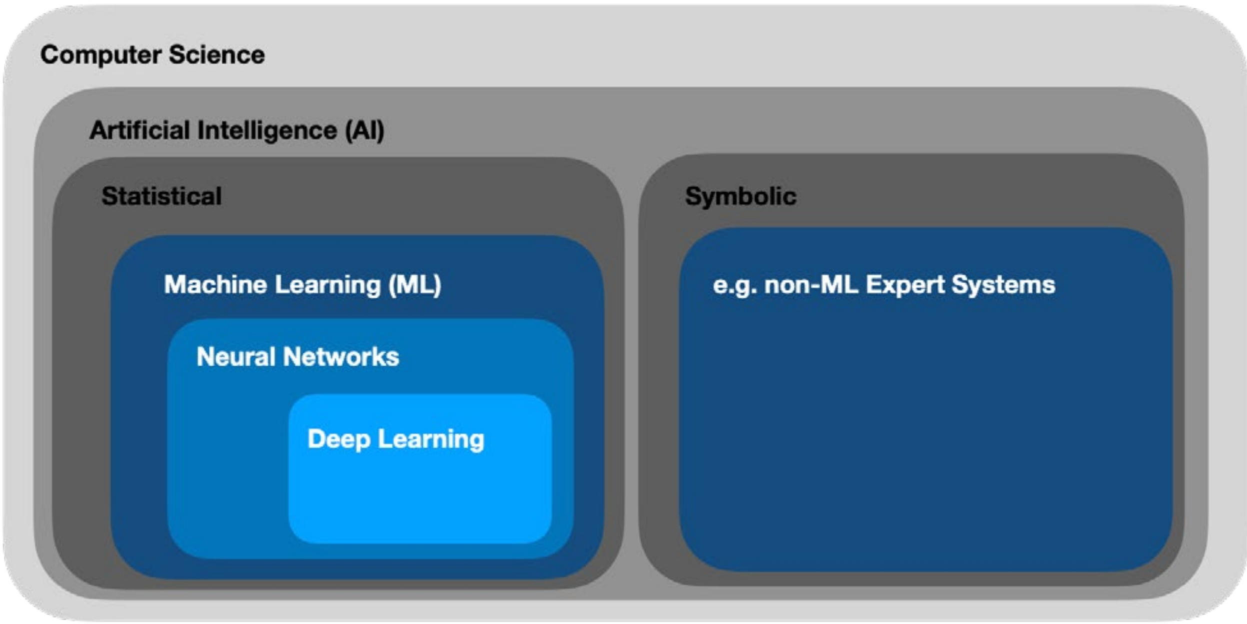
(44) AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 5. See also: Frank Pasquale, The Black Box Society, 2016, https://www.hup.harvard.edu/catalog.php?isbn=9780674970847. (generally on the concept of black box in relation to algorithms); Sandra Wachter, Brent Mittelstadt, and Chris Russell, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, October 6, 2017), https://doi.org/10.2139/ssrn.3063289. (on the concept of black-box and explainability int the context of the GDPR).

(45) On this point, the AI-HLEG definition states "AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions". See AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines.".

(46) See, for example, OECD, Artificial Intelligence in Society, chap. 1. (AI Research).

(47) This figure is based on OECD, Artificial Intelligence in Society, chap. 1. ("Figure 1.6. The relationship between AI and ML") and Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019, 3. (unlabelled figure on section "Machine Learning and Evolutionary Algorithms as Subfields of Artificial Intelligence").

Figure 1.1 AI approaches and techniques



Symbolic AI is also often referred to as "good old-fashioned AI". It covers systems relying on "the organisation of abstract symbols using logical rules"[48] to "build detailed and human-understandable decision structures to translate real-world complexity and help machines arrive at human-like decisions".[49] Many of the techniques within the symbolic AI category are also referred to as "reasoning" techniques.[50]

The failure of symbolic AI to deliver on the promise of machine reasoning – to which some attribute a period of lack of investment in the 1970's[51] sometimes referred to as the "AI winter"[52] – led to a shift in research to focus on specific problems with the assistance of data and statistical representations. Statistical AI refers to a set of techniques of this type that "induce trends from a set of patterns".[53] This group of techniques is also referred to as "learning". It includes machine learning, neural networks, deep learning, decision trees, and many other sorts of learning techniques.[54] The advantage of these techniques over their symbolic counterparts is that they can better deal with tasks that are challenging to define, or that describe something in a detailed fashion "comprehensively… by symbolic behavioural rules", especially in interpreting unstructured data.[55] As a rule, symbolic AI systems reflect a higher degree of human contribution or intervention – including a closer connection to an output – than statistical AI systems. For the purposes of this Report, it is important to offer additional detail on the concepts of machine learning, deep learning, and natural language processing.

Machine Learning

Machine Learning (ML) normally refers to the branch of AI that is focused on developing systems that learn from data. Rather than being explicitly told how to solve a problem, ML algorithms can create solutions by learning from examples (referred to as "training" the ML algorithm). Often, the terms ML and AI are used interchangeably, and their meaning has certainly changed over the last two decades. From a more recent perspective, ML has grown to encompass data-driven approaches, including traditional computational statistics models.[56] In modern

(48) Calo, "Artificial Intelligence Policy: A Primer and Roadmap," 404–5.
(49) OECD, Artificial Intelligence in Society., exemplifying with "optimisation and planning tools."
(50) AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 3.
(51) Rockwell Anyoha, "The History of Artificial Intelligence," Science in the News (blog), August 28, 2017, http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/.
(52) See, e.g.. Calo, "Artificial Intelligence Policy: A Primer and Roadmap"; OECD, Artificial Intelligence in Society.
(53) OECD, Artificial Intelligence in Society.
(54) See AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 3.: "These techniques allow an AI system to learn how to solve problems that cannot be precisely specified, or whose solution method cannot be described by symbolic reasoning rules. Examples of such problems are those that have to do with perception capabilities such as speech and language understanding, as well as computer vision or behaviour prediction ".
(55) AI-HLEG, 3–4.
(56) Examples include polynomial regression and logistic classification.

terminology, the term AI is used to describe “deeper” models, which have the ability to learn (almost) arbitrarily complex mappings from input to outcome. Such models include deep neural networks and Gaussian processes. ML can be viewed as an extension of AI , augmenting models that learn from example with approaches such as expert systems, logical and statistical inference methods, and planning[57].

ML can be categorised as **supervised, unsupervised and reinforcement learning**. In the supervised learning sub-area, “learning” is the ability of an AI system to reproduce regularities or patterns. Results or facts that are known through laws of nature or expert knowledge are used to teach the system. A learning algorithm attempts to find a hypothesis that makes accurate predictions, a hypothesis being a figure that assigns an assumed output value to each input value. The method is therefore based on a pre-determined output to be learned, the results of which are known. The results of the learning process can then be compared with the known, correct results, i.e. “monitored”.[58]

Unsupervised learning refers to ML without known targets. The (learning) machine tries to detect patterns in the input data, patterns that deviate from any associated structureless data noise. An artificial neural network orients itself based on the similarity of the data to the input values, and adapts the weighting accordingly. Different things can be learned. Popular forms are the automatic segmentation (clustering) or the compression of data for dimension reduction.

Reinforcement learning is a set of ML methods in which an agent independently learns a strategy to maximise the rewards received. The agent is not shown which action is the best in which situation, but rather receives a ‘reward’ at certain points in time, which can also be a negative ‘reward’. Using these rewards, the agent approximates a utility function that describes the value of a certain state or action. The term ‘reinforcement learning’ is borrowed from psychology and has been used since the beginnings of cybernetics. Marvin Minsky, for example, used the term in his dissertation of 1954.

The models of reinforcing learning attempt to imitate learning behaviour in nature. In the spectrum of “goal-directed AI”, reinforced learning techniques allow the highest level of freedom for the system (in relation to human developers or users) in defining the path to achieve its goal.[59]

At present, ML is the most widely used or dominant subfield or branch of AI.[60] It also appears to be the case that most AI-assisted outputs susceptible of protection by copyright or patent law result from ML systems, or a subfield thereof, such as neural networks, deep neural networks and deep learning techniques.

(Deep) Neural Networks and Deep Learning

Neural networks are a type of model used in AI operational logic.[61] The structure of the model “consists of layers of neurons connected by weights” and “imitates the functioning of a human brain.”[62]

Work on neural networks began in the 1950s, progressing significantly in the 1980’s and 1990’s. Despite initial investigations of deep neural networks in the 1990’s, high-performance computing of the time did not allow for training using large datasets, in realistic time periods, for well over a decade. It is only more recently that the impressive ability of deep learning to solve certain classes of problem has been seen[63].

(57) OECD (2018). Artificial intelligence and machine learning in science, 5.

(58) https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

(59) For a basic description and explanation of “reinforcement learning”, see Wolfgang Ertel, “Reinforcement Learning,” in Introduction to Artificial Intelligence, ed. Wolfgang Ertel, Undergraduate Topics in Computer Science (Cham: Springer International Publishing, 2017), 289–311, https://doi.org/10.1007/978-3-319-58487-4_10.

(60) [EC, Artificial Intelligence for Europe, COM(2018)237 final, p.10]. See also WIPO, “WIPO Technology Trends 2019,” 14.: “Machine learning is the dominant AI technique disclosed in patents and is included in more than one-third of all identified inventions (134,777 patent documents). Filings of machine learning-related patent have grown annually…”.

(61) For a basic description and explanation of “neural networks”, see Wolfgang Ertel, “Neural Networks,” in Introduction to Artificial Intelligence, ed. Wolfgang Ertel, Undergraduate Topics in Computer Science (Cham: Springer International Publishing, 2017), 245–87, https://doi.org/10.1007/978-3-319-58487-4_9.

(62) Drexl et al., “Technical Aspects of Artificial Intelligence,” October 8, 2019, 5, 12.

(63) Ibid.

In simple terms, neural networks "modify their own code to find and optimise links between inputs and outputs".[64] The AI-HLEG definition paper provides a good explanation:

> *A neural network has, as input, the data coming from the sensors (in our example, the picture of the floor) and, as output, the interpretation of the picture (in our example, whether the floor is clean or not). During the analysis of the examples (the network's training phase), the connections' weights are adjusted to match as much as possible what the available examples say (that is, to minimise the error between the expected output and the output computed by the network). At the end of the training phase, a testing phase of the behaviour of the neural network over examples never seen before checks that the task has been learnt well.*[65]

Both deep learning (DL) and (deep) neural networks are a type of ML. In common with other "non-parametric" methods, DL does not specify the functional form of solutions. Instead, it has enough flexible complexity to learn arbitrary mappings, from input to outcome, from many training examples[66]. Recently, DL has transformed the way in which algorithms achieve (or exceed) human-level performance in areas such as game playing and computer vision. DL owes its success to the easy availability of vast amounts of data and powerful computers, as well as novel algorithmic insights.

Natural Language Processing

According to becominghuman.ai[67], Natural Language Processing (NLP) can be defined as '*the application of computational techniques to the analyses and synthesis of natural language and speech*'. Natural Language Processing, usually shortened as NLP, is a branch of AI that deals with the interaction between computers and humans using natural language (i.e. a human language rather than a computer language). The ultimate objective of NLP is to read, decipher, understand, and make sense of human language in a manner that is, in some way, valuable. Most NLP techniques rely on machine learning to derive meaning from human language[68].

1.3. Focus of the analysis: "AI-assisted" outputs

In this report, we use the term "**AI-assisted outputs**" as a catch-all term to mean outputs, applications or productions generated by or with the assistance of AI systems, tools or techniques. This terminology allows us to engage with existing scholarship and policy materials that often use different terms with varying degrees of precision to mean a wide range of AI outputs or applications susceptible of IP protection.

In the context of IP, some scholars make a distinction between **"direct outputs" and "applications"** resulting from an AI system, arguing that they give rise to "distinct legal issues".[69] Direct outputs may refer to predictions, recommendations (e.g., for a "daily mix" playlist on Spotify), correlations, clustering, etc. "Applications" are outputs that can "be put into practical use", such as in the area of IP subject matter.[70] Still, in the area of IP this distinction is not always easy to draw and may, in some cases, be artificial.

(64) OECD, Artificial Intelligence in Society, chap. 1.
(65) AI-HLEG, "A Definition of AI: Main Capabilities and Disciplines," 4.
(66) Ibid.
(67) https://becominghuman.ai
(68) https://becominghuman.ai/a-simple-introduction-to-natural-language-processing-ea66a1747b32
(69) Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019, 9. NB that this is a distinct use than AI "application fields" (i.e. "different fields, areas or disciplines where AI techniques or functional applications may find application, such as transportation, agriculture or life and medical sciences.") and AI "functional applications" (e.g. speech and computer vision), as used in WIPO, "WIPO Technology Trends 2019," 13, 24.
(70) Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019, 9.

In the field of copyright, AI applications include, to name but a few, paintings (such as "The Next Rembrandt"[71] or "The Portrait of Edmond Belamy"[72]); text translations produced by the GP-T2[73] or (more recently) the GP-T3 text generator from Open AI[74] or the DeepL translator[75]; and musical compositions such as the "emotional soundtrack music" composed by AIVA[76] or the tracks produced for the "AI Song Contest".[77] AI applications relevant for patent law might include, for example, technology for "the functioning of a self-driving car, optimisation of a car design, development of medical treatments, virtual assistants".[78]

An important point is that **AI applications thus defined will always require human input**. What varies is the "required degree of human contribution" for each application."[79] This view is consistent with technical descriptions of the "lifecycle" of AI systems, which identify a number of human interventions and contributions at different stages: planning and designing (e.g. defining requirements and objectives, prototyping), collecting and processing data (e.g. developing a dataset), and modelling (e.g. development of models or algorithms); "verification and validation" (e.g. performance tests and calibration); "deployment" (e.g. piloting and compatibility with external software); and "operation and monitoring" (e.g. assessment and correction of outputs for consistency with objectives).[80]

Despite our use of the term "AI-assisted outputs", we recognise (in line with the examples above) that many of the outputs under consideration in this Report fall under the heading of "applications" or "productions". They denote some level of human contribution at different stages of the AI system lifecycle, often between the generation of an output and its publication, or otherwise making available.

This does not mean that such intervention or contribution is always material to the output and its protection under IP law. Rather, it means that **the focus of the legal inquiry is on assessing the level of human contribution to (or intervention in) the lifecycle of an AI system, and the extent to which it determines IP protection of an AI-assisted output**. The hypothetical scenario of fully autonomously generated AI output – often discussed in pop culture, science fiction and some IP scholarship and policy – is not the focus of our analysis.[81]

As explained in the state of the art section of this Report (Section 2), an increasing number of large IT companies and start-ups nowadays offer "artificial intelligence as a service" (AIaaS), which offer access to powerful cloud-based AI capabilities. AIaaS allow users to produce a full range of AI outputs, without the users having to develop the AI technology in-house.[82] For IP law, the use of AIaaS necessarily implies a shift of the focus of the legal analysis away from the developer of the AI technology to the user (business or individual) that employs the AI system to generate a specific output.

(71) The Next Rembrandt, https://www.nextrembrandt.com/. See also Microsoft News, "Blurring the Lines Between Art, Technology and Emotion: The Next Rembrandt" (13 April 2016), https://news.microsoft.com/europe/features/next-rembrandt/

(72) See Christies, "Is artificial intelligence set to become art's next medium?" (12 December 2018) https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx ("Portrait of Edmond Belamy (detail) created by GAN (Generative Adversarial Network)…";" This portrait, however, is not the product of a human mind. It was created by an artificial intelligence, an algorithm defined by that algebraic formula with its many parentheses. And when it went under the hammer in the Prints & Multiples sale at Christie's on 23-25 October, Portrait of Edmond Belamy sold for an incredible $432,500, signalling the arrival of AI art on the world auction stage.")

(73) See Open AI, GPT-2: 1.5B Release (5 November 2019), https://openai.com/blog/gpt-2-1-5b-release/ . See also James Vincent, "OpenAI Has Published the Text-Generating AI It Said Was Too Dangerous to Share," The Verge (blog), November 7, 2019, https://www.theverge.com/2019/11/7/20953040/openai-text-generation-ai-gpt-2-full-model-release-1-5b-parameters.

(74) ON GPT-3, see: Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. "Language Models Are Few-Shot Learners." ArXiv:2005.14165 [Cs], July 22, 2020. http://arxiv.org/abs/2005.14165; and OpenAI. "OpenAI Licenses GPT-3 Technology to Microsoft." OpenAI (blog), September 22, 2020. https://openai.com/blog/openai-licenses-gpt-3-technology-to-microsoft/. For criticism, see Marcus, Gary, and Ernest Davis. "GPT-3, Bloviator: OpenAI's Language Generator Has No Idea What It's Talking about." MIT Technology Review. Accessed September 28, 2020. https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/.

(75) Deep L, https://www.deepl.com/en/translator.

(76) AIVA, https://www.aiva.ai/

(77) VPRO, AI Song Contest, FAQ, https://www.vprobroadcast.com/titles/ai-songcontest.html

(78) Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019, 9. *See also* Ian Sample, "Powerful Antibiotic Discovered Using Machine Learning for First Time," *The Guardian*, February 20, 2020, sec. Society, https://www.theguardian.com/society/2020/feb/20/antibiotic-that-kills-drug-resistant-bacteria-discovered-through-ai; Fleming, "How Artificial Intelligence Is Changing Drug Discovery"; European Patent Office et al., "Patents and Self-Driving Vehicles."

(79) Drexl et al., "Technical Aspects of Artificial Intelligence," October 8, 2019, 9.

(80) In this paragraph we rely on the technical description of the "AI System Lifecycle" in OECD, *Artificial Intelligence in Society*, chap. 1.

(81) *See e.g.* WIPO, "Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence," 4., which appears to define the terms "AI-generated" and "generated autonomously by AI" in such a way: "terms that are used interchangeably and refer to the generation of an output by AI without human intervention. In this scenario, AI can change its behaviour during operation to respond to unanticipated information or events.", *See also* Kim, "'AI-Generated Inventions.'", with a discussion on the lack of suitability of the term AI-generated as compared to AI-aided.

(82) *See* Daniel Newman, "Why AI As A Service Will Take Off In 2020," *Forbes*, accessed July 19, 2020, https://www.forbes.com/sites/danielnewman/2020/01/07/why-ai-as-a-service-will-take-off-in-2020/.

1.4. Policy and Institutional Context

This section briefly examines recent legal policy developments in the EU and abroad on the intersection of IP and AI, with a focus on the protection of AI-assisted outputs by copyright and patent law. The aim is to provide the necessary policy context to the legal analysis by setting out relevant initiatives chronologically.

1.4.1. EU Institutional Level: the Commission and the European Parliament

At the **EU institutional level**, there has been significant and wide-ranging policy work on the regulation of AI, namely from the European Parliament and EC, including their associated research services.[83]

Most policy initiatives and research do not directly touch IP protection of AI-assisted outputs. This is well illustrated by the work of the EC, which in the last years has published an increasing number of Communications on the topic of AI. These include, to name the most prominent, communications on "Artificial Intelligence for Europe" [84], a "Coordinated Plan on Artificial Intelligence" [85], "Building Trust in Human-Centric Artificial Intelligence"[86], the "European Strategy for Data"[87], and the "White Paper on AI: a European approach to excellence and trust".[88] As mentioned, the EC has also set up the AI-HLEG, which has developed important groundwork in this area.[89] Still, these communications either do not discuss IP at all or contain only minor references to it, none of which is directly relevant to the issue of IP protection of AI-assisted outputs.[90] To be sure, the 2018 Communication on "Artificial Intelligence for Europe" calls for reflection on the interactions between AI and IP rights "from the perspective of both intellectual property offices and users, with a view to fostering innovation and legal certainty in a balanced way". It adds that the use of "AI to create works can have implications on intellectual property, with questions arising for instance on patentability, copyright and right ownership".[91] Beyond that, the most relevant initiatives on IP protection of AI-assisted outputs are the Call for Tender from which originated this Study[92] and a subsequent Call for a study on "Copyright Data Management and Artificial Intelligence".[93]

By contrast, the European Parliament (EP) has produced policy documents relevant to this Study. In February 2017, a Resolution with recommendations on "Civil Law Rules on Robotics" called on the EC to support a horizontal and technologically neutral approach to IP that could be applied to the various sectors in which robotics could be employed.[94] More directly linked to our work here is an Explanatory Statement on "Civil Law Rules on Robotics" by

(83) See, for a recent overview, Jędrzej Niklas and Lina Dencik, "European Artificial Intelligence Policy: Mapping the Institutional Landscape," Working Paper DATAJUSTICE project (DATAJUSTICE, 2020). For a gateway to the multiple EC initiatives in this area, European Commission, Strategy, Shaping Europe's digital future, Policies, "Artificial intelligence", https://ec.europa.eu/digital-single-market/en/artificial-intelligence. Regarding the work of research services associated with the EP and EC – e.g. the EPRS and the JRC – such work is not detailed in this section but is referred to throughout the text, where directly relevant. Among the most relevant recent work in this respect, see: Iglesias Portela, Shamuilia, and Anderberg, "Intellectual Property and Artificial Intelligence - A Literature Review."; Ronan Hamon, Henrik Junklewitz, and Ignacio Sanchez Martin Jose, "Robustness and Explainability of Artificial Intelligence," EUR - Scientific and Technical Research Reports (Publications Office of the European Union, 2020), JRC119336, https://publications.jrc.ec.europa.eu/repository/handle/111111111/58835.; Delipetrev et al., "AI Watch." WIPO, Impact of Artificial Intelligence on IP Policy: Call for Comments, https://www.wipo.int/about-ip/en/artificial_intelligence/call_for_comments/.

(84) EC, 'Artificial Intelligence for Europe', Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2018) 237 Final, 25 April 2018.

(85) EC, 'Coordinated plan on Artificial Intelligence', Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2018) 795 Final, 7 December 2018.

(86) EC, 'Building Trust in Human-Centric Artificial Intelligence', Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2019) 168 final, 8 April 2019.

(87) EC, European strategy for data, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Study Guidelines 2020, COM (2020) 66 final, 19 February 2020.

(88) EC, 'White paper On Artificial Intelligence – A European approach to excellence and trust', COM (2020) 65 final, 19 February 2020.

(89) On which see European Commission, Strategy, Shaping Europe's digital future, Policies, "High Level Expert Group on Artificial Intelligence", https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence .

(90) The Communication on "An European Strategy for Data" does however discuss possible legislative intervention in the area of IP: ""evaluating the IPR framework with a view to further enhance data access and use (including a possible revision of the Database Directive41and a possible clarification of the application of the Trade Secrets Protection Directive42as an enabling framework)"". See EC, European strategy for data, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Study Guidelines 2020, COM (2020) 66 final, 19 February 2020, p. 13

(91) EC, 'Artificial Intelligence for Europe', Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2018) 237 Final, 25 April 2018. p. 15 (fn. 52)

(92) EC, 'Trends and Developments in Artificial Intelligence – Challenges to the Intellectual Property Rights Framework', Shaping Europe's digital future, *SMART 2018/0052,* March 2019 ("Call for Tender").

(93) EC, CNECT/2020/OP/0009 — Study on Copyright and New technologies: Copyright Data Management and Artificial Intelligence — Smart 2019/0038.

(94) EP, Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), paras 136 – 137 (Interoperability, access to code and intellectual property rights).

the EP's Committee on Legal Affairs (JURI). The Statement proposes (inter alia) "the elaboration of criteria for 'own intellectual creation' for copyrightable works produced by computers or robots".[95] Two years later, in February 2019, the EP passed a Resolution on a "comprehensive European industrial policy on artificial intelligence and robotics". This resolution renewed the earlier call to the EC to "support a horizontal and technologically neutral approach to intellectual property applicable to the various sectors in which robotics could be employed" and underscored the "need to monitor the relevance and efficiency of rules on intellectual property rights to govern the development of AI."[96] Finally, on 20 October 2020, the European Parliament adopted a resolution on "intellectual property rights for the development of artificial intelligence technologies".[97] Regarding AI-assisted creations, the EP resolution

> *"[t]akes the view that technical creations generated by AI technology must be protected under the IPR legal framework in order to encourage investment in this form of creation and improve legal certainty for citizens, businesses and, since they are among the main users of AI technologies for the time being, inventors; considers that works autonomously produced by artificial agents and robots might not be eligible for copyright protection, in order to observe the principle of originality, which is linked to a natural person, and since the concept of 'intellectual creation' addresses the author's personality; calls on the Commission to support a horizontal, evidence-based and technologically neutral approach to common, uniform copyright provisions applicable to AI-generated works in the Union, if it is considered that such works could be eligible for copyright protection; recommends that ownership of rights, if any, should only be assigned to natural or legal persons that created the work lawfully and only if authorisation has been granted by the copyright holder if copyright-protected material is being used, unless copyright exceptions or limitations apply; stresses the importance of facilitating access to data and data sharing, open standards and open source technology, while encouraging investment and boosting innovation".*[98]

1.4.2. The WIPO, EPO and National IPOs

At the international institutional level, WIPO[99] is carrying out significant work on the intersection of IP and AI.[100] In addition to its flagship publication of technology trends in AI[101] and ongoing dialogues with national Intellectual Property Offices (IPOs)[102], WIPO launched, in December 2019, a "Call for Comments on the Impact of AI for IP Policy". The underlying draft issues paper[103] identifies a number of questions that are relevant to our legal analysis.[104] More than 250 submissions were received by WIPO follwoing this call for comments[105], leading to

(95) EP JURI, 'Explanatory Statement', *European Parliament* 2017, # par. Intellectual property rights, data protection and data ownership (elaborating also on the on the need to come forward 'with a balanced approach to intellectual property rights when applied to hardware and software standards and codes that protect innovation and at the same time foster innovation').

(96) EP, 'Resolution on a comprehensive European industrial policy on artificial intelligence and robotics', *(2018/2088 (INI)),* 12 February 2019.

(97) European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)), P9_TA-PROV(2020)0277, available at https://www.europarl.europa.eu/doceo/document/TA-9-2020-0277_EN.html.. See also the preceding DRAFT REPORT on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)), Committee on Legal Affairs of the EU Parliament, (Rapporteur: Stéphane Séjourné), 24.4.2020, https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.html?redirect.

(98) European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI), para. 15.

(99) World Intellectual Property Organisation https://www.wipo.int/portal/en/index.html

(100) *See* generally WIPO, Artificial Intelligence and Intellectual Property, (WIPO, 2019) .

(101) WIPO, "WIPO Technology Trends 2019.". *See also* WIPO, WIPO Technology Trends – Artificial Intelligence, https://www.wipo.int/tech_trends/en/artificial_intelligence/ ("This first report in our "Technology Trends" series offers evidence-based projections to inform global policymakers on the future of AI. It analyses data in patent applications and scientific publications to better understand the latest trends in the field.")

(102) *See, e.g.,* WIPO, 'WIPO Conversation on Intellectual Property and Artificial Intelligence', *WIPO* 2019, https://www.wipo.int/meetings/en/details.jsp?meeting_id=51767 .

(103) WIPO, 'WIPO Conversation on IP and AI', Draft issues paper on intellectual property policy and artificial intelligence', *WIPO/IP/AI/2/GE/20/1*, 13 December 2019, https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_ge_20/wipo_ip_ai_2_ge_20_1.pdf

(104) Id., on Patent Inventorship and Ownership (*Issue 1, Qs 6 and 7*); on Patentable Subject Matter and Patentability Guidelines (*Issue 2, Q8*); on Inventive Step or Non-Obviousness (*Issue 3, Q9*); on Disclosure (*Issue 4, Q10*); on General Policy Considerations for the Patent System (*Issue 5, Q11*); on Copyright Authorship and Ownership (*Issue 6, Q 12*); on General Policy Issues on Copyright (*Issue 9, Q 16*).

(105) All responses to the call for comments are available at https://www.wipo.int/about-ip/en/artificial_intelligence/policy.html#submissions. *See, e.g.,* the critical submission of the MPI, https://www.ip.mpg.de/fileadmin/ipmpg/content/stellungnahmen/2020-02-11_WIPO_AI_Draft_Issue_Paper__Comments_Max_Planck.pdf.

a "Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence"[106] and a second session of the WIPO Conversation on IP and AI in July 2020.[107]

In Europe, the EPO[108] has also been active in this area.[109] Firstly, since 2018, it has been working in the context of the IP5 on issues at the intersection of AI and IP.[110] The IP5 in 2019 launched a task force on AI and emerging technologies.[111] Secondly, the EPO has commissioned research and surveys on topics at the crossroads of patent law and AI.[112] Thirdly, and perhaps most importantly, in two recent decisions applying Article 81 EPC (and Rule 19E EPC Implementing Regulations) on inventor designation, the EPO has refused two European patent applications designating an AI system as the inventor.[113] The rejected applications (on appeal at the time of writting) where made by the team behind the "Artificial Inventor Project" [114], which has also filed similar applications – with similar results – in the UK and US, as detailed below.[115] Finally, in 2019, the EPO made changes to its Examination Guidelines regarding the technical character of AI inventions. In particular, relevant changes were made to Guidelines G-II, 3.3 – Mathematical methods and G-II, 3.3.1 – Artificial intelligence and machine learning.[116]

In the United States, two developments at the United States Patent and Trademark Office (USPTO) are worth noting. In late 2019, the USPTO announced its request for public comments on IP Protection for AI Innovation[117], which was followed in 2020 by a consultation report.[118] Also in 2020, the USPTO published a petition decision explaining that, under current US law, only natural persons may be named as an inventor in a patent application.[119] The decision, published pursuant to 37 CFR 1.14(e), mirrors those of the EPO and UK IPO in relation to similar applications made by the team behind the "Artificial Inventor" project.

Two important developments in the UK echo the EPO and USPTO interpretations. The UK IPO recently updated its Formalities Manual to clarify that "[a]n 'AI inventor' is not acceptable as this does not identify a 'person' which is required by law", with the result that such a designation would lead to withdrawal of the application.[120] It also issued a decision to the effect that a non-human inventor, such as an AI machine, could not be named as the inventor of the patent. Furthermore, because such a machine lacks legal personality, it cannot own property in

(106) WIPO, "Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence."

(107) WIPO, WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI): Second Session (9 to 7 July 2020) https://www.wipo.int/meetings/en/details.jsp?meeting_id=55309. (recordings available at https://webcast.wipo.int/). A third session is scheduled on November 4, 2020; see WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI): Third Session, WIPO/IP/AI/3/GE/20, https://www.wipo.int/meetings/en/details.jsp?meeting_id=59168.

(108) European Patent Office https://www.epo.org

(109) *See* generally EPO, Artificial Intelligence, https://www.epo.org/news-events/in-focus/ict/artificial-intelligence.html.

(110) *See, e.g.,* FiveIPOffices, Report of the IP5 expert round table on artificial intelligence, EPO, Munich, 31 October 2018, https://www.fiveipoffices.org/material/ai_roundtable_2018_report/ai_roundtable_2018_report . More recently, see FiveIPOffices, First IP5 NET/AI task force meeting takes place in Berlin, Germany (17 January 2020), https://www.fiveipoffices.org/news/20200117 .

(111) EPO, 'World's five largest patent office's agree on joint task force for emerging technologies and AI', *EPO News* 2019, https://www.epo.org/news-events/news/2019/20190613a.html.

(112) *See*, respectively, Shemtov, "A Study on Inventorship in Inventions Involving AI Activity.", and Heli Pihlajamaa, "Legal Aspects of Patenting Inventions Involving Artificial Intelligence Summary of Feedback by EPC Contracting States," Presentation (EPO, 2020).

(113) *See*: summary of decisions at https://www.epo.org/news-events/news/2020/20200128.html. The applications are: EP3564144 (FOOD CONTAINER), https://register.epo.org/application?number=EP18275163#_blank ; and EP3563896 (DEVICES AND METHODS FOR ATTRACTING ENHANCED ATTENTION), https://register.epo.org/application?number=EP18275174#_blank.

(114) The Artificial Inventor Project, http://artificialinventor.com.

(115) For additional information and analysis on the EPO decisions, *see*: Ryan Abbott, "The Artificial Inventor Project," *WIPO Magazine*, December 2019, https://www.wipo.int/wipo_magazine/en/2019/06/article_0002.html.; Heli Pihlajamaa, "Legal Aspects of Patenting Inventions Involving Artificial Intelligence Summary of Feedback by EPC Contracting States," Presentation (EPO, 2020).; Kluwer-Patent-Blogger, "EPO: A Machine Cannot Be an Inventor," *Kluwer Patent Blog* (blog), January 29, 2020, http://patentblog.kluweriplaw.com/2020/01/29/epo-a-machine-cannot-be-an-inventor/..

(116) See: EPO Guidelines G-II, 3.3 – Mathematical methods, https://www.epo.org/law-practice/legal-texts/html/guidelines2018/e/g_ii_3_3.htm; EPO Guidelines G-II, 3.3.1 – Artificial intelligence and machine learning, https://www.epo.org/law-practice/legal-texts/html/guidelines2018/e/g_ii_3_3_1.htm.

(117) USPTO, Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation, Federal Register Docket No. PTO-C-2019-0038 (Document Number 2019-23638) 30 October 2019, https://www.federalregister.gov/documents/2019/10/30/2019-23638/request-for-comments-on-intellectual-property-protection-for-artificial-intelligence-innovation.

(118) USPTO, Public Views on Artificial Intelligence and Intellectual Property Policy (October 2020), https://www.uspto.gov/sites/default/files/documents/USPTO_AI-Report_2020-10-07.pdf.

(119) USPTO, Decision on Petition, Application No 16/524,350, https://www.uspto.gov/sites/default/files/documents/16524350_22apr2020.pdf?utm_campaign=subscriptioncenter&utm_content=&utm_medium=email&utm_name=&utm_source=govdelivery&utm_term=.

(120) UKIPO, Formalities Manual (Online Version), Chapter 3: The Inventor, Sections (3.01 - 3.38) (last updated: October 2019), https://www.gov.uk/guidance/formalities-manual-online-version/chapter-3-the-inventor#ref3-02.

the patent and transfer that ownership to a (human) applicant.[121] This decision was appealed and the appeal was rejected by the England and Wales High Court (Patents Court).[122]

In addition to these institutional developments, it is important to briefly mention some recent developments that speak directly to the topic of IP protection of AI-assisted outputs.

As regards copyright, a recent study of the International Association for the Protection of Intellectual Property (AIPPI) examines the topic "Copyright in artificially generated works" across more than thirty jurisdictions, including fourteen Member States of the EU.[123] The resulting Resolution rejects copyright protection for "AI generated works" absent human intervention, while entertaining the possibility of protection of such output through related rights, although it considers it too early to take a position on this matter.[124] In addition, the influential French *Conseil supérieur de la propriété littéraire et artistique* (CLSPA) in 2020 published a Report on the "Mission Intelligence Artificielle et Culture" (27 January 2020), examining *inter ali*a the protection by copyright of AI-assisted outputs.[125] The report *inter alia* concludes that for the time being current copyright law is sufficiently flexible to deal with AI-assisted creations.[126]

Finally, regarding patent protection of AI-assisted outputs, it is worth mentioning the ongoing AIPPI[127] study on the topic of inventorship in this context[128], as well as a recent position paper by the UK Chartered Institute of Patent Attorneys (CIPA), which calls for "clarity regarding the patenting of innovations created using AI systems".[129]

The next section of this Report provides a state of the art review of AI algorithms and their applications in the three priority domains.

(121) UKIPO, Patent Decision O/741/19, https://www.ipo.gov.uk/p-challenge-decision-results/p-challenge-decision-results-bl?BL_Number=O/741/19.

(122) Thaler v The Comptroller-General of Patents, Designs And Trademarks [2020] EWHC 2412 (Pat), https://www.bailii.org/ew/cases/EWHC/Patents/2020/2412.html.

(123) For a short overview, see Jan Bernd Nordemann, "AIPPI: No Copyright Protection for AI Works without Human Input, but Related Rights Remain," *Kluwer Copyright Blog* (blog), November 21, 2019, http://copyrightblog.kluweriplaw.com/2019/11/21/aippi-no-copyright-protection-for-ai-works-without-human-input-but-related-rights-remain/.. The EU member states covered in the study are: Austria, Belgium, Czech Republic, Denmark, Estonia, France, Finland, Germany, Hungary, Italy, Latvia, The Netherlands, Spain and Sweden.

(124) J.P. Osha et al., "AIPPI Resolution on Copyright in Artificially Generated Works," Resolution (London, 2019).

(125) CLSPA, Alexandra Bensamoun, and Joëlle Farchy, "Mission du CSPLA sur les enjeux juridiques et économiques de l'intelligence artificielle dans les secteurs de la création culturelle" (CLSPA - Conseil Supérieur de la Propriété Littéraire et Artistique, 2020), https://www.culture.gouv.fr/Sites-thematiques/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-les-enjeux-juridiques-et-economiques-de-l-intelligence-artificielle-dans-les-secteurs-de-la-creation-culturelle.

(126) Id., p. 48.

(127) The International Association for the Protection of Intellectual Property https://aippi.org

(128) AIPPI, 'Inventorship of inventions made using Artificial Intelligence', Study Guidelines 2020, https://aippi.info/wp-content/uploads/2020/02/Q272-SGL-P-2020-Inventorship-of-inventions-made-using-AI.pdf.

(129) CIPA, "Patenting Inventions Created Using an AI System: A CIPA Discussion Paper" (CIPA - Chartered Institute of Patent Attorneys, 2020), https://www.cipa.org.uk/_resources/assets/attachment/full/0/260456.pdf; Kluwer-Patent-Blogger, "CIPA Calls for Clarity about Artificial Intelligence and Patents," *Kluwer Patent Blog* (blog), June 5, 2020, http://patentblog.kluweriplaw.com/2020/06/05/cipa-calls-for-clarity-about-artificial-intelligence-and-patents/.

STATE OF THE ART

2. State of the art review of AI algorithms in the three priority domains

2.1. Introduction

2.1.1. Objectives

The objective of this Section[130] is to provide an overview of the state of the art of AI algorithms used to generate or produce content and knowledge in the three priority domains (pharmaceutical research, science/meteorology and journalism) and to identify major trends in the market of AI-assisted outputs or products in those three priority domains. It also addresses the role of human intervention in AI augmented processes in the priority domains.

The Section also:

- considers (in 2.2) AI in general – background and likely future trends;
- presents (in 2.3) the use of AI in the area of pharmaceutical research, concluding with examples including two detailed case studies on the companies IKTOS and BenevolentAI. The case studies explore the company, the technology, the use case(s), business models and IP rights. The section concludes with some future trends in AI use in pharmaceutical research;
- reports (in 2.4) on AI in science, specifically in meteorology, concluding with a case study on Ubimet and future trends;
- identifies (in 2.5) the use of AI in journalism, with a case study on Retresco, and consideration of current developments and future trends.
- provides (in 2.6) conclusions from the overview of the state of the art and the case studies.

2.2. General development and trends in AI

2.2.1. Background

The previous Section identified different definitions of AI and associated terminology. To better understand the development and trends of the technology, here we set out briefly some further historical background relevant to the state of the art review that follows.

AI was proposed in 1956 as a field of research with the vision that, one day, machines should speak, formulate abstract concepts, solve problems whose solution was previously only possible for humans, and be able to improve themselves[131]. As early as 1950, Alan Turing had proposed, in a famous article, what is today called the "Turing Test", i.e. that in a conversation you should be able to determine which of two participants is a human and which is not (i.e. is a 'bot'). Today, the test is more topical than ever, given concerns about the influence of social bots on our opinions including, for example, election results.

(130) The state of the art review was authored by Christian HARTMANN (JIIP-Joanneum Research) and reviewed and commented by the Study Team, final editing being undertaken by JIIP.

(131) OECD (2017). OECD Digital Economy Outlook 2017, OECD Publishing, Paris, 296.

AI has undergone various peaks and troughs in its development. In the first commercially-relevant wave of AI continuing to the 1980's, expert systems and planning systems for narrowly-defined tasks in very structured areas such as chess, mathematical proofs and error diagnosis were developed. These used manually-entered symbolic knowledge of various kinds: class hierarchies with inheritable properties, "if-then" rules, logical formulas and consistency conditions. This allowed for heuristic searches and logical (or even fuzzy) reasoning. Natural language, vision and robotics were researched separately. As noted previously in this Report, the success was moderate, resulting in the so-called "AI winter" and AI became termed as "good old-fashioned AI" (GofAI)[132].

From about 1990 onwards, another new approach, traceable back to Marvin Minsky, emerged in the form of distributed AI. In his book "Society of Mind", he describes the human mind as a kind of society: "intelligence", according to Minsky, is composed of small units that perform primitive tasks and whose interaction generates intelligent behaviour. Minsky called on the AI community to overcome this individualistic impasse and design completely different, socially-inspired algorithms for parallel computers.

An AI milestone in the 1990's was the first victory by an AI chess machine over the World Chess Champion. In 1997, the IBM Deep Blue computer defeated Garry Kasparov, the reigning world chess champion, in an official tournament[133]. This event was considered a historic victory of machine over man. The event caused a worldwide sensation and brought IBM significant attention and reputation. Today, computers are considered unbeatable in chess. However, the victory of the IBM computer was overshadowed by the fact that Deep Blue owed its success less to its artificial, cognitive intelligence, and more to the brute force of being able to generate all conceivable moves.

The current phase of AI developments and application began around 2010 with the beginning of AI commercialisation. AI applications left research laboratories and spread into everyday applications, led by the two most visibly applied AI areas – ML and natural language processing. In addition, neural networks are currently experiencing their second – this time very successful – rebirth.

The main reasons for the commercial turnaround were both improved AI processes and more powerful software and hardware. On the software side, the more advanced neural networks – and especially the DL variant[134] – have proved to be robust and versatile. Other trends – such as multi-core architectures, improved algorithms and super-fast in-memory databases – have made AI applications particularly attractive to the corporate sector. An additional factor is the increasing availability of large amounts of structured and unstructured data from a variety of sources – such as sensors and digitised documents and images – that can be used to "train" learning algorithms.

In parallel with these improved technical and economic possibilities, the large IT corporations also discovered (or became more involved in) AI. In 2011, IBM took an important step forward with its computer, Watson[135]. Watson can understand natural language and answer complex questions very quickly. In 2011, Watson was the winner in the US-American television quiz, *Jeopardy*, defeating two human quiz champions. Subsequently, IBM further developed Watson into a cognitive system, combining algorithms of natural language processing (NLP) and information retrieval, methods of ML, knowledge representation and automatic inference. Since then, Watson has been successfully applied in fields including medicine and finance, and IBM has focused much of its business on Watson.

Other large IT companies also followed the path taken by IBM. Google, Microsoft, Facebook, Amazon and Apple currently invest many millions of euro/dollars in AI and provide both AI applications and services. For example, in January 2016, Google's AlphaGo beat probably the world's best Go player by 4 to 1. Because of the greater complexity of Go compared to chess, the Japanese board game is practically unbreakable when using traditional brute force algorithms such as those still used in Deep Blue. DL and other current AI methods led to the AlphaGo success.

Several aspects differentiate the current situation from previous AI "hype-cycles": the underlying computer technology has improved; there is vastly more data; AI is better understood; and – perhaps most importantly as a point of historical difference – the amount of corporate money being invested has increased, with large profits being made from using AI technologies.[136]

(132) Ibid.
(133) https://www.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/
(134) https://nl.mathworks.com/discovery/deep-learning.html; https://developer.ibm.com/technologies/deep-learning/
(135) https://www.ibm.com/ibm/history/ibm100/us/en/icons/watson/
(136) OECD (2018). Artificial intelligence and machine learning in science, 122.

At present, AI solutions offer, above all, the possibility to more strongly automate processes. This applies, for example, to machine image recognition in the medical field or in industrial production. In the past, a visual inspection of a product was necessar: now, sensors and algorithms can be used. Also, taking an example in the field of machine processing of natural language, communication can be automated, for example, written communication with customers through "chatbots". In the short term, these advances will reduce costs, optimise processes and reduce waiting times. In addition, Intelligent Personal Assistants have already become part of the everyday life of many people.

Weak AI or artificial narrow intelligence (ANI) refers to systems that focus on solving concrete application problems. The solution of the problem is based on methods of mathematics and computer science, which are developed and optimised for the respective requirements. The resulting system is able to optimise itself. ANI systems function reactively on a superficial intelligence level and do not achieve a deeper understanding of the problem solution. ANI is mainly focused on the fulfilment of clearly defined tasks and does not vary its approach to problems. Instead, weak AI relies on the methods that are made available to it for problem solving. Even the most developed currently available AI systems (e.g. IBMs Watson, Google's AlphaGo) only apply ANI[137].

The goal of "strong AI", also called "super intelligence" or "artificial general intelligence" (AGI), is to achieve or exceed the intellectual abilities of humans. Strong AI does not only act reactively, but also on its own initiative, acting both intelligently and flexibly[138]. As of today, it has not yet been possible to develop strong AI. The discussion continues about whether the development of such an intelligence is even possible. However, the majority of researchers now agree that strong AI will be developed, without consensus on *when* this will happen[139].

2.2.2. Future Trends

Artificial Intelligence as a Service (AIaaS)

AI as a Service (AIaaS), i.e. resources and development tools based on SaaS, is increasingly in demand. The analysts at Allied Market Research forecast the global market for AIaaS to be around USD 77 billion by 2025. By comparison, in 2017 the market was still estimated at just under USD 2.4 billion. This corresponds to an average annual growth rate of over 50 percent[140].

There are four main different forms of AIaaS:

1. Tailor-made platforms and frameworks for ML, which can create data models and "derive" patterns from existing data pools (e.g. the Spaya platform for finding new routes for drug synthesis[141]);

2. AI-based bots, based on the ever-improving NLP (e.g. Microsoft Azure Bot Service[142]).

3. Fully managed ML services that use drag-and-drop tools, cognitive analysis and custom data models. The goal for these is to generate more value, i.e. a greater benefit compared to general ML frameworks (e.g. Bigml's Comprehensive Machine Learning Platform[143], Dataiku DSS[144]).

4. The fourth type of AIaaS includes third-party Application Programming Interfaces (APIs) designed to add functionality to any new/existing application (e.g. craft ai[145]).

AIaaS enables everyone to use AI regardless of their level of knowledge. Simple APIs are available for developers and, by virtue of graphical user interfaces (GUIs) with detailed instructions for users without programming skills

(137) OECD (2017). OECD Digital Economy Outlook 2017, OECD Publishing, Paris, 300.
(138) Ibid.
(139) Goertzel, B. and Pennachin, C. (2006). Artificial General Intelligence, Springer, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-540-68677-4.
(140) https://www.alliedmarketresearch.com/artificial-intelligence-as-a-service-aiaas-market
(141) For more details see page 35
(142) https://azure.microsoft.com/de-de/services/bot-service/
(143) https://bigml.com/features
(144) https://www.dataiku.com/product/
(145) https://www.craft.ai/

a data processing pipeline can be made to "click" together – simply the so called "click-and-drop". The algorithms are often a 'black box', in that commercial AIaaS providers do not publish their embedding of algorithms and clients just have to trust in their effectiveness.

A next wave in AI development?

A recent study by the European Parliamentary Research Service for the Panel for the Future of Science and Technology (STOA) provided a speculative discussion of possible paths to strong AI and addressed theoretical future trends in AI development[146], topics that are considered below.

Self-explanatory and *contextual AI* could offer completely new functions and, while maintaining the performance of neural networks, could provide the human accessibility and explainability of expert systems. It could combine ML with more comprehensive contextual knowledge of the world, achieving better results with less training data. For example, a handwriting recognition system would, on the one hand, train on images of written text and, on the other, benefit from contextual knowledge about how people use their hands and pens.

Quantum computers use technology based on the principles of quantum theory, which explains the nature of energy and matter at the atomic and subatomic levels and is based on the existence of quantum mechanical phenomena such as superposition and entanglement. According to the superposition principle (as illustrated in the well-known 1930's thought experiment "Schrödinger's cat"), quantum systems can exist simultaneously in several states until they are observed or measured.

Based on this technology, and while some companies have already developed the first quantum computers, their overall development is still in its infancy. If high quality quantum computers become reality in the future, they would be superior to any supercomputer currently running. If well-developed quantum computers were to be available in the near future, the development of AI would be certainly strongly impacted. The dramatic increase of computer power could lead to both a dramatic increase of effectiveness of, for example, deep neural networks, and completely new approaches in AI.

The next part of the Report considers the three priority domains, beginning with pharmaceutical research.

2.3. Pharmaceutical research

2.3.1. Introduction

Advances in AI, especially in its sub-field of ML, have empowered a wide variety of innovative business applications.[147]. Based on its abiilty to find hidden patterns within large data sets and to automate many predictions[148], AI has been promoted as the means to accelerate drug development[149].

Developing drugs is perhaps one of the most expensive processes in the world, costing over USD 2 billion for a typical drug[150]. It is also risky: 90 percent of drugs fail to obtain the approval of the European Medicines Agency (EMA)[151]. As the pharmaceutical industry is extremely competitive, with no single firm occupying a market share greater than 6% of the industry[152], it is critical for pharmaceutical firms to continuously develop novel drugs, despite the high costs, in order to stay competitive[153]. Although scientists have made considerable progress in identifying the underlying patterns and mechanisms of how a biological target could cause disease, the drug development process remains slow[154]. The human biological system is extremely complex, with 25,000 genes and millions of

(146) Boucher, Ph. (2020), Artificial intelligence: How does it work, why does it matter, and what can we do about it? Scientific Foresight Unit (STOA)PE 641.547 – June 2020
(147) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 2.
(148) Agrawal, A., Gans, J., Goldfarb, A. (2018). Prediction Machines: The simple economics of artificial intelligence. Harvard Business Press.
(149) Fleming, N. (2018). How artificial intelligence is changing drug discovery. Nature. 557(7707) p55.
(150) Bowen, L., Wu, L. (2020), Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 6.
(151) Ibid.
(152) Ibid.
(153) Ibid.
(154) Ibid.

proteins, all of which can create complex interactions with each other[155]. Difficulty in managing this complexity is a key factor in the high failure rate of developing drugs[156]. Furthermore, there may be a huge cost in meeting regulatory requirements – to demonstrate safety and efficacy – especially since only a small fraction of the drugs pass through the final stage of drug discovery despite showing promise in earlier stages[157].

The potential of AI to identify novel drugs that human researchers alone cannot detect has attracted investment from both start-ups and established pharmaceutical companies[158]. For example, Atomwise, a biomedical start-up that pioneered DL models to optimise drug designs, has claimed a substantial shortening of the process of discovering new chemical compounds. It was able to screen over 8 million compounds and predict which compound could address a particular disease target[159]. Another firm, BenevolentAI, has developed a DL platform that can generate a hypothesis about a drug candidate by analysing academic publications, patents, and clinical trials. It can then validate the hypothesis to some extent by predicting the interplay between chemical/ biological entities, such as genes, proteins and drug candidates[160].

2.3.2. Types of AI used in pharmaceutical research

AI used in pharmaceutical research is mostly based on ML, which can take three forms – supervised, unsupervised and reinforcement learning.[161] Outputs from *supervised* ML in pharmaceutical research include disease diagnosis, prediction of drug efficacy and ADMET[162] prediction[163], while *unsupervised* learning can help in disease subtype discovery from clustering, and disease target discovery from feature-finding methods[164]. *Reinforcement learning* can support *de novo* and experimental drug design[165].

DL can help to discover new compounds (i.e. that can potentially be new drugs), to identify or repurpose drugs that could be more potent (when used individually or in combination) and to improve the area of personalised medicine based on genetic markers[166].

2.3.3. The drug discovery process

Drug development typically has three main stages. The first is the discovery and pre-clinical trials stage (that often involves animal testing) where drug candidates are proposed to address certain biological targets that are responsible for a disease[167]. The next main stage involves three phases (Phase I/ Phase II/Phase III) of human clinical trials. Lastly, if the drug succeeds in these trials, the European Medicines Agency may grant its approval[168].

While the failure rate at each of the stages is high, the reasons for failure differ[169]. During the discovery and preclinical trial stage, the bottleneck lies in screening drug candidates that can recognise and modify the disease targets to achieve therapeutic effects[170]. This process is labour- and data-intensive because the amount of searching needed to find the drug candidate-target pair (the ‘search space’) is large and requires the understanding of

(155) Pisano, G.P. (2006). Science business: The promise, the reality, and the future of biotech. Harvard Business Press.
(156) Dougherty, D., Dunne D.D. (2012). Digital science and knowledge boundaries in complex innovation. Organisation Science. 23(5) 1467-1484.
(157) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 2.
(158) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 2.
(159) Wallach, I., Dzamba, M., Heifets, A. (2015). AtomNet: A deep convolutional neural network for bioactivity prediction in structure-based drug discovery. arXiv preprint arXiv:1510.02855.
(160) Fleming, N. (2018). How artificial intelligence is changing drug discovery. Nature. 557(7707) p55.
(161) See supra at 1.2.3
(162) Initialism of absorption, distribution, metabolism, excretion, toxicity: A set of test categories used together in drug discovery to provide insight into how a pharmaceutical drug interacts with the body as a whole.
(163) Guncar, G. et al. (2018). An application of machine learning to haematological diagnosis. Sci. Rep. 8, 411
(164) Young, J.D. et al. (2017). Unsupervised deep learning reveals prognostically relevant subtypes of glioblastoma. BMC Bioinf. 18, 381
(165) Chen, H. et al. (2018). The rise of deep learning in drug discovery. Drug Discov. Today 23, 1241–1250
(166) Mak K.K., Pichika, P.P. (2019). Artificial intelligence in drug development: present status and future prospects, Drug Discovery Today, Volume 24, Number 3 .March 2019, 773.
(167) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 7
(168) Ibid.
(169) Ibid.
(170) Gashaw, I., Ellinghaus, P., Sommer, A., Asadullah, K. (2011). What makes a good drug target? Drug Discovery Today. 16(23-24) 1037-1043; Hughes JP, Rees S, Kalindjian SB, Philpott KL (2011). Principles of early drug discovery. British Journal of Pharmacology. 162(6) 1239-1249; Keserű GM, Makara GM (2006). Hit discovery and hit-to-lead approaches. Drug Discovery Today. 11(15-16) 741-748.

complex interactions within the human biological system. It can range from genetics, to protein synthesis, from biological and chemical synthetic processes to disease mechanisms[171].

Once a drug candidate/target pair is found and verified during the pre-clinical trial phase, it enters the clinical trials stage where the focus shifts more toward testing for drug safety and efficacy in people, complying with the European Medicines Agency regulations, communicating with patients, and ensuring the integrity of the process.

Each phase of a clinical trial has different sample sizes, safety and efficacy requirements, and different patient profiles require the tailoring of dosages. These efforts all require substantial interactions with the trial participants to understand their health conditions, treatment regimens, and the potential side effects. Drug developers need to continually monitor the patients and communicate evidence of drug safety and efficacy to the European Medicines Agency[172]. This development process can cost billions of dollars for a single drug and can take five to fifteen years to bring a new drug to market[173]. Accordingly, the largest share of domestic expenditure on research and development is spent by the pharmaceutical sector[174].

2.3.4. How AI is used in drug discovery

The use of AI in the search for new chemical compounds began decades ago and forms the basis for cheminformatics and bioinformatics studies[175]. Earlier approaches often relied on expert systems and data mining techniques to construct chemical structures in *in silico* representations (i.e. using computer modelling or simulation)[176]. However, these approaches have not really succeeded in improving the drug discovery process. Only the recent paradigm shift in AI from expert systems to ML is in the process of fundamentally changing the manner of drug discovery[177].

There are two key drivers for the recent AI advances that could significantly accelerate drug discovery:

- First, the digitisation of existing scientific knowledge has enabled the discovery of new compounds that could address certain disease targets. For example, digitising the human genome has drastically increased the search space for new chemical compounds[178]. Rich and multimodal digital data (e.g. chemical structures, X-ray imaging) have further enlarged the search space. However, the size and modality of the data also increased the complexity of navigating in the search space[179].

- Second, recent advances in AI – specifically in the field of machine learning – can help in managing the complexity of navigating the vast search space for potential drugs[180]. AI can automatically and intelligently collect, digest, analyse and detect complex patterns in the existing data on chemical compounds and biological reactions in the human body. ML algorithms can generate a large feature space to find hidden linkages, a process that is extremely labour-intensive for medical experts to perform. They can also search systematically to predict whether new drug compounds can treat a target and generate novel hypotheses faster than scientists can[181] Accordingly, by overcoming the barriers in computation and data management, AI technologies can reduce the search costs to find drug-target pairs even when the search space has

(171) Dougherty, D., Dunne, D.D. (2012). Digital science and knowledge boundaries in complex innovation. Organisation Science. 23(5) 1467-1484; Vamathevan J, Clark D, Czodrowski P, Dunham I, Ferran E, Lee G, Li B, Madabhushi A, Shah P, Spitzer M (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery 1.

(172) Sacks, L.V., Shamsuddin, H.H., Yasinskaya, Y.I., Bouri, K., Lanthier, M.L., Sherman, R.E. (2014). Scientific and regulatory reasons for delay and denial of FDA approval of initial applications for new drugs, 2000-2012. Jama. 311(4) 378-384.

(173) Harrer, S., Shah, P., Antony, B., Hu, J. (2019). Artificial Intelligence for Clinical Trial Design. Trends in Pharmacological Sciences; Hughes, J.P., Rees, S, Kalindjian, S.B., Philpott, K.L. (2011). Principles of early drug discovery. British Journal of Pharmacology. 162(6) 1239-1249.

(174) Kinch, M. (2016). A Prescription for Change: The Looming Crisis in Drug Development. UNC Press Books.

(175) Agrawal, A., Gans, J., Goldfarb, A. (2018). Prediction Machines: The simple economics of artificial intelligence. Harvard Business Press

(176) Brown, F.K .(1998). Chemoinformatics: what is it and how does it impact drug discovery. Annual Reports in Medicinal Chemistry. 33 375-384.

(177) Zhavoronkov, A., Ivanenkov, Y.A., Aliper, A., Veselov, M.S., Aladinskiy, V.A., Aladinskaya, A.V., Terentiev, V.A., Polykovskiy, D.A., Kuznetsov, M.D., Asadulaev, A. (2019). Deep learning enables rapid identification of potent DDR1 kinase inhibitors. Nature Biotechnology. 37(9) 1038-1040.

(178) Drews, J. (2000). Drug discovery: a historical perspective. science. 287(5460) 1960-1964.;Jayaraj and Gittelman 2018

(179) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 8

(180) Harrer, S., Shah, P., Antony, B., Hu, J. (2019). Artificial Intelligence for Clinical Trial Design. Trends in Pharmacological Sciences.

(181) Gil, Y., Greaves, M., Hendler, J., Hirsh, H. (2014). Amplify scientific discovery with artificial intelligence. Science. 346(6206) 171-172; Vamathevan J., Clark D., Czodrowski P., Dunham I., Ferran E., Lee G., Madabhushi A., Shah, P, Spitzer, M. (2019). Applications of machine learning in drug discovery and development. Nature Reviews Drug Discovery 1.

expanded[182]. The resulting drug candidates could then serve as useful starting points to examine whether they are safe and effective in treating their disease targets[183].

Once a list of potential candidates for a drug compound is identified, AI can also be used to choose which molecules possess suitable characteristics to address the biological target of interest[184]. For example, AI can find the optimal chemical structures to reduce toxicity and to satisfy metabolic requirements, both of which searches can be costly and data-intensive[185]. Similarly, human errors – which have proven to constitute a barrier during this discovery process – are less likely to occur if the process is primarily driven by algorithms[186]. By facilitating the discovery and verification of drug candidates, AI is expected to have a strong effect on the early stage of drug development (before clinical trials)[187].

While AI can accelerate the discovery of drug candidates at the earliest stage, it is unclear whether the drug candidates thus uncovered represent incremental or radical improvements on existing drugs[188]. Research has shown that discovering drug candidates that can address new diseases is much harder than discovering incremental "me-too" drugs[189], but on average, the return on these drugs is substantially higher than "me-too" drugs[190]. Although it is difficult to assess a drug's therapeutic impact at the early stage (without going through lengthy clinical trials to ascertain efficacy and safety), the chemical novelty of a drug is often used as a proxy for measuring its potential impact for addressing new diseases or conditions, providing an *ex-ante* measure of drug quality[191]. Thus, it is important to find a new drug candidate with novel chemical properties, both for public health and for firms to stay competitive. Investing in AI is an attractive option if it can accelerate the process of discovery for drugs with a novel chemical property (a novel drug, in short) that can address new disease and conditions[192].

AI can help to identify new chemical compounds by finding hidden patterns in digitised data, much faster than humans. For example, Watson from IBM identified six new p53 kinases – a cancer suppressor – within a month. It would have taken researchers about six years to do this. This accelerated discovery was possible because there were p53 kinases with established chemical properties that were already known, making it relatively easy for AI to find other types of p53 proteins that would fit into existing patterns.

In contrast, there are virtually no precedents for novel drugs that by definition are radically different from existing compounds. Here, machine learning offers no real advantages[193]. When data are scarce, it is not easy to draw conclusions about the functionalities of drugs through ML. Conclusions based on limited data can depend heavily on implicit knowledge, the acquisition and transmission of which is inherently costly and therefore difficult to digitise and use as input for AI systems[194]. The development of novel medicines also requires a deeper understanding of a narrow field of implicit knowledge, to which AI can only make a limited contribution, being better suited to tasks that require a broad search across several innovation areas[195].

In contrast, the example of the discovery of artemisinin – for the treatment of malaria – illustrates well the limits of AI capabilities compared to the discovery of p53 proteins. Artemisinin was discovered mainly through limited data and human ingenuity. Dr. Youyou Tu, the inventor of artemisinin, combined her clinical experience with a sentence in a third century book not directly related to malaria to bring about a paradigm shift in the development of

(182) Dougherty, D., Dunne, D.D. (2012). Digital science and knowledge boundaries in complex innovation. Organisation Science. 23(5) 1467-1484; Pisano, G.P. (2006). Science business: The promise, the reality, and the future of biotech. Harvard Business Press.
(183) Bowen, L., Wu, L. (2020), Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 8
(184) Ibid.
(185) Lo, Y-C, Rensi, S.E., Torng, W, Altman, R.B. (2018). Machine learning in chemoinformatics and drug discovery. Drug Discovery Today.; Schneider, G. (2018). Automating drug discovery. Nature Reviews Drug Discovery. 17(2) 97.
(186) Gil, Y., Greaves, M., Hendler, J., Hirsh, H. (2014). Amplify scientific discovery with artificial intelligence. Science. 346(6206) 171-172.
(187) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 8
(188) Ibid.
(189) A drug that is structurally very similar to already known drugs, with only minor differences. The term "me-too" carries a negative connotation. However, me-too products may create competition and drive prices down, https://www.medicinenet.com/script/main/art.asp?articlekey=33748
(190) Krieger, J.L., Li, D., Papanikolaou, D. (2018). Developing Novel Drugs. National Bureau of Economic Research.
(191) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 8
(192) Ibid.
(193) Wu L, Lou B, Hitt LM (2019). Data Analytics Supports Decentralised Innovation. Management Science. Forthcoming.
(194) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 8
(195) Taylor, A., Greve, H.R. (2006). Superman or the fantastic four? Knowledge combination and experience in innovative teams. Academy of Management Journal. 49(4) 723-740; Weisberg RW (1999). 12 Creativity and Knowledge: A Challenge to Theories. Handbook of Creativity. 226.

antimalarial drugs[196]. In contrast, current AI technology is not able to understand the meanings behind the ancient text in order to make the necessary connection to treat malaria[197]. Even if it could, a single data point would not help ML algorithms to draw useful conclusions[198].

2.3.5. Strategic cooperation between the pharma industry and AI companies

In recent years, there has been an increase in AI applications in the healthcare sector. Big pharma has invested in AI companies, or entered into joint ventures with them, driven by the expectation to use AI to develop better healthcare tools. A recent report from Deep Knowledge Analytics (July 2019) shows that globally there are 170 AI companies, 50 corporations, 400 investors and 35 major R&D centres active in pharmaceutical research. According to Deep Knowledge Analytics the global AI R&D market increased from USD 200 million in 2016 to more than USD 700 million in 2018, and it is expected to reach USD 20 billion in the next five years[199]. The following examples demonstrate how big pharma companies liaise with specialised AI companies to speed up drug discovery and thereby to reduce costs.

Atomwise (San Francisco, USA)[200] has invented the first adaptive AI technology for the structure-based development of small molecule drugs. Founded in 2012, Atomwise currently conducts hundreds of projects each year in collaboration with some of the world's largest pharmaceutical and agrochemical companies and more than 200 universities and hospitals in 40 countries. Atomwise has received over USD 50 million from leading venture capital firms to support the development and application of its AI technology.

Since 2012, the Atomwise deep learning platform has been used to develop new drugs. This includes investigating the possibility of whether already known and tested drugs can combat new diseases. The technology is based on a technology similar to deep learning methods for image recognition. Atomwise extracts knowledge from millions of experimental measurements and protein structures to predict the binding of molecules to proteins. This enables chemists to make precise predictions, for example about the toxicity of a substance. Atomwise describes its proprietary technology as follows: *"AtomNet is based on convolutional neural networks – the same AI technology that recognises faces in a crowd, enables self-driving cars, and allows you to talk to your phone. Our technology uses a statistical approach that extracts the insights from millions of experimental affinity measurements and thousands of protein structures to predict the binding of small molecules to proteins. This fundamental tool makes it possible for chemists to pursue hit discovery, lead optimisation and toxicity predictions with unparalleled precision and accuracy"*[201]. In March 2020, Atomwise and Bridge Biotherapeutics announced a deal to unlock blockbuster potential in treating inflammation with the help of AI[202].

Exscientia (Cambridge, UK)[203] Exscientia, founded in 2012, is a leading company in AI-based drug research and development. In addition, it is considered to be the first company to automate drug discovery and development, outperforming conventional methods by connecting AI technology with experienced researchers.

In January 2020, the pharmaceutical and agrochemical group Bayer announced a three-year cooperation with Exscientia on AI-supported research into active ingredients to combat heart disease and cancer. The collaboration is aimed at early-stage research projects for the treatment of cardiovascular and oncological diseases. The alliance could result in up to EUR 240 million in payments to the British company[204].

Exscientia and the Japanese pharmaceutical company Sumitomo Dainippon Pharma recently developed a drug entirely by artificial intelligence for the first time. The drug (DSP-118) is for the treatment of obsessive-compulsive disorders (OCD). It was created by algorithms that searched a huge database for suitable chemical structures for the new compound to stimulate a specific receptor in the brain that is active in OCD. Another special feature of

(196) Tu, Y. (2011). The discovery of artemisinin (qinghaosu) and gifts from Chinese medicine. Nature Medicine. 17(10) 1217.
(197) Marcus, G., Davis, E. (2019). Rebooting AI: Building Artificial Intelligence We Can Trust. Pantheon.
(198) Bowen, L., Wu, L. (2020). Artificial Intelligence and Drug Innovation: A large scale examination of the pharmaceutical industry, 9
(199) Deep Knowledge Analytics, AI for drug discovery, biomarker development and advanced R&D landscape overview 2019 / Q2, https://ai-pharma.dka.global/quarter-2-2019/
(200) https://www.atomwise.com/
(201) https://www.atomwise.com/our-technology/
(202) https://www.businesswire.com/news/home/20200309005771/en/Atomwise-Bridge-Biotherapeutics-Sign-1B-Deal-Unlock
(203) https://www.exscientia.ai/
(204) https://www.boerse-frankfurt.de/news/52e5d655-274f-44e4-bb7b-65e73f390088

the drug was its rapid development as it was ready for use in just under twelve months. Normally, drug development takes around four and a half years. If the OCD drug is successful in the first test phase, further studies are to follow worldwide.[205].

Cyclica (Toronto, Canada) is a start-up that applies DL and computational biophysics to innovate the drug discovery process. In 2018, Merck announced the signing of a license agreement with Cyclica for the use of its cloud-based in-silico platform Ligand Express for proteome screening[206]. Ligand Express® is a structure-based platform for proteome screening that is enhanced by artificial intelligence (AI). It is used to identify novel targets that are modelled to interact with small molecule compounds. Ligand Express® leverages computational biophysics and DL to screen small-molecule ligands against the structurally characterised proteome, ML to predict modulatory effects of ligand-protein interactions, and bioinformatics and systems biology to provide insights into disease pathways.

The conventional development of therapies based on small molecule drugs focuses on specific, disease-associated protein targets. However, as a drug enters the body, it undergoes dozens, if not hundreds, of interactions with proteins before being eliminated from the body. With Ligand Express, it is possible to present a unique all-round view of the proteome for a specific small molecule. Since the technology can model how a small molecule will interact with all proteins (of known structure), it can help identify both "on-targets" (interactions that have a desired effect in a particular disease) and "off-targets" (interactions that can cause side effects)[207]. In 2018, Bayer announced a long term strategic cooperation with Cyclica with the aim to jointly develop new drugs with the help of Cyclica's AI technology[208].

2.3.6. Case study – The role of human intervention in AI augmented de novo drug design (IKTOS)

The company

IKTOS[209], based in Paris (France), was founded in 2016 by Quentin Perron, Nicolas Do Huu, and Yann Gaston-Mathé. The aim of the company was to develop an innovative and user-friendly technology platform for DL-based *de novo* drug design, leveraging a proprietary algorithm developed by Quentin and Nicolas, who initially wanted to apply to chemistry the deep generative models previously used in fields such as image processing and NLP. Since that time, academic and industry publications in the field of DL for drug design have flourished. However, IKTOS has retained a pioneering positioning and specific expertise in the field of ligand-based *de novo* design for multi-parameter optimisation.

The company currently employs about 30 people and has collaborated actively with several biopharmaceutical companies, where IKTOS AI technology is used to accelerate the design phase of promising compounds. Among its cooperation partners are Merck, Grünenthal, Syngenta, Servier, and Sanofi. As IKTOS understands itself as a AI service provider, its current main challenge is in establishing a sustainable business model. The aim is to move from a project-based approach to one offering AI solutions in the framework of an AI as a service model that would leave the whole process of drug discovery with the client.

Furthermore, IKTOS and SRI International (SRI), a research centre headquartered in California, jointly announced early this year that they have entered into a collaboration agreement designed to accelerate the discovery and development of novel anti-viral therapies. Under that collaboration, the generative modelling technology of IKTOS will be combined with SRI's SynFini™, a fully automated end-to-end synthetic chemistry system, to design novel, optimised compounds and accelerate the identification of drug candidates to treat multiple viruses, including influenza and the coronavirus (COVID-19).

(205) https://t3n.de/news/vollstaendig-ki-entwickeltes-1248880/
(206) https://www.pharmatechnik-online.com/merck-will-ki-basierte-screening-plattform-nutzen/
(207) https://www.laborpraxis.vogel.de/amp/der-steinige-weg-zum-blockbuster-a-783089/
(208) https://www.businesswire.com/news/home/20181126005455/en/Cyclica-Drives-Drug-Discovery-AI-Augmented-Integrated-Network
(209) http://iktos.ai/

The technology

Basic description

IKTOS uses deep generative models – inspired by natural language analysis algorithms used in translation – to imagine new molecules. Intelligent, they learn from what already exists and from *in silico* tests to predict the therapeutic effectiveness of their creations. Structural activities and relationships at atom-level are identified in an automated process that is based on specifically designed algorithms.

More precisely, the company uses Long Short-Term Memory (LSTM)[210] combined with re-enforcement learning for multi-parameter optimisation (MPO) of molecules. Accordingly, new molecules are generated in a first step with the help of the generative model based on LSTM (trained on 86 million molecules). In the second step, molecules are scored by a multi-objective fitness function. The objectives are set using data from traditional approaches such as QSAR models[211], docking scores[212], and metrics/descriptors, etc. In a third step, the weights of the model are adjusted using a policy gradient algorithm to maximise the probability of generating molecules similar to those maximising the global score[213]. The process runs iteratively until the target score has been maximised.

IKTOS is also using AI to develop new pathways to generate specific compounds. The application (called Spaya) is currently being offered for free on the web to chemical engineers[214]. The process works in a fully automated manner. "Spaya.ai" is an online platform designed to help users in their practice of retrosynthetic analysis and chemistry in general. Purely data-driven, its algorithms empower users with machine-augmented capacity to navigate potential retrosyntheses in minutes.

Technology development and learning

The technology was invented by the company founders in 2016. At that time, they were the first and only persons to develop such a method. In the same year, by coincidence, Toronto University published a paper dealing with the same approach. Since then, dozens of researchers have engaged worldwide to develop this approach further. Currently about 10 papers are being published each month on the subject.

The model is therefore under continuous improvement, with progress on new algorithms, not on new data. Current research and development activities include scaffold hopping (a strategy in medicinal chemistry for developing new active ingredients). The aim is to modify – with the help of deep generative learning – the basic structure of a known active ingredient in such a way that new active ingredients with the same mechanism of action are obtained. In addition, efforts are underway to use deep learning for better exploitation of High-Throughput-Screening[215] (HTS) to accelerate the hit-to-lead[216] stage of drug discovery.

(210) Long short-term memory (LSTM) is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections. It can not only process single data points (such as images), but also entire sequences of data (such as speech or video), https://en.wikipedia.org/wiki/Long_short-term_memory

(211) Quantitative structure-activity relationship (QSAR) models - collectively referred to as (Q)SARs - are mathematical models that can be used to predict the physicochemical, biological and environmental fate properties of compounds from the knowledge of their chemical structure. These models are available for free or as commercial software, https://echa.europa.eu/support/registration/how-to-avoid-unnecessary-testing-on-animals/qsar-models

(212) In the fields of computational chemistry and molecular modelling, scoring functions are mathematical functions used to approximately predict the binding affinity between two molecules after they have been docked. Most commonly one of the molecules is a small organic compound such as a drug and the second is the drug's biological target such as a protein receptor, https://en.wikipedia.org/wiki/Scoring_functions_for_docking

(213) The goal of reinforcement learning is to find an optimal behaviour strategy for the agent to obtain optimal rewards. The policy gradient methods target at modelling and optimizing the policy directly, https://lilianweng.github.io/lil-log/2018/04/08/policy-gradient-algorithms.html#what-is-policy-gradient

(214) https://spaya.zendesk.com/hc/en-150

(215) HTS (high throughput screening) is often the first step on the road to drug discovery. Once a potential target has been identified, screening of a large number of substances can help find a molecular starting point.

(216) Hit to lead (H2L) also known as lead generation is a stage in early drug discovery where small molecule hits from a high throughput screen (HTS) are evaluated and undergo limited optimisation to identify promising lead compounds, https://en.wikipedia.org/wiki/Hit_to_lead

The use case – de novo drug design with the Makya project

Brief description

Designing and producing a new drug candidate is no small task. The process is particularly long (4-5 years on average), costly (USD 50 to 100 million of investment in R&D) and rarely successful (only 1% of the molecules tested reach the preclinical development stage and only one in 10,000 is finally marketed as a drug). In other words, pharmaceutical research must meet very precise criteria to guarantee efficacy and patient safety and for a drug to be eligible for future development. The Makya project has the objective to drastically accelerate research processes in the field of medicinal chemistry using an algorithm capable of imagining the "right molecule" at very high speed. This is a way to completely revolutionise the design of new drugs.

According to the founders, the Makya project is like the automated resolution of a Rubik's cube. The question to be addressed is: How to obtain the ideal molecule, simultaneously respecting a certain number of characteristics (simultaneous optimisation on activity, potency, ADME, toxicity, etc.) while knowing that the modification of a single parameter impacts the rest of the system (i.e. gain on one objective usually results in loss on the other ones)? Current techniques do not have the capacity to automatically and efficiently create ideal virtual molecules, so the IKTOS teams relied on the contributions of deep learning and big data to design an algorithm trained on a public database of more than 86 million molecules, in order to teach IKTOS to imagine new molecular structures itself. In fact, it is a question of finding the "key", in other words the ideal molecule, to trigger the required biological mechanism.

Makya proposes to generate molecules using algorithms that have learned the rules of chemical syntax and are capable of distinguishing between molecules. In this way, Makya is able to imagine new molecular structures at very high speed.

The role of human intervention

Human intervention in a specific *de novo* drug discovery project is present in several steps, as follows:

1. **Preparation/definition of the initial dataset:** this step requires decisions involving medicinal chemists, in particular with regard to the parameters/dimensions and type of molecules to include. Here two options exist:

   a. The initial dataset is provided and defined by the client: data on potential molecules including specific properties are owned by the client and handed over to IKTOS.

   b. The initial dataset is provided by IKTOS. The work build then solely on Makya and the 86 million molecules included in the dataset.

2. **Setting up of specific objectives for the multi-parameter optimisation process:** objectives and structural requirements are usually defined by medicinal chemists of the client.

3. **Virtual compound generation**: based on the predefined properties, molecules are generated by means of deep generative learning. This process runs in an automated manner at IKTOS.

4. **Preselection of compounds**: ideal "virtual molecules" are preselected by medicinal chemists for the optimisation process.

5. **Running of the optimisation process**: this iterative trial-and-error process requires interventions by the engineers at IKTOS, repeatedly setting the weights of the model until the goal score has been reached.

6. **Selection of the compounds**: the client (i.e. their medicinal chemists) select those compounds with high goal score values for further research (i.e. synthesis and testing).

Business Models and IP rights

Business models

The business model of IKTOS currently builds both on research collaborations with universities and/or drug discovery companies, and on services sold confidentially to pharmaceutical companies.

For the future, the development of a SaaS (software as a service) solution for Makya is also planned. Eventually, the platform will be directly accessible to all the laboratories' chemical researchers, who will be able to use it autonomously. In this future stage, revenue for the company will come mainly from service fees for the access to and use of the Makya platform.

IKTOS offers its DL-based SPAYA service - an AI empowered of-the-shelf solution - for free. It enables customers to obtain output data created by the service based on customer input data in the field of chemistry. The Spaya community of chemists has a unique opportunity to share expert knowledge and help one other. Enriched by the community and for the community, and as the community grows, the SPAYA algorithms will improve over time.

Patenting

New lead compounds are the main patentable output of the process described above. The use of the software is a major input in the invention process (like, for example, laboratory equipment or other research infrastructure). The selection of a compound at the end of the process, or the preselection of a compound for an optimisation process, is done by the client and thus subject to human intervention.

Whether and how patenting is done in a particular project depends on the type of client. Big pharma companies do not allow shared patenting and, at the end of a project, all IP rights revert to the client. In collaborative projects with a drug discovery company or an university, IP rights are shared among the partners, the designated inventors typically being the scientists of the client team and engineers at IKTOS.

2.3.7. Case study - The role human intervention in AI augmented target identification (BenevolentAI)

The company

BenevolentAI is a small biotechnology company that uses biomedical data and AI – in particular ML – to improve all levels of drug development: basic science, target identification, molecular design and clinical outcomes. AI is being used to help find new medicines for a number of serious diseases, including Parkinson's disease.

Founded in November 2013 by Ken Mulvany, Benevolent AI was initially known as Stratified Medical Ltd. Mulvany is a serial entrepreneur, investor, and healthcare and technology veteran with over 20 years of business leadership experience. Previously, he was CEO of Proximagen, a UK-based biotech company committed to delivering novel drugs and innovative new treatments for central nervous system disorders. Since its establishment, the unicorn start-up BenevolentAI has raused USD 292 million in funding. The company's main challenges related to IP include the strong need to both protect knowledge – on newly identified targets and matching compounds – and its own technological innovations.

BenevolentAI has its headquarters in London, a research facility in Cambridge (UK) and further offices in New York and Belgium. It has active R&D drug programmes – from discovery to PhaseIIb – in disease areas such as ALS, Parkinson's, ulcerative colitis and sarcopenia.

The technology

Basic description

The company's *Benevolent Platform* is an experimental drug discovery platform that uses AI for scientific innovation. The Platform absorbs and analyses biomedical information and combines it with AI deep learning to create a bioscience knowledge graph. Scientists use the automated platform to hypothesise and validate new ways to treat disease and personalise drugs for patients.

The Platform focuses on three key areas – (1) target identification, (2) molecular design, and (3) precision medicine – to better understand the underlying mechanisms of diseases and to develop new treatments. It contains structured and unstructured data and applies Large Scale NLP, ML, DL and Neural Reasoning Structured Data Technology.

In target identification, it is mainly relation inference AI models that are used to predict potential disease targets that may be easily overlooked by scientists. The BenevolentAI knowledge graph integrates biomedical data from structured and unstructured sources. Using molecular, clinical, pharmacological data and scientific literature, it derives contextual relationships across all relevant biomedical entities, leading to the proposal of novel and optimal drug targets. It is queried by a fleet of algorithms to identify new relationships to suggest new ways of tackling disease.

Molecular design builds mainly on BenevolentAI's *EvoChem* platform. EvoChem designs *de novo* compounds based on multiparametric optimisations with a scoring function that factors in all the features to be optimised for that molecule. The compound ideas are ranked, the best candidates being selected directly for synthesis, others informing chemists on where to furtherexplore the chemical space.

In precision medicine, machine learning models are used to *identify patient groups* by the molecular signature of their disease and design, enabling faster clinical trials. This allows the firm to identify patient subtypes more likely to respond to drugs, further increasing the probability of clinical success. The approach has benefits for existing medicines too: it can be used to elucidate the mechanism of action, identify new patient responders, improve diagnosis and more precisely target treatment.

Technology development and learning

All AI technologies at BenevolentAI have been developed in-house and are under continuous improvement. Since its formation in 2013, BenevolentAI has been developing a knowledge pipeline based on data from various structured and unstructured biomedical data sources, curating and standardising this knowledge via a "data fabric", an architecture to enable incoming disparate data to be used in a harmonised manner. The data is fed into a proprietary knowledge graph, which extracts and contextualises the relevant information. The knowledge graph is made up of a vast number of contextualised, machine curated relationships between diseases, genes, drugs and over 20 types of biomedical entities.

Use case – AI augmented target identification

Brief description

A target is a biomolecule to which a drug can bind, enabling the active ingredient to have its effect. Active ingredients can be drugs, for example, so targets play an important role in drug research. Targets are often proteins, for example an enzyme, a receptor or an ion channel. The identification of targets – as well as their investigation in signal transduction processes and the understanding of their interactions with ligands – are central elements of modern biomedical research. In addition to molecular biology and molecular pharmacology, this applies in particular to the pharmaceutical sciences. Targets are responsible for the effect of drugs, but also their kinetics in an organism. The elucidation of the structure, conformational behaviour and catalytic properties of specific targets enables the rational design of active ingredients and biotechnological processes.

BenevolentAI uses its Benevolent Platform in target identification. Using AI in the target identification process – synthesising the data and making inferences on the best targets – enables scientists to explore all the available evidence to better understand a disease and its underlying biology, too large a task for one, or even for a team of people. Scientists are facilitated to make better informed decisions about which targets are most likely to succeed and for which patients. For example, in the identification of new potential treatments for ALS, billions of sentences and paragraphs from millions of scientific research papers within the BenevolentAI knowledge graph were first analysed with the help of AI, leading to a hundred potential hypotheses to be validated by researchers at BenevolentAI. The process led initially to five final hypotheses. These were further refined into two targets in line with ongoing research paths and two that were new to the researchers.

A concrete target identification project leads to the suggestion of new drug targets as its main outputs. Such target identification is usually organised as an internal project within BenevolentAI.

The role of human intervention

Human intervention is present in several steps of a specific target identification project, as follows:

1. **Formulation of the biological question**: this first step requires human intervention. Drug discovery expertise is needed to do proper disease profiling (i.e. to investigate into the systemic roots of a specific disease), to identify data sets and quality, and to ensures biological relevance.

2. **Prediction of new therapeutic targets**: structured and unstructured data is analysed with the help of ML based on the inputs defined in step 1. Human interaction is present in this step in adjusting the parameters of the ML model and in looking at the outputs in an iterative mode.

3. **Assessment of the relevance of targets proposed by the model**: here human intervention is needed to decide whether predictions are sufficiently different from standards in each case. This judgment needs to done against the landscape of existing therapies (i.e. how the new target fits).

4. **Identification of the most appropriate target**: the modelling of the most relevant disease mechanisms involves both ML and human intervention. ML is used to manage the breadth and depth of the knowledge graph, while further research studies around targets is performed mainly by humans with a high degree of *in silico*[217] background.

5. **Testing and experiments**: pre-clinical testing is done by BenevolentAI in its Cambridge-based premises; some tests are also done with contract research partners but it is noteworthy that all assays are designed by BenevolentAI. Also, all analysis is done within the company.

Use case – AI augmented lead optimisation

Brief description

A lead is a chemical substance that is investigated in pharmaceutical research in the course of drug design as a starting point for the development of a drug candidate. A lead structure already shows a desired biological effect *in vitro*, but does not yet have the qualities required for a drug substance in terms of potency, selectivity and pharmacokinetics. A lead structure differs from any hit from high-throughput screening in that its structure enables the synthesis of analogue molecules. A lead structure can also be a natural substance that needs to be improved in its pharmacological properties by chemical modifications. Using suitable chemical, biochemical and *in silico* methods, medicinal chemistry attempts to optimise the molecular structure, and thus the mode of action, of the lead structure to such an extent that the drug candidate can be tested in the next phases of drug development.

(217) In silico modelling is used to calculate how the pharmacology and pharmacokinetics of a drug, determined in laboratory and animal experiments, might act in humans. Complex computer simulations of the physiology of humans and animals are used to determine how a substance affects humans before the first data from tests on humans are available. In silico models of diseases are also being developed to gain insights into their causes and underlying disease mechanisms or how patients' symptoms will develop over time with or without new treatments. These technologies are used to identify potential drug targets, predict the relationship between dosage, drug concentration in the body and treatment efficacy, identify appropriate patient populations or markers of drug effects, optimise clinical trial design and analyse clinical trial data to achieve the desired outcome.

BenevolentAI's lead optimisation projects usually build upon promising targets stemming from internal target identification projects. In such projects, the company applies AI with help of its chemistry lead platform. ML is used to support the design process of drug-like molecules.

The role of human intervention

Human intervention is present in a specific lead optimisation project in the following steps:

1. **Design of molecules**: human intervention is needed here in two stages: for setting the parameters for the model (i.e. the design algorithm) and for the triage of the most potent output compounds. ML is used to find optimal compounds. The algorithm helps to better understand the pharmacokinetics and toxicology of each compound; and characteristics of each compound are predicted by the model accordingly.

2. **Selection of the compounds**: once a set of optimal compounds has been generated by the model, the most promising compounds will be selected for further testing.

3. **Testing of most promising compounds**: pre-clinical testing is done for candidate selection. The main aim is here to identify those compounds that might in later stages lead to patentable knowledge. In this stage, human intervention is present in interpreting the test results and in further selection.

4. **Clinical trial stages 1 and 2**: for the most promising in-house designed compound(s), a proof of concept is established. Again, human intervention is present in analysing the results of the trial stages.

Once a compound has successfully passed stage 2 of the clinical trials, IP rights will be protected by patenting and the company will look for partners (e.g. big pharmaceutical companies) for further trials and market development.

Business Models and IP rights

Business modes

While BenevolentAI began its business as a ML start-up, today it understands itself as a drug discovery company that even plans in the future to become a drug producer.

While the company also cooperates with external partners from the pharmaceutical sector, long-term value creation builds, for the most part, on the commercial exploitation of internal projects.

Although Software as a Service (SaaS) is a major trend in the domain of AI-augmented drug discovery, BenevolentAI does not plan to offer such a service to external partners.

Patenting

BenevolentAI pursues an active IP strategy and files patents in the domains of both pharmaceutical and AI technology, each having designated IP rights managers within the company.

Patent applications for new compounds are a standard option in internal projects. AI is seen in this context only as a technological input into the process of invention. In the company view, there is no difference whether the data for validation is generated by human researchers or an ML model.

IP in the domain of AI has been the subject of an active protection strategy for three years. Before then, BenevolentAI's ML technologies were only protected as trade secrets. Increasing competition, and long term plans to become a producer of new drugs, have motivated the company to alter its IP strategy for its AI technologies.

2.3.8. Future Trends

The use of AI in pharmaceutical research is likely to become mainstream in the next ten years with exponential growth expected in the number of companies using AI in precision medicine. This also means that the trend towards outsourcing drug discovery processes to external specialised AI companies will increase massively by

2030. Such companies will conduct their research primarily *in silico* (i.e. using computer modelling or simulation) in close cooperation with research organisations such as universities[218]. The main drivers of this development will be significant time and cost savings in the development of new potential drug candidates. It is estimated that the time from screening to preclinical testing will be reduced from five or six years to a few months[219].

On a technological level, it is now assumed that, by 2030, next-generation AI technologies (DL, GANs and reinforcement learning) will be used primarily for biomarker development, drug discovery and drug reuse[220].

New applications for AI in pharmaceutical research

AI will, in the near future, support researchers in keeping up with the constantly ***increasing stream of new knowledge***. Eight to ten thousand new scientific publications appear every day – and the trend is rising. It is becoming more and more difficult for individual scientists to keep track of even one subject area. Research-based pharmaceutical companies rely already on artificial intelligence in this situation. Using specially developed algorithms, data scientists extract the most important knowledge from the global ocean of knowledge for research colleagues, day after day. In this way, tailor-made collections of specialist literature are created, which summarise new developments in the researcher's own discipline and provide insights into neighbouring areas. This facilitates the exchange of ideas across disciplines and stimulates new ideas. There could be in the future even a kind of 'Spotify for science' - a streaming service that compiles new literature in an individually adapted way and broadens the horizon with clever suggestions.

Clinical trials are complex, lengthy and expensive. Significant improvements can be achieved with *so-called* ***virtual control arms***. Up to now, a group of patients in clinical trials receive the new therapy to be tested in addition to the standard therapy. The control group also receives a sham drug. This comparison can be used to prove that the new drug may be more effective. In some cases, however, this control group can now be virtually simulated with existing patient data. This enables smaller studies in which patients only receive the new and possibly better therapy. It also saves both time and money. Data from comparable patients who have received the usual treatment outside of clinical studies are used as a basis. Roche, a research-based pharmaceutical company, has already gained practical experience with this method. The data for the artificial control group come from a subsidiary that specialises in processing treatment data from US hospitals for cancer research.

The application of ML to support the ***development of new vaccines*** is still in its nascent state. Current models used in immunology are trained with far smaller datasets than, for example, models for speech or image recognition that are showcases for demonstrating the potential of ML. As the quality of a ML model depends greatly on its initial training data set, there is plenty of room for improvement in the near future. It can be expected that datasets in immunology will, in the next couple of years, become larger and more diverse thus leading to better output quality of ML models that augment the development process of new vaccines. Thus a significant reduction in the development time for new vaccines seems likely to be in reach already in this decade.

The provision of AI technologies as Software as a Service – the move to AIaaS

Software as a Service – abbreviated SaaS – refers to a distribution model for applications via a web browser. SaaS is understood as a subset of cloud computing, since requested applications are never directly available on the user's device. As noted, there has been an increasing trend of offering AI technologies under a similar "as a service" model, which we have referred to as 'AIaaS'.

As mentioned earlier, IKTOS is using AI to develop new pathways to generate specific compounds. The application, Spaya, is free on the web for chemical engineers and the process is fully automated. The "Spaya.ai" online platform helps users in retrosynthetic analysis and chemistry in general. Entirely data-driven, Spaya's algorithms empower users with machine-augmented capacity to quickly navigate potential retrosyntheses.

(218) Deloitte (2019). Intelligent drug discovery Powered by AI, 31

(219) Ibid.

(220) Ibid.

This new software tool to design new pathways to compound (retro)synthesis works without human intervention and accordingly a new process for compound synthesis could, in theory, be developed without a human inventor.

The second domain area is science and specifically meteorology and its application to weather forecasting.

2.4. Science/Meteorology

2.4.1. Introduction

The aim of weather forecasting is to predict the state of the atmosphere at a certain time in a certain place or over a specific area. In fact, this not only means weather phenomena that affect the ground, but taking into consideration the entire atmosphere of the Earth. As a physical event, the weather can be described by the laws of nature. The basic idea of a weather forecast is to predict a state of the atmosphere in the future based on knowledge of past atmospheric conditions, and the corresponding weather, and the current state of the atmosphere, using known physical rules. However, the mathematical constructs that describe these physical rules are non-linear (based on "non-linear equations").

A distinction is made between *manually-prepared (or synoptic) weather forecasting* and *numerical weather forecasting*, although a combination of both methods is still used today. This is due to the fact that even the current numerical forecast models provide insufficient results. In order to take into account the local climatology of weather stations carrying out the measuring of weather characteristics, statistical methods, such as the MOS methods (Model Output Statistics), are still used downstream of numerical calculations. The data on the current state of the atmosphere come from a network of ground measuring stations that measure wind speed, temperature, air pressure and humidity, as well as precipitation levels. Data from radiosondes, weather satellites, commercial aircraft and weather ships are also used.

Since the 1950's, the data basis for computational weather forecasting has also been improved, for example, by weather satellites, a network of radio weather probes and weather radars. The relatively reliable forecast period in the mid-latitudes of the Earth has increased from about three days to four to five days, which has meant a noticeable improvement in forecasting for many sectors of the economy, including transport and construction, as well as for planning in agriculture. The automation and networking of weather stations has also brought great progress. Whereas weather stations used to have to be operated and maintained by specialists, modern weather stations automatically send data to the weather services. Weather stations with their own power supply (solar cells and batteries) and a satellite modem can be set up and operated for years even in the most remote areas. More recent developments in data science make it possible, for example, to recognise patterns in the measured values themselves and to find out, for example, which measuring stations provide the most significant or most reliable values under which circumstances. In this way, even with private, barely standardised weather stations, useful insights can be gained.

2.4.2. How AI is used in weather forecasting

Meteorology is one of the application areas in which machine learning technologies are already used as standard. These are applications that are largely automated and run, once coded and implemented, without further human intervention. The following fields of application can be described concretely:

Bias correction of meteorological observations: a global network of measuring stations at representative locations is required to obtain the data needed to simulate the weather. To enable the three-dimensionality of the forecast model, the physical quantities and weather events are measured and observed on the ground and at different altitudes. Such measurements are recorded by automated measuring stations, staffed weather stations, weather balloons and geostationary and polar orbiting weather satellites. Additional data is provided by merchant ships, buoys and commercial aircraft equipped with measuring technology. ML is mainly used to correct biases in satellite observation data. Such biases are, for example, varying with time (e.g. diurnal or seasonal), varying with the scan position of the satellite instrument, or varying with the position of the satellite around its orbit.

Parameterisation of meteorological models: typical parameterisations are corrections for: cloud microphysics (formation and change of cloud and precipitation particles), radiation, turbulence, ground model, schemes for shallow and high-reaching convection. Deep neural networks are being used to substitute the physical parameterisation scheme in NWP models to reduce simulation time without any loss of quality.

Data assimilation: the determination of the initial state of the model atmosphere is important for the success of the model prediction. For this purpose a weighted combination of measured values and older model predictions is interpolated to the model grid using different mathematical methods (variable methods such as 3D-Var[221], 4D-Var[222], Optimal Interpolation[223], Nudging[224], Kalman filter[225]). Measured quantities from remote sensing devices (radar, lidar, satellites) must be transformed into model variables and vice versa. Conventional measurements of temperature, humidity, pressure, etc. from weather stations, weather balloons, airplanes, ships and buoys need only be interpolated in space and time. Furthermore, erroneous measurements must be sorted out and systematic measurement errors (bias) must be corrected. Neural network algorithms are used for the learning of model error within data assimilation. They can for example successfully emulate existing mathematical methods such as Kalman filters with better computational performance, saving time and costs.

Local downscaling of model output to improve predictions: weather is a complicated phenomenon: air masses that are constantly in motion, fluctuating temperatures and changing air pressure conditions make it difficult to calculate the atmosphere above the Earth. Meteorologists forecast weather using measurements of the current state of the atmosphere and simulating alternative scenarios, such as changing temperature or humidity, calculating how each modification could affect the weather. However, this method is subject to systematic uncertainties that distort the results. Computer scenarios cannot represent some physical relationships in a sufficient level of detail or spatial resolution. For example, predictions on the temperature at certain locations are always too mild and at others too high because local conditions, some of which vary over time, cannot be taken into account by the models. Therefore, it is necessary to post-process the results of the simulations with elaborate statistical methods and expert knowledge in order to obtain better forecasts and more accurate probabilities of occurrence of weather events. ML (i.e. Random Forest[226], Boosting[227]) is used to automate Model Output Statistics (MOS)[228] that is used for local downscaling.

Post processing of weather data: when the numerical forecast of the computers is available, the work of the meteorologists is far from over. That's when the post-processing starts: First of all, numerous weather maps and other visualisations are automatically generated from the digital data. Furthermore all numerical weather forecasts have systematic errors. Before these can be passed on to users, the forecasts must be calibrated with statistical means. Traditionally this has been done with simple linear methods. Neural networks offer a flexible method to model nonlinear relationships between different predicted variables and station-specific properties. The neural network provides better predictions than previous methods and is faster.

Predictive analytics uses historical data to predict future events, including financial, meteorological, security, economic, insurance, mobility and marketing events. Generally, historical data is used to create a mathematical

(221) Three dimensional variational assimilation system

(222) Four dimensional variational assimilation system

(223) This procedure provides an estimate of the state of the atmosphere by a weighted least squares fit to observations and a background field, usually provided by a NWP model forecast, accessed 28.07.2020 http://glossary.ametsoc.org/wiki/Optimum_interpolation

(224) It is also known as Newtonian damping. Instead of changing the state of the model, a corrective force is introduced on the right side of the equation system. This force is proportional to the difference between the model state and the observation, so the model is permanently driven in the direction of the measurements, even if the model dynamics are not sufficient, Schröter, J. (2003), Assimilierung von Messdaten, Ozean, promet, Jahrg. 29, Nr. 1-4, 55-62

(225) The Kalman filter is a mathematical procedure for the iterative estimation of system parameters on the basis of erroneous observations. It is used to estimate system quantities that cannot be measured directly, while the errors of the measurements are optimally reduced. For dynamic quantities, a mathematical model must be added to the filter as a constraint to take into account dynamic relationships between the system quantities. For example, equations of motion can help to precisely estimate variable positions and speeds together, https://de.wikipedia.org/wiki/Kalman-Filter, accessed 28.07.2020

(226) Random forest is a classification method consisting of several uncorrelated decision trees. All decision trees have grown under a certain type of randomisation during the learning process. For a classification, each tree in that forest is allowed to make a decision and the class with the most votes decides the final classification. Random forests can also be used for regression, https://de.wikipedia.org/wiki/Random_Forest, accessed 11.07.2020.

(227) In machine learning, boosting is an ensemble meta-algorithm for primarily reducing bias, and also variance in supervised learning, and a family of machine learning algorithms that convert weak learners to strong ones, https://en.wikipedia.org/wiki/Boosting_(machine_learning), accessed 11.07.2020

(228) Model Output Statistics (MOS) is a statistical procedure in modern weather forecasting that was developed in the USA in the 1960s/1970s. It often involves multi-linear regression equations that are applied to numerical weather models. Nowadays, MOS methods are used worldwide and serve as an aid especially for local weather forecasting.

model that captures key trends. This predictive model is then applied to current data to predict what will happen next, or to suggest actions that will achieve optimal results.

Predictive analytics has received much attention in recent years as supporting technologies have made great strides, especially in the areas of big data and machine learning. An important prerequisite for successful predictive analytics projects is a powerful Big Data infrastructure, which can also include cloud components. The predictive analytics solution collects data from a wide variety of sources, integrates systems such as Enterprise-Resource-Planning (ERP) or Customer-Relationship-Management (CRM) with each other, handles different types of data, standardises and connects them. The data should also be of high quality, i.e. correct, consistent, complete, up-to-date and uniform[229]. In times of climate change, extreme weather situations will occur again and again. Therefore, there is a growing need to be better prepared for these external conditions and to minimise risks with the help of predictive analytics. To predict future developments, data on weather patterns are also linked with logistics information, sales figures, forecasts of retail demand, product data, etc. The choice of data always depends on the specific issue that a company wants to address. In the retail industry, predictive analytics is mainly used for sales forecasts. The forecasts are based on historical sales figures, data from suppliers and market research and, of course, current or expected weather conditions. Precise sales forecasts prevent costly over- or under-stocking in the warehouse, avoid short sales and ensure that the product range is always available.

2.4.3. Case Study – the role of human intervention in AI supported weather forecasting (UBIMET)

The company

The company was founded in September 2004 as a start-up by Michael Fassnauer and Manfred Spatzierer under the name Meteomedia GmbH as a subsidiary of the Meteomedia Group in Vienna. The aim of the company founders was to significantly improve weather forecasts in Austria and to break up the existing oligopoly of the state weather services (ZAMG, Austro Control). One year after its foundation, Die Presse[230] and ÖBB[231] (the national railway company) and the Uniqa Insurance Group were won as customers, and a separate forecasting centre for extreme weather events - the Austrian Storm Centre - was established. The severe weather centre supplies more than two million customers in Central Europe with severe weather warnings via cooperations with insurance companies. From January 2009, the service was also offered in several countries in Eastern and South-Eastern Europe.

In December 2008, the Meteomedia Group shares were bought back by the company founders. With the repurchase of the shares, the name was changed from Meteomedia GmbH to UBIMET GmbH, whereby UBIMET stands for "Institute for UBIquitäre (ubiquitous) METeorologie". Headquartered in Vienna, UBIMET serves customers in Eastern and Central Europe and Australia, in addition to the Austrian market. The company offers weather forecasts in newspapers, radio, on the Internet and as wireless data services on mobile phones. The last of these includes severe weather warnings in seven languages. The largest customer groups are the infrastructure, government and insurance sectors. The two-man company has become one of the largest and fastest growing private providers of meteorological services with now more than 400 employees.

UBIMET applications of artificial intelligence-based services are used in the following areas, amongst others, where UBIMET has successfully completed customer-specific projects:

- Energy management (runoff modelling, weather-dependent energy production, household demand and predictive maintenance)
- Logistics and infrastructure (efficiency and capacity management)
- Insurance and financial management (reliable prediction of loss figures to control internal processes)
- Retail (consumer demand).

(229) https://www.predictiveanalyticstoday.com/what-is-predictive-analytics/
(230) Main Austrian newspaper
(231) Austrian national railway corporation

The technology

Basic description

In the field of meteorology, different approaches to ML are used. These are, for example, multi-linear regressions, logarithmic regressions and boosting[232]. In addition, neural networks with one hidden layer (both feed forward and backward propagation networks) are also used.

Examples of the use of neural networks include the following application areas: horizontal visibility prediction (e.g. for Quatar airport), water level predictions in complex river systems, and predictions of electricity production in the wind and solar energy sector. Machine learning is used, for example, in correlation models (with multi-linear regressions being used to predict the influence of weather data on the expected sales of barbecue products in supermarkets over the coming weekend).

Currently, only supervised learning is used: i.e. there is (still) no self-configuration of the systems. However, a research project with Reinforcement Learning for Now Casting applications has been submitted recently to a major national research funding agency in Austria.

Technology development and learning

For AI applications, no third-party software is used. All developments in the field of AI are coded in-house.

UBIMET annually invests more than 25% of its turnover in research and development. Partnerships with international research institutes, renowned universities and industrial companies ensure innovation leadership and help in further expansion.

Technology development is mainly carried out through cooperative research projects submitted to national and European programmes. The basic motivation here is the further development of existing methodologies. The ultimate goal is, of course, to gain competitive advantages over competitors. Since 2006, UBIMET has implemented 35 projects in the fields of Intelligent Systems, Energy, Mobility and Safe Systems. UBIMET develops concepts for energy-efficient buildings, precise warnings for disaster management and weather models for innovative agriculture.

Different core systems – frameworks for the processing of different questions – are being further developed. For example, multi-linear regressions are a framework within which model output statistics for meteorological models can be developed as well as regional sales forecasts for yogurt. In contrast, neural networks with a hidden layer are mainly used in the fields of energy production or insurance (damage prognosis due to extreme weather). Each framework is a system that has grown over the years and is scalable for different questions.

In cooperative research projects, a consortium agreement regulates the allocation of IP ownership and rights between the partners. In the case of industrial contract researchs (i.e. contracts with UBIMET), the IP created remains with the client in most cases.

Use case: predictive analytics for the retail sector

Basic description

For grocery retailers, the key question is: how can the ordering process be optimised so that, on the one hand, storage capacities are used optimally and planned with foresight for fresh goods, and, on the other hand, spoilage is minimised as far as possible and sell-out (rupture of product availability) is kept low at the same time? This requires not only short-term but also seasonal forecasts. Using these, the demand for fresh meat, fruit, vegetables or dairy products during a rainy summer, for example, can be easily estimated. More than 30 percent of all fresh produce is weather-elastic, i.e. sales depend very closely on the actual weather or the weather or temperature forecast. Mozzarella, for example, shows a clearly pronounced seasonality. While a medium-sized supermarket

(232) In machine learning, boosting is an ensemble meta-algorithm for primarily reducing bias, and also variance in supervised learning, and a family of machine learning algorithms that convert weak learners to strong ones, https://en.wikipedia.org/wiki/Boosting_(machine_learning), accessed 11.07.2020

usually sells only between 10 and 30 units per day in Autumn, Winter and Spring, the sales volume increases significantly in Summer, when over 120 units are sold on some days. In addition, a much greater fluctuation in sales is observed in Summer, usually between 30 and 100 units per day.

The weather is only one of several important factors that, in combination, determine turnover. For example, shopping in a supermarket that has a parking space is very different from shopping in a branch of the same retail chain in the pedestrian zone and without parking facilities. If it snows, walk-in customers are lost and the turnover drops towards zero.

Using predictive analytics methods, relevant forecast and analysis parameters can be identified from unstructured data sets. These include retailer data with day- and store-specific sales, information on special promotions and the general purchasing frequency at the respective location. In addition, high-resolution historical weather data are used, which are put into a quantitative context with the retail data of the stores using artificial intelligence (i.e. multilinear modelling). In this way, purchasing and warehousing can be significantly improved. The results reduce the error rates in forecasts of product sales per store environment by an average of 10 percent. The decisive factor is the meteorologists' deep understanding of the weather data and the characteristics of customers (based on the knowledge of the retailer).

The role of human intervention

The development of a sales forecasting system for a customer in the retail sector (supermarket chain) has the following steps, which build on each other:

1. **Initial contact with the potential customer and briefing for the UBIMET team**: in this step, the main points of the requested forecasting system are defined (i.e. characteristics of the goods; type, number and location of stores, etc.).

2. **Transmission of the defined data** (gathered preferably over several years) **in the format used by the client**.

3. **Data processing at UBIMET**: here, manual human intervention is required in two areas:

    a. Making the data readable for UBIMET. The data is manually transferred into a format used by UBIMET.

    b. Quality control on the transmitted data. Data gaps, outliers and artefacts are identified and still largely corrected manually. It is important to ensure the consistency of the data in order to achieve a good forecast quality (to avoid that there will be "garbage in - garbage out").

4. **Correlation of the data set with weather data** (using at least four dimensions): in addition to sales data, the dataset also contains data on all factors that can influence sales (e.g. calendar day, special sales promotion, etc.). This process is semi-automated and thus still requires human intervention. Automated feature selection[233] is not yet possible.

5. **The training of the system**, which runs in an automated manner without human intervention.

6. **The control of the results of the training is again carried out manually by humans**: if necessary, the input data will be checked and adjusted again. This is an iterative process (adaptation of input data - training - result checking) repeated until a satisfactory quality of forecast is reached.

7. **Handing over the study** (i.e. final dataset, forecasts and interpretation) **to the customer and decision on implementation:** if the customer is satisfied with the quality of the forecasts, the next step is a permanent operationalisation.

8. **Operationalisation**: from now on, the forecasts are calculated for the customer on the basis of a continuous data feed of sales figures and other indicators. IT-monitored daily forecasts are generated and

(233) Feature selection or variable selection is an important step in different machine learning tasks

automatically verified every five to ten days. With the exception of any trouble shooting, this process runs without human intervention. Human intervention is only necessary in the event of a new training session (e.g. when the system has to be adapted).

Use case: post processing of weather data with AI

Basic description

UBIMET operates its own high-resolution weather model RACE (Refined Atmospheric Condition Evolution) with 100 metre by 100 metre resolution. The latest weather data (from precipitation radar, satellite data, aircraft altitude profiles, etc.) are fed into RACE on an hourly basis at data centres in New York, Melbourne and Vienna, to refine its forecasts. Highly complex data assimilation (3dVar) and nudging techniques (LHN[234]) are also used in this process.

The most accurate geodata is used by UBIMET in downscaling algorithms to achieve model resolutions of a few metres. Thus, UBIMET is able to create spatial and temporal forecasts tailored to the needs of a wide range of industries worldwide.

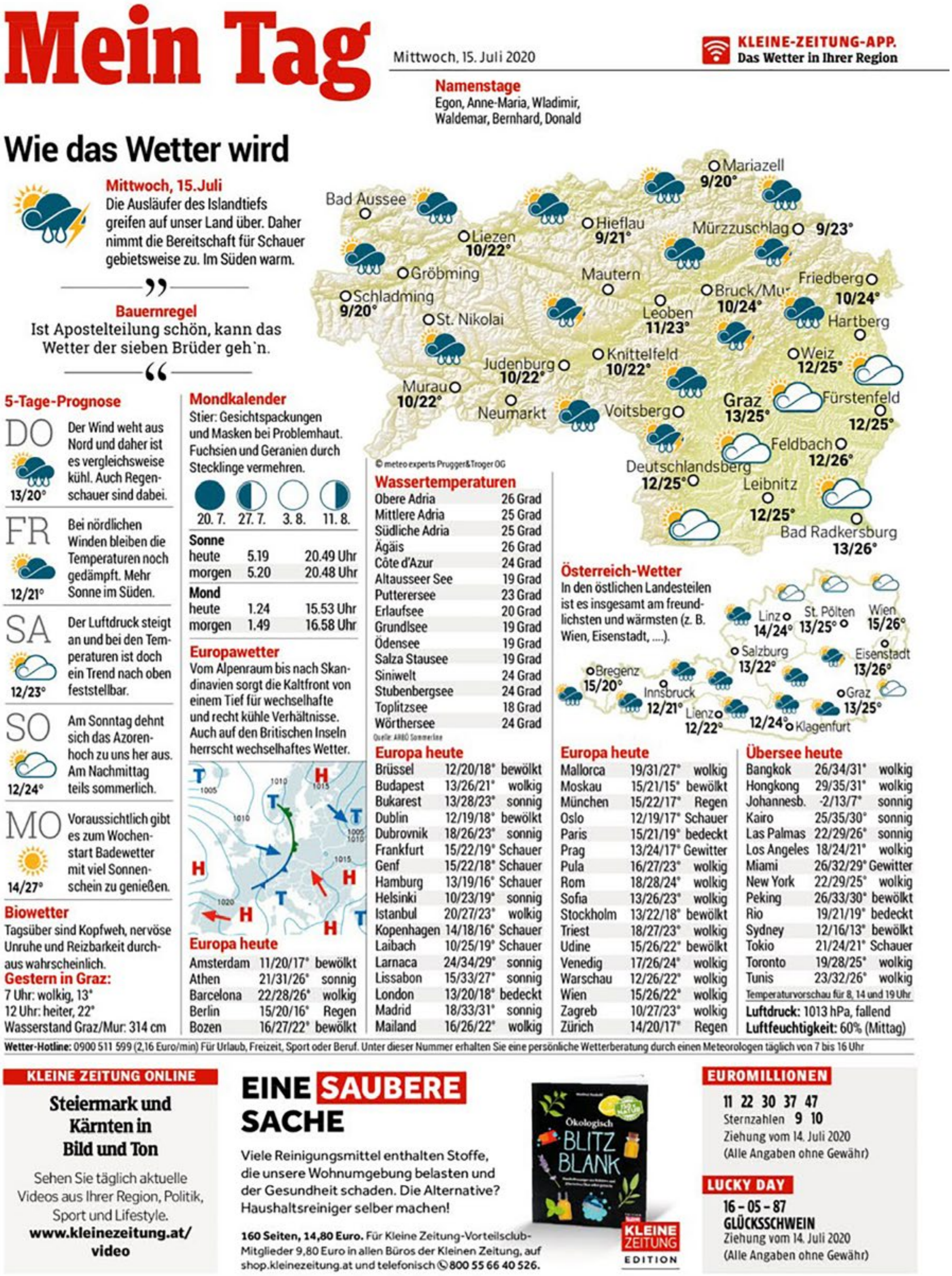
UBIMET's model development experts work to increase forecast precision through continuous evaluation and model refinement. Leading global weather models, such as ECMWF or GFS, and specially developed model output statistics methods are also used in a multi-model approach at UBIMET.

UBIMET has been supporting media companies in weather reporting and forecasting for many years. Its weather clients include major national and regional newspapers (e.g. 'Die Presse', 'der Standard', 'die Kronenzeitung', 'die Kleine Zeitung', Niederösterreichische Nachrichten').

UBIMET's high-precision and precise weather forecasts are prepared daily in several formats depending on the reporting medium. For daily and weekly newspapers, data and graphically appealing information – which can be individually designed – are prepared in UBIMET's weather editorial system (see example below). The possibility for customers to integrate weather widgets on their websites offers information about weather events in an additional form to online readers in an up-to-date and location-specific way. UBIMET also has its own studio for the production of weather films.

(234) Latent Heat Nudging is a data assimilation method

Sample data from an Austrian newspaper



Mein Tag

Mittwoch, 15. Juli 2020

KLEINE-ZEITUNG-APP. Das Wetter in Ihrer Region

Namenstage
Egon, Anne-Maria, Wladimir, Waldemar, Bernhard, Donald

Wie das Wetter wird

Mittwoch, 15.Juli
Die Ausläufer des Islandtiefs greifen auf unser Land über. Daher nimmt die Bereitschaft für Schauer gebietsweise zu. Im Süden warm.

Bauernregel
Ist Aposteltteilung schön, kann das Wetter der sieben Brüder geh`n.

Mariazell 9/20°
Bad Aussee
Hieflau 9/21°
Liezen 10/22°
Mürzzuschlag 9/23°
Gröbming
Mautern
Friedberg
Schladming 9/20°
Bruck/Mur 10/24°
Leoben 11/23°
10/24°
St. Nikolai
Hartberg
Knittelfeld 10/22°
Weiz 12/25°
Judenburg 10/22°
Murau 10/22°
Neumarkt
Voitsberg
Graz 13/25°
Fürstenfeld 12/25°
Feldbach 12/26°
Deutschlandsberg 12/25°
Leibnitz 12/25°
Bad Radkersburg 13/26°

5-Tage-Prognose

DO 13/20° – Der Wind weht aus Nord und daher ist es vergleichsweise kühl. Auch Regenschauer sind dabei.

FR 12/21° – Bei nördlichen Winden bleiben die Temperaturen noch gedämpft. Mehr Sonne im Süden.

SA 12/23° – Der Luftdruck steigt an und bei den Temperaturen ist doch ein Trend nach oben feststellbar.

SO 12/24° – Am Sonntag dehnt sich das Azorenhoch zu uns her aus. Am Nachmittag teils sommerlich.

MO 14/27° – Voraussichtlich gibt es zum Wochenstart Badewetter mit viel Sonnenschein zu genießen.

Biowetter
Tagsüber sind Kopfweh, nervöse Unruhe und Reizbarkeit durchaus wahrscheinlich.

Gestern in Graz:
7 Uhr: wolkig, 13°
12 Uhr: heiter, 22°
Wasserstand Graz/Mur: 314 cm

Mondkalender
Stier: Gesichtspackungen und Masken bei Problemhaut. Fuchsien und Geranien durch Stecklinge vermehren.

20. 7. 27. 7. 3. 8. 11. 8.

Sonne		
heute	5.19	20.49 Uhr
morgen	5.20	20.48 Uhr
Mond		
heute	1.24	15.53 Uhr
morgen	1.49	16.58 Uhr

Europawetter
Vom Alpenraum bis nach Skandinavien sorgt die Kaltfront von einem Tief für wechselhafte und recht kühle Verhältnisse. Auch auf den Britischen Inseln herrscht wechselhaftes Wetter.

T 1005 1010 H 1015 T 1010 T 1005 1010 1015 H H T 1020 H H T

Europa heute

Amsterdam	11/20/17°	bewölkt
Athen	21/31/26°	sonnig
Barcelona	22/28/26°	wolkig
Berlin	15/20/16°	Regen
Bozen	16/27/22°	bewölkt

© meteo experts Prugger&Troger OG

Wassertemperaturen

Obere Adria	26 Grad
Mittlere Adria	25 Grad
Südliche Adria	25 Grad
Ägäis	26 Grad
Côte d'Azur	24 Grad
Altausseer See	19 Grad
Puttersee	23 Grad
Erlaufsee	20 Grad
Grundlsee	19 Grad
Ödensee	19 Grad
Salza Stausee	19 Grad
Siniwelt	24 Grad
Stubenbergsee	24 Grad
Toplitzsee	18 Grad
Wörthersee	24 Grad

Quelle: ARBÖ Sommerline

Europa heute

Brüssel	12/20/18°	bewölkt
Budapest	13/26/21°	wolkig
Bukarest	13/28/23°	sonnig
Dublin	12/19/18°	bewölkt
Dubrovnik	18/26/23°	sonnig
Frankfurt	15/22/19°	Schauer
Genf	15/22/18°	Schauer
Hamburg	13/19/16°	Schauer
Helsinki	10/23/19°	sonnig
Istanbul	20/27/23°	wolkig
Kopenhagen	14/18/16°	Schauer
Laibach	10/25/19°	Schauer
Larnaca	24/34/29°	sonnig
Lissabon	15/33/27°	sonnig
London	13/20/18°	bedeckt
Madrid	18/33/31°	sonnig
Mailand	16/26/22°	wolkig

Österreich-Wetter
In den östlichen Landesteilen ist es insgesamt am freundlichsten und wärmsten (z. B. Wien, Eisenstadt, ...).

Linz 14/24°
St. Pölten 13/25°
Wien 15/26°
Salzburg 13/22°
Eisenstadt 13/26°
Bregenz 15/20°
Innsbruck 12/21°
Graz 13/25°
Lienz 12/22°
12/24° Klagenfurt

Europa heute

Mallorca	19/31/27°	wolkig
Moskau	15/21/15°	bewölkt
München	15/22/17°	Regen
Oslo	12/19/17°	Schauer
Paris	15/21/19°	bedeckt
Prag	13/24/17°	Gewitter
Pula	16/27/23°	wolkig
Rom	18/28/24°	wolkig
Sofia	13/26/23°	wolkig
Stockholm	13/22/18°	bewölkt
Triest	18/27/23°	wolkig
Udine	15/26/22°	bewölkt
Venedig	17/26/24°	wolkig
Warschau	12/26/22°	wolkig
Wien	15/26/22°	wolkig
Zagreb	10/27/23°	wolkig
Zürich	14/20/17°	Regen

Übersee heute

Bangkok	26/34/31°	wolkig
Hongkong	29/35/31°	wolkig
Johannesb.	-2/13/7°	sonnig
Kairo	25/35/30°	sonnig
Las Palmas	22/29/26°	sonnig
Los Angeles	18/24/21°	wolkig
Miami	26/32/29°	Gewitter
New York	22/29/25°	wolkig
Peking	26/33/30°	bewölkt
Rio	19/21/19°	bedeckt
Sydney	12/16/13°	bewölkt
Tokio	21/24/21°	Schauer
Toronto	19/28/25°	wolkig
Tunis	23/32/26°	wolkig

Temperaturvorschau für 8, 14 und 19 Uhr

Luftdruck: 1013 hPa, fallend
Luftfeuchtigkeit: 60% (Mittag)

Wetter-Hotline: 0900 511 599 (2,16 Euro/min) Für Urlaub, Freizeit, Sport oder Beruf. Unter dieser Nummer erhalten Sie eine persönliche Wetterberatung durch einen Meteorologen täglich von 7 bis 16 Uhr

KLEINE ZEITUNG ONLINE
Steiermark und Kärnten in Bild und Ton
Sehen Sie täglich aktuelle Videos aus Ihrer Region, Politik, Sport und Lifestyle.
www.kleinezeitung.at/video

EINE SAUBERE SACHE
Viele Reinigungsmittel enthalten Stoffe, die unsere Wohnumgebung belasten und der Gesundheit schaden. Die Alternative? Haushaltsreiniger selber machen!
160 Seiten, 14,80 Euro. Für Kleine Zeitung-Vorteilsclub-Mitglieder 9,80 Euro in allen Büros der Kleinen Zeitung, auf shop.kleinezeitung.at und telefonisch 0800 55 66 40 526.
Ökologisch BLITZ BLANK
KLEINE ZEITUNG EDITION

EUROMILLIONEN
11 22 30 37 47
Sternzahlen 9 10
Ziehung vom 14. Juli 2020
(Alle Angaben ohne Gewähr)

LUCKY DAY
16 – 05 – 87
GLÜCKSSCHWEIN
Ziehung vom 14. Juli 2020
(Alle Angaben ohne Gewähr)

© Kleine Zeitung GmbH & Co KG

The role of human intervention

The implementation of a weather forecasting system for a customer in the media sector has the following steps (here described for a newspaper), which build upon each other:

1. **Contracting between the client and UBIMET**: in this step, the elements of the daily weather forecast to be provided are defined, based on a pre-set list of options. Among the elements are:
    a. Forecasts for specific parameters (e.g. temperature, air pressure, humidity etc.)
    b. Localised weather parameters (district level and below)
    c. Images (e.g. weather maps, meteograms, etc.)
    d. Short texts explaining the weather forecasts
    e. Additional output channels. In addition to outputs for the print version of each issue of the newspaper, web-based versions and widgets can be ordered.
2. **Specification of the daily output to be provided to the client**: in this step, the whole model output (including images) are defined by the UBIMET team.
3. **Generation of weather forecasts**: Data will be generated with UBIMET's own model (RACE, see above). ML is used both in the stage of data assimilation (without human intervention) and in the post processing of raw model output data (e.g. quality control) (with human intervention).
4. **Automated generation of the final model output**: This step includes the automated generation of images and sets of forecast parameters as specified by each client in step 1.
5. **Final editing and data provision**: here, human intervention is needed as texts need to be generated by the UBIMET team explaining the overall outlook on the weather for the next day for each specific dimension.

Business Models and IP Rights

UBIMET is active in patents, but mainly in areas outside software development (e.g. 3D lightning detection). In the field of AI, it is difficult to obtain patentable results for meteorological applications. Usually only new algorithms are developed for already existing data sets. Thus, rights to existing IP are more likely to be secured here through common holdings and license agreements with customers.

The business model of UBIMET in the area of business-to-business (B2B) solutions is based on the development of individual forecasting solutions for its customers. In various sectors (e.g. energy, insurance, retail, food, etc.), solutions are developed at UBIMET by combining data feeds from the client with weather forecast data from UBIMET. The company sets its strategic focus in those domains where market prospects equally exist as specific customer know-how that has already been built up over a long period of time, which enables UBIMET to develop and offer tailor-made forecasting solutions for each customer.

UBIMET thus differentiates itself from large competitors who offer predictive analytics as "of the shelf" solutions with weather as input variable (e.g. SAP). Customers who buy such services, however, have to live with a worse forecast quality (about 30% higher error variance compared to customised systems) and therefore cannot gain any competitive advantage. Nevertheless, there is a clear trend towards such broadly offered solutions in B2B markets and this trend is likely to increase.

2.4.4. Future trends

If emerging trends in the field of AI for weather forecasting are considered, the main question is whether forecast models based on deep learning and trained using atmospheric data can compete with weather forecast and climate models based on physical knowledge and the fundamental laws of nature and equations of motion – and whether they could even be superior to them in the future.

The main drivers of this development towards DL and enhanced data-related training are, on the one hand, the availability of computer power and, on the other hand, the significantly improved availability of data (i.e. increasing number of meteorological observations), especially since the beginning of the satellite age. The plethora of data means that neural networks can be better trained and the use of AI will no longer be limited to parameterisation schemes and certain model components. Some in the scientific community believe that neural networks will be able to perform actual weather forecasts in the coming decade, using both observations of the past for training and observations of the present as input for the generation of forecasts. Currently established global weather prediction models that solve three-dimensional, non-linear equations could become obsolete as a result. It can also be assumed that neural networks will become much cheaper and easier to optimise for high-performance computing in the future compared to conventional models[235].

However, there is an essential caveat here: there is still a lack of evidence that neural network-based weather forecasting systems can stand direct comparison with conventional models. This applies on the one hand to global forecasts and on the other hand to forecast periods longer than 24 hours. An additional field in which AI first has to prove itself is the wide range of physical parameters provided by numerical models with physical consistency[236].

2.5. Journalism

2.5.1. Introduction

Algorithmic journalism can be understood as the software-supported or automated aggregation, production and distribution of content of any kind (data, text, images, audio, video). Applications from the field of AI – especially ML – are often used in this context. A general distinction can be made between:

1. *Assistive technologies*, which support journalists in the creation of media content.
2. *Generative technologies*, capable of creating media content largely autonomously and with very little human intervention.
3. *Distributing technologies* that also allow the mediation and publication, or other communication, of content generated with the help of algorithms. Different formats and formal requirements of different social media platforms can be addressed.

In the following, we will focus on automated journalism, as it is currently the most relevant *generative technology*, while detailed information on assistive and distributing technology and further applications of generative technologies is provided in **Annex IA**.

2.5.2. Automated journalism

How automated journalism currently works

Text generation is based on two main components: a strategic component, which usually uses AI search or planning strategies to determine which information is relevant ("content planning"), and a tactical component, which also determines the linguistic form ("language planning"). The quality of the generated text depends on the performance of the components as well as the combination of content and language planning[237].

The starting point for the use of text generation – so called "natural language generation" (NLG) - is a database or application interface providing constantly updated, thematically-relevant data. That data is structured by the

(235) https://www.geosci-model-dev.net/
(236) Ibid.
(237) Hess M. (2005). Einführung in die Computer Linguistik

software and organised in a so-called "text plan"[238]. Within the framework of language generation, text criteria such as text length and tonality are also defined. Algorithms then select relevant elements from a data set and determine the linguistic structures of the text. In the last step, a trained algorithm selects the words to be used and, based on grammatical rules, puts them in the correct order based on syntax.

Using NLG technologies, it is possible to create journalistic texts largely automatically. With the exception of feeding the relevant data set (if the application does not have an interface), defining the text criteria (length, tonality, etc.) and, if necessary, quality control, the text generation and publishing process usually take place without human intervention[239].

In practice, automatically-generated texts are currently used primarily for standardised messages and notices for which current data or unique information can be accessed on a rule-based basis. This is due to the fact that current applications mainly refer to structured data in machine-readable format. A human editor writes a template (for example, an advertising text for a mobile phone), then links words or phrases from this text to the columns of a pre-designed machine-readable table. The application converts each table row into text.

Based on statistical methods, current information is inserted into text modules pre-written by humans. The modules are reusable text fragments, which makes text automation particularly suitable for journalism topics based on reporting information, such as sports, finance, weather or events where information and terms have a high repetition rate from one report to another.

Unlike human journalists, algorithms are able to systematically avoid repetition, thus reducing the risk of recurring formulations or linguistic boredom[240]. Modern applications also have a good command of grammar, based on rules programme by humans, enabling them to find the right articles and endings for the contents of the data records. The software does not learn to write by throwing thousands of texts at it but has to be programmed with rules, ML playing a secondary role. The set-up effort can be reduced as the applications analyse data records and associated texts and thus learn which data fields belong to which positions in the text.

Technology providers

Recent analyses of the market identified fourteen companies in different countries that offer technology solutions for automated content creation[241]. It is noteworthy that these enterprises do not see themselves as journalistic organisations or media companies; neither do their names indicate a relationship to journalism; nor are their products specifically geared toward providing journalistic content The technological solutions offered by these companies are generic in nature and can be applied to data from all industries. Major applications range from writing product descriptions to preparing patient summaries in hospitals.

Of those fourteen, five are based in Germany (AX Semantics[242]; Text-On[243]; 2txt NLG[244]; Retresco[245]; Textomatic[246]); three in France (Syllabs[247]; Labsense[248], Yseop[249]) two in the United States (Narrative Science; Automated Insights); and one each in the United Kingdom (Arria[250]), the Russian Federation (Yandex[251]), Spain (Narrativa[252]) and China (Tencent[253]). Ten of the companies offer automated content creation in only one language, while four offer their

(238) Bateman, J., Zock, M. (2003). Natural language generation. in: The Oxford Handbook of Computational Linguistics, p. 5f.
(239) Tusch, R. (2017). Medienforscher über Roboter-Journalismus: ‚In der Massenproduktion schneiden Maschinen besser ab als Menschen, online: http://meedia.de/2017/03/20/medienforscher-ueber-roboter-journalismus-in-der-massenproduktion-schneiden-maschinen-besser-ab-als-menschen/
(240) Ibid.
(241) Dörr, K. (2016). Mapping the field of Algorithmic Journalism
(242) https://en.ax-semantics.com/
(243) https://text-on.de/
(244) https://2txt.de/
(245) https://www.retresco.de/
(246) https://www.textomatic.ag/
(247) https://www.syllabs.com/fr/
(248) https://www.lab-sense.com/
(249) https://www.yseop.com/
(250) https://www.arria.com/
(251) https://yandex.com/company/
(252) https://www.narrativa.com/
(253) https://www.tencent.com/

services in multiple languages[254]. The algorithms of the German company AX Semantics, for example, work in up to 110 languages[255].

News companies that have developed internal solutions for the generation of automated news[256] include: Xinhua News Agency (China), Tencent (China), MittMedia/United Robots (Sweden), NTB/Bakken & Baeck (Norway), Washington Post (USA), Los Angeles Times (USA), Bloomberg (USA), Thomson Reuters (USA), Austria Press Agency (Austria), and the Berliner Morgenpost (Germany). The Chinese news agency Xinhua, for example, launched the "Kuaibi Xiaoxi" project in 2015, which translates as "Little Xinhua who writes fast" and is used to automatically generate sports and financial reports[257].

Automated News or "robot journalism"

While automated news is a rather young field, its basic idea has a long history starting in 1963 with the BASEBALL programme. The principle was that users could ask BASEBALL questions about baseball games, receiving simple answers as information.

In a study published in 1970, a simple algorithm was reported that generated weather forecasts comprising four parts. In 1992, this capability was extended in the form of the "Forecast Generator", which led to longer weather forecasts, which were eventually published[258].

The application of AI for the generation of news is particularly interesting in cases where the data basis is sufficiently large and of good quality and where the AI is therefore provided with a large amount of text on which it can train. Automated news journalism or "robot journalism" is especially useful for clearly structured texts that are constantly updated, e.g. for very localised (hyper-local) topics. Robot journalists can produce texts on topics that human colleagues are unable to due to their high specificity and the need for too large a workforce (e.g. to write dozens of articles a day on particulate matter pollution in specific areas). The BBC has been quoted as predicting that, by 2022, about 90 per cent of all news texts will be written by robots, but that this does not mean that 90 percent of what we read will be texts produced by robots[259]. Automatically-created texts will be found mainly in niche markets. For example, currently no-one reports lower level matches in a little known sport but, in the future, AI may fill that gap.

In Europe, the British newspaper **The Guardian** has been among the pioneers in using automated content creation. Already in 2010, this media corporation carried out two projects on automated sports news focussing on statistics of games and historical information of teams and players. To compose short stories, they combined the data with standard phrases and connectors[260]. In 2014, they started an AI application called "Guarbot" that complemented financial information with complex data, avoiding the need to use journalists for this task.[261]

Since 2016, **textOmatic** has been publishing numerous automatically-generated weather news reports for all major cities and regions in Germany via **FOCUS Online**[262]. Since the beginning of 2018, the cooperation has been further expanded as part of a premium partnership between FOCUS Online and textOmatic and, since May 2018, financial news has also been generated for FOCUS Online, based on the information in the **Handelsblatt**[263].

MittMedia launched the so-called "Homeowners Bot[264]" for its local newspapers at the end of 2017. It has only one purpose: to collect local information on the property market and turn it into articles. The articles tell the reader who has sold and bought - and at what price. Readers receive a 'push notification' of a purchase or sale in the

(254) Graefe, A. (2016). Guide to Automated Journalism, New York City: Columbia Journalism Review.
(255) https://en.ax-semantics.com/about-ax-semantics
(256) Reichelt, P. (2017). Einführung in den Roboterjournalismus. Bedrohung oder Chance?, Baden-Baden: Tectum.
(257) http://www.scmp.com/news/china/article/1876422/think-robot-can-write-pahchinas-state-news-agency-begins-publishing-work.
(258) Reichelt, P. (2017), Einführung in den Roboterjournalismus. Bedrohung oder Chance?, Baden-Baden: Tectum.
(259) https://www.rolandberger.com/fr/Point-of-View/Journalists-vs-robots-writing-with-AI.html
(260) Bunz, M. (2010). In the US, algorithms are already reporting the news. The Guardian, 30 March. https://www.theguardian.com/media/pda/2010/mar/30/digital media algorithms reporting journalism
(261) Gani, A., Haddou, A. (2014). Could robots be the journalists of the future? The Guardian, https://www.theguardian.com/media/shortcuts/2014/mar/ 16/could robots be journalist of future
(262) https://textomatic.info/News/detail.191.html
(263) https://www.handelsblatt.com/boersenberichte/textomatic-boersen-informationsdienst/
(264) https://medium.com/mittmedia/the-homeowners-bot-36d2264e2d34

neighbourhood, depending on their geographical interest. Since its launch in September 2017, the "Homeowners Bot" has produced an enormous 34,000 articles - 480 per week - which have contributed significantly to attracting new subscribers. According to DigiDay, MittMedia has, through robot journalism, gained a thousand new subscribers willing to pay ten euro a month to follow the property market[265].

AI is already being used particularly intensively by the **Associated Press news agency**[266], for example, to report on developments in the share price of companies or to summarise quarterly balance sheets. For this purpose, corresponding text modules have been pre-formulated by editors. An algorithm combines these modules plausibly and adds the corresponding figures. Every three months, up to 4,000 reports are written using the programming. AP is aiming to have up to 80 percent of the articles produced this way in future., leaving more time for human employees to focus on their other activities including, for example, investigative journalism[267].

The editorial team of the **Stuttgarter Zeitung,** together with the **Stuttgarter Nachrichten,** publishes the daily "Particulate Matter Radar[268]", a further example of how media companies, and especially regional media, can benefit from the use of AI. Readers can view online the particulate pollution levels in certain residential areas. To do this, the service accesses weather data and data from hundreds of particulate matter sensors that private individuals have hung up and which produce user-generated content. An algorithm provided by **AX Sematics** uses the information to generate texts, a forecast in the morning and a report on actual air pollution in the evening.

In June 2019, the **German Football Association** (DFB)[269] presented a development project for nationwide coverage of amateur football, made possible by automated text creation. The AI-based technology behind it comes from the Berlin company **Retresco**, with project partner is **Sportplatz Media**[270]. With the start of the 2019/2020 season, FUSSBALL.DE, Germany's largest amateur football platform operated by the DFB, has started to publish automatically-generated pre- and post-match reports on league matches in the men's, women's and older youth sectors nationwide, right down to the lowest divisions. The DFB is thus following the request of many amateur clubs for regional reporting on FUSSBALL.DE. The basis of the match reports are the available official match data from DFBnet. NLG creates a natural language text from relevant data such as match results, goal scorers and substitutions.

The Dutch news agency **Algemeen Nederlands Persbureau BV** (ANP) uses AI differently. Here, the so-called "Natural Language Generation (NLG) Systems" edit the texts of human colleagues to make them more knowledgeable or even more understandable[271].

A study by the Ludwig-Maximilians-Universität (LMU) in Munich found that readers rated robot texts as particularly credible because they contained a particularly large number of figures. In 2018, the researchers asked around 1,000 people about the quality of robotic articles[272]. The study participants were asked to evaluate the extent to which different texts, created either with or without "AI", were convincing. In fact, the texts written by bots scored very well although sometimes robot texts read the same and are monotonous, lacking complexity and creativity. These are reasons why the editors are sometimes sceptical about them. In addition, algorithms can be led to make mistakes. For example, in 2017, the Los Angeles Times reported on an earthquake that had actually occurred in 1925. Its editorial staff were using a bot to write reports on earthquakes and, in this case, scientists at the California Institute of Technology had updated historical earthquake data and supplied the algorithm with incorrect information[273]. Such errors can have considerable consequences.

Albeit there are plenty of interesting examples, experts have expressed the view that it does not look like technology is going to be a disruptive force in the media industry. Few interesting stories are based on data alone, even in sport and finance. Where automation is possible, it is unlikely to destroy jobs, but rather to expand coverage. In

(265) https://digiday.com/media/swedens-mittmedia-increases-subscriber-conversions-20-time-wall/
(266) https://www.ap.org/discover/artificial-intelligence
(267) https://www.tcs.com/perspectives/articles/why-artificial-intelligence-is-a-big-part-of-APs-Future
(268) https://www.stuttgarter-zeitung.de/feinstaub
(269) https://www.dfb.de/
(270) https://sportplatz-media.com/
(271) https://www.anp.nl/over-ons-anp
(272) Graefe, A., Haim, M. (2018). Automatisch interessant? Der Einfluss von Involvement auf die Wahrnehmung computergenerierter Texte, In: Rössler, P. / Rossmann C. (Eds.): Kumulierte Evidenzen. Replikationsstudien in der empirischen Kommunikationsforschung. Wiesbaden: Springer Fachmedien 2018, S. 189 –206.
(273) https://laist.com/2017/06/22/quakebot_error.php

newspapers and broadcasts, practically no one converts tables into continuous text on a full-time basis: at most, some employees do so in agencies. AI is having a faster effect on other industries through the media: for example, the retail sector. Up to now, many online shops have been putting manufacturers' product data online unchanged. However, if an online-shop wants to climb up in Google rankings, it needs texts that no other retailer has, a task suited to text creation software. In a few hours, an editor can write a standard text, assign a table with product data, and the software writes hundreds of thousands of texts by itself. This makes traders such as Home24[274], Euronics[275] and Hornbach[276] customers already of companies such as AX Semantics.

2.5.3. Case study – The role human intervention in automated text generation (Retresco)

The company

Retresco is an expert in the automated exploitation of content and data. Using semantics and AI, Retresco automates the communication of its customers and their work with content. Retresco's technology is used by Axel Springer, FAZ.net, United Internet, Simpleshow and the German Federal Ministry of Health, amongst others. Retresco has over 70 employees, and the number is growing, and it is active in several European countries.

Alexander Siebert is the founder, managing director and creative director of Retresco. Before his passion for both data and voice led him to create his own company in 2008, he worked at the Berlin-Brandenburg Academy of Sciences and Humanities. Johannes Sommer has been Managing Director at Retresco since 2013 and, in addition to being responsible for sales and marketing, is responsible in Retresco for activities on media and the digitalisation of the publishing industry.

Cost and time pressure is a significant issue in all industries. Working with content is complex, time-consuming and expensive. Retresco's technology can replace many editorial and communication activities, making it a key technology in any digitisation strategy. In media, technology is increasingly replacing editorial activities (e.g. for weather and sports reporting). Technology also supports the identification of topics and the automatic compilation and playout of content for target groups. In the financial sector, Retresco technology is used primarily for generating asset performance reports. Based on extensive data and its analysis as well as AI, Retresco generates fund, portfolio and investment reports, as well as index and individual value reports, in various languages in real time.

With its self-developed software, Retresco takes over the evaluation and analysis of large amounts of data and generates reports, analyses or instructions for action from the information obtained. This content is efficiently created, informative and generated in real time and can be customised for each individual user.

The Technology

Basic Description

Natural Language Generation (NLG) is a type of NLP and refers to the creation of natural language texts by a software. In a nutshell, NLG is the process of adopting structured data and converting it into text. Wherever structured data is generated - for example, in e-commerce, on the stock exchange or in reporting for sports, business or weather - NLG software can create reader-friendly texts from data.

Data is the foundation of an NLG application. For texts in natural language, templates and conditions must also be defined in advance. The templates are essentially pre-formulated sentences that are individualised with the help of data and lexicalisation algorithms. Conditions here are the circumstances that must be fulfilled for the template to be used.

In the creation process, the software arranges the templates in a certain "narrative" order. The given sequence in which the templates are arranged is called "storyplot". NLG can also include "Big Data" in the text creation

(274) https://www.home24.com/
(275) https://www.euronics.de/
(276) https://www.hornbach.de/

process. Depending on the input data and the complexity of the defined conditions, NLG is thus able to present facts or interpret numbers in detail.

Two approaches have become established in the concrete implementation of NLG projects. If the requirements are standardised and the users do not have any programming skills, a SaaS solution is the best choice. A web-based NLG platform enables the connection to a data source via API, the formulation of text modules and conditions as well as the connection of the output, for example to a CMS.

Alternatively, there is a supervised service, where a team of experienced developers, data experts and linguists adapts and implements the NLG project to specific needs.

Technology development and learning

RETRESCOs NLG solutions have all been developed solely in-house since the start of the company in 2008. The starting points have been developments in the domain of natural language understanding (NLU) for clients in the German media sector such as 'Frankfurter Allgemeine Zeitung' or 'Die Zeit'. Automated summaries have also been sold to other sectors (e.g. financial auditing companies, etc.) Based on these developments, RETRESCOs NLG technology has been developed continuously. In the beginning, automated texts were generated in-house for customers, but the more developed its NLG solutions became, the more text generation was moved into the domain of its clients.

Core to RETRESCOs NLG technology are customer-specific models that reflect the characteristics of each customer's domain. In each case, NLG technology comprises the use and analysis of structured data and automated generation of text, with applications reaching from automatically-generated descriptions for online-retailers to medical doctors' reports in the information system of large hospitals.

The core NLG technology is continuously improved and developed further, on the one hand, by means of internal projects and, on the other hand, through external projects with new or existing customers.

The use-case of FUSSBALL.DE

Basic description

FUSSBALL.DE has been operated by the German Football Association (DFB) since July 2014 and is the largest amateur football platform in Germany. Last year, it set a record with 4.5 billion page impressions and 337 million visits. The FUSSBALL.DE app has been downloaded 3.5 million times. The core tasks of FUSSBALL.DE are to increase the visibility of amateur football, to support the clubs in their public perception, to provide content services to the grassroots and to make the action on amateur sports' fields tangible online. Therefore content created with NLG is given for free to local newspapers.

Since the start of the 2019/2020 season, the DFB's, automatically-generated pre- and post-match reports on league matches in the men's, women's and older youth sectors have been appearing on the portal FUSSBALL.DE- nationwide, right down to the lowest divisions. Every weekend, up to 75,000 texts about amateur football are produced based on AI and published on FUSSBALL.DE.

The basis of the match reports is the available official match data from the DFBnet, a networked IT system for organisational tasks in sports associations and clubs. The DFBnet match report is the electronic version of the official match report sheet, on which team line-ups and match events such as substitutions, goals or penalty cards are noted. The software, which was originally developed for the Bundesliga, has increasingly replaced the paper form in the amateur sector since 2006. Through the integration with the modules DFBnet Match Operation, DFBnet Referee Scheduling and DFBnet Pass, the match date, venue, referee and player are all taken into account. The home and visiting teams can put together their squad online in advance; the data is only visible to the opponent after approval and shortly before the start of the match.

NLG forms a text in a natural sounding language (here German) from relevant data such as match results, goal scorers and substitutions. The technological heart of the project is the NLG platform textengine.io from Retresco.

The texts are displayed on FUSSBALL.DE and in the FUSSBALL.DE app on both the Team page, the Game Details, and the League page marked with a small symbol. All users who have registered on FUSSBALL.DE and set up a personalised homepage can also see the robot texts under "My News".

Example of an automatically generated match report text[277]

> *"match report | 25.03.2019 | 05:04*
>
> *SV Fellbach's success story continues*
>
> *Landesliga 1: TSV Pfedelbach - SV Fellbach, 0:3 (0:1), Pfedelbach*
>
> *Last Saturday TSV Pfedelbach lost clearly with 0:3 against SV Fellbach. In the run-up to the match, it was agreed that the winner could only be SV Fellbach. The course of the 90 minutes finally confirmed this impression.*
>
> *Filip Jaric brought SV Fellbach forward in the twelfth minute. In the time up to the break no further goal was scored, it remained the lead for the guest. Martin Wöhrle of TSV Pfedelbach made a change at the restart: Kaan Uzuner stayed in the dressing room. Michael Blondowski came in his place. Jaric tied a double with his second goal and put his team on the winning track (65th). The 170 spectators witnessed Tim Lück "aiming" at the wrong goal and overcoming his own goalkeeper. The win was in the bag and coach Giuseppe Greco gave Altin Gashi a little time to play. In return Jamie Miller left the field (180.). In the end, SV Fellbach claimed a foreign victory and put TSV Pfedelbach in their place 3-0.*
>
> *This defeat caused TSV Pfedelbach to drop back to tenth place in the table.*
>
> *SV Fellbach have been undefeated for four games. The season balance of the league primus thus continues to look very positive. With eleven wins and six draws, SVF have lost only three matches. SV Fellbach is following the old football wisdom that attack wins matches, but defence wins championships. With just 24 goals to their name, SV Fellbach are the best defenders in the Landesliga 1. Who else can stop SV Fellbach? SV Fellbach scored the next three points against TSV Pfedelbach and continues to lead the Landesliga 1 field.*
>
> *Next up for TSV Pfedelbach is an away game. On Sunday (16:00 hrs) they will play TSV Schwaikheim I. SV Fellbach will play SV Allmersbach two days earlier (19:30 hrs)."*
>
> *Author: FUSSBALL.DE (This text was automatically generated on the basis of the official match data available to the DFB on 25.03.2019 at 05:04 am)*

The role of human intervention

The role of human intervention needs to be considered separately in the two stages of automated text generation within FUSSBALL.DE.

In the development stage, there was a clear role for humans. This included the development of templates, definition of specific conditions, and the training of the machine. All this was done by (human) editors with good domain knowledge of sports journalism and, in particular, football.

Training of the system is based on the data from past games. The machine uploads data into sentences that it has been taught. One could say that the machine "only" puts text blocks together, but there are many text blocks

(277) Source: www.fussball.de

stored in the programme according to the “if, then” logic that must be properly applied in the reportig. So, for example, if team A has won with a goal difference x:y against team B:

1. ‘SV Fellbach’s success story continues’ should be used as headings.

2. ‘Last Saturday TSV Pfedelbach lost clearly with 0:3 against SV Fellbach’ as an introduction to the sentence.

3. ‘In the run-up to the match, it was agreed that the winner could only be SV Fellbach’ as second sentence.

Even the conditions for using individual adjectives must be “trained” (e.g. at what goal difference can the robot call a win or loss “clear”?). The machine will only be able to decide this in the future – and automatically – once the editor has determined it.

Things look significantly different once the implementation stage has been reached. Here the whole process of text generation runs without any human involvement. The process of databased automated text generation has the following steps:

1. Data-feed on individual games: The system is able to cover between 70,000 to 80,000 games on a match day. The data-feed is provided by DFB within its DFBnet system.

2. Analysis: Here trends are identified automatically on the basis of the data-feed. This means that issues are covered such as “the favourite in this game” or “the player most likely to score for the most goals”. This analysis is fully automated.

3. Generation of the text: The text for each game is generated automatically in 150 to 200 milliseconds, based on the defined template and conditions.

In such a way, pre- and post-match reports are generated and published for each game. This is done without any further human control or editing. Furthermore, links and/or photos are added automatically to the published text.

Business Models and IP rights

RETRESCO has not filed patents on its AI technologies. Its know-how is protected partly through licensing models for its core NLG technology. In addition, large amounts of tacit knowledge are involved in the development and training of customer-specific models, forming high barriers for competitors to enter the market. Furthermore, long-term cooperation leads to a lock-in effect with high switching costs for the client to change the partner for NLG applications[278].

RETRESCO currently uses two licensing models for its core NLG technology. Either – as in the case of FUSSBALL.DE – a complete custom-tailored solution is developed and transferred to the client, or the NLG platform itself is licensed to a client who designs and implements their project autonomously.

2.5.4. Future trends

Fully automatically-generated content

Most recent developments point in the direction of a greatly automated process of content generation. In addition, technologies such as speech-to-text and face recognition from images are on the way to vastly improving the workflow speed of journalists and to adding value to their existing archives by making them machine-searchable in new ways[279].

(278) It is noteworthy that German companies active in NLG technologies do all follow a similar strategy to protect IPRs, while Arria (UK), Yseop (FR), Yandex (RU), Narrative Science (US), and Automated Insights (US all have filed specific patents at the USPTO.

(279) Fanta, A. (2018). Putting Europe’s Robots on the Map: Automated journalism in news agencies

Monok[280] (Sweden) runs a news generation system that is an extractive-abstractive multi-document multi-modal summariser, combining neural pre-processing for co-reference, named entities, etc., with a coherence-preserving, rule-based extractor, and a neural abstractor (which rewrites the text in order to create a sufficiently new piece of writing). The resulting articles are validated via third-party plagiarism detectors, and augmented with scraped videos, social media embeds, Wikipedia links, etc.

Radar AI[281] (UK) uses a system based on ArriaNLG Studio[282] with a template-scripting language used to control text generation from structured data. As input, RadarAI uses periodic-release public data sets (e.g. health data from the NHS). For each of these, its editors write templates structured into decision trees that control which templates are used, which data is selected and where it is inserted, depending on date, identity and location of the client (e.g. a national or local newspaper), etc. After this initial one-off effort, article generation is fully automatic for a given data source.

Both Monok's and Radar's articles are published without human post-processing (although in the latter case, editors add local information), indicating a level of quality that readers find hard to distinguish from that written by humans. While the industry still assumes "no one wants AI to fully replace humans in the newsroom, and no one is actively working towards that end"[283], start-ups like Monok and Radar already show that full AGN is about to enter general news production.

At the beginning of 2019, the American OpenAI Foundation, which deals with the potential dangers of artificial intelligence, made the headlines when researchers developed a language AI (called GPT-2) able to independently write journalistic and literary texts. Although OpenAI always releases its own developments as open source, GPT-2 has remained protected, with only a greatly reduced version having been published to date. The creators call the AI a "deep fake for texts", able to create articles that can no longer be clearly distinguished from those written by humans. For example, the AI could write pseudo-news stories with invented quotations, or otherwise provide journalistic commentaries with plausible chains of argumentation[284]. In 2020, the American OpenAI Foundation released GPT-3, which offers further applications such as layouts of websites and development of roleplays.

Automated text generation as Software as a Service

As with pharmaceutical research (see above at 2.3.8), an increasing trend on the market in this journalism area is the offering of automated text generation with structured data as SaaS (or in this case AIaaS) solution to clients. Several of the fourteen companies currently active on the market already offer such a service to their customers[285]. Nevertheless, it is noteworthy that such services do not work for all kinds of texts. NLG as an "of-the-shelf-solution" for customers is currently, and in the foreseeable future, not possible for texts that also include structured data, as these might have very heterogenous formats depending on the specific application domain. There are, for example, significant differences between the product description for a washing machine to be sold in an online store and a medical doctor's report in a hospital. While freely accessible online translation services mainly rely upon large bodies of terms that are further developed by crowdsourcing through users, this is not the case for automated text generation that builds on templates, small text pieces and a set of common rules.

2.6. Conclusions

The information and evidence collected on the use of AI technology in **pharmaceutical research** suggest that, although AI has the promise of becoming an important general purpose technology to facilitate drug discovery, its principal use is in facilitating the earliest stages in the drug development process where tasks rely heavily on automatic data processing and reasoning. Currently, beyond possible assistance in identifying patients, it does not seem to have substantive effects on helping the drug through the expensive and risky clinical trials that require substantial human engagement and intervention.

(280) https://www.monok.com/
(281) https://app.radarai.com/
(282) https://www.arria.com/studio/studio-overview
(283) Chiusi, F., Beckett, C. (2019). The journalism AI global survey: What we've learned so far.
(284) https://openai.com/blog/better-language-models/
(285) See e.g. https://en.ax-semantics.com/the-tool or https://www.arria.com/studio/studio-overview/

Human intervention is still needed in several steps of AI-augmented target identification, *de novo* design of new compounds, and lead identification. Due to the complexity of the work, tacit domain knowledge is still necessary to set initial parameters or specific weights in modelling tasks based on ML. In addition, the selection of molecules for further testing and optimisation is still subject to human decision. In the views of domain experts, AI is seen only as a technological input in the invention process.

In **meteorology**, an area of science in which AI technology is the most developed and used, AI is currently mainly finding applications in the correction of observation errors, data assimilation, and post-processing of meteorological data. These tasks usually run in an almost fully automated manner with only very limited human intervention. This also holds true for applications for ML in numerical weather prediction models. However, many weather companies still provide a final post processing step for clients involving human intervention, such as writing of texts to explain forecasts and editorial work.

While there are several potential application areas for AI in **journalism**, only automated journalism is currently showing many existing uses. Tasks performed by AI include, in particular, the reporting of sports, weather and finance. Local news is also already an important application area. AI systems can produce texts on topics that humans can or do not cover because they are too specific and resource intensive. Nevertheless, human intervention is still necessary in the setting up and supervising such systems. Nevertheless, most recent developments point in the direction of a greatly automated process of content generation.

The biggest barrier to a faster development of the market for automated text generation is the low degree of digitisation of the EU. Even European media companies have to buy data from external sources, which leads currently to only very limited promising scenarios for commercially successful application of automated text generation (e.g. sports, traffic news, weather forecasts etc.). Accordingly, “robo-journalism” offering more than a presentation of text and data seems to have no real prospects in the next five to ten years’ time.

The following sections build on the introduction and the state of the art review above to carry out the legal analysis of the IP protection of AI-assisted outputs that is at the heart of this assessment study. We first examine challenges to such protection in EU copyright law (Section 4) and then European patent law (Section 5).

LEGAL ANALYSIS[286]

3. Artificial intelligence and copyright law

3.1. Outline, research question for copyright law

This section examines the question whether, and under which conditions, AI-assisted output[287] qualifies as "works" protected under copyright, and how authorship and copyright ownership are allocated in respect of such works. In addition, this section queries whether such outputs might benefit from protection under existing regimes of related rights, notably the database right and phonographic rights.

The section is mainly descriptive in that it analyses these – closely related – questions applying to current law. Whereas copyright in the EU Member States is still governed mostly by national law, the extensive state of harmonisation of the law of copyright in the EU, in particularly with regard to the concept of the work of authorship, justifies examining these issues primarily through the lens of harmonised EU law. Nevertheless, national law, doctrines and case law from various Member States are occasionally discussed whenever relevant. Moreover, the Report regularly references relevant law and doctrine from the United States as well.

To begin with, in , we briefly examine the international law of copyright, notably the Berne Convention, which although not formally part of the EU legal order can be considered quasi-*acquis* because of its adherence by all Member States of the EU. In 3.3, we present an overview of the current framework of EU copyright law as it pertains to copyright protected "works" and issues of authorship. In 3.4, we discuss core question whether – and under which conditions – AI-assisted outputs might qualify as "works" protected under harmonised copyright standards. To this end, we closely scrutinise the notions of "work", "originality" and "creative freedom", as developed in the decisions of the CJEU. As is demonstrated, the CJEU's decisions in *Painer* and *Football Dataco* are of particular relevance and offer important clues in answering this question.[288] In 3.5, we address the issue of authorship of AI-assisted outputs, which is inextricably linked to that of the notion of "work". Nevertheless, since this notion has so far remained largely unharmonised, nationally divergent solutions, such as the rules on authorship of computer-generated works in UK and Irish copyright law, have been developed.[289] 3.6 examines to what extent, and under what conditions, existing EU rules on the protection of related rights (notably, the rights of phonogram producers, broadcasters and news publishers) and on *sui generis* database rights might adequately protect AI productions that lack originality or authorship. This additionally queries whether an extension of current related rights to (better) protect "authorless" AI productions might be justified. The section closes with concise case studies on the priority domains of science and media, where we apply our analytical framework and findings to AI-assisted outputs in weather forecasting and automated journalism (3.7).

3.2. The international copyright framework

The three main international treaties in the field of copyright (Berne Convention, WCT and TRIPS) oblige their contracting states – currently, 179 countries, including all EU Member States – to provide international protection of eligible authors of "works", according to the minimum standards of these treaties. While the Berne Convention

(286) The legal analysis (copyright and patent law) was authored by P. Bernt HUGENHOLTZ, João Pedro QUINTAIS and Daniel GERVAIS (IViR – University of Amsterdam) and reviewed and commented by the Study Team, final editing being undertaken by JIIP.

(287) See the definition of AI-assisted output" above at section 1.3. of this report.

(288) Case C-145/10 Eva-Maria Painer (2011) ECLI:EU:C:2011:798 (*Painer*); Case C-604/10 Football Dataco Ltd and Others v Yahoo! UK Ltd and Others (2012) ECLI:EU:C:2012:115 (*Football Dataco*).

(289) See UK Copyright, Designs and Patents Act 1988 (as updated), s178 (on minor definitions: "computer-generated", in relation to a work, means that the work is generated by computer in circumstances such that there is no human author of the work"); Copyright and Related Rights Act 2000 (Ireland), ss 2 ("computer-generated", in relation to a work, means that the work is generated by computer in circumstances where the author of the work is not an individual").

creates obligations for the Member States only, the EU is directly bound by the TRIPS (part of the WTO package of international trade agreements) and the WCT.

Art. 2(1) of the Berne Convention (Paris Text) defines "literary and artistic works" in very broad terms to "include every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression", leaving ample room for work categories not enumerated in the lengthy list of examples in the remaining part of this provision. The TRIPS and the WCT incorporate the Berne Convention's definition of "literary and artistic works" by reference.[290] Art. 2(1) Berne Convention therefore is indirectly integrated into the EU legal order.[291]

The Berne Convention repeatedly speaks of protection offered to "original" works,[292] but offers no definition of the requirement of originality,[293] contrary to a number of European directives. Nevertheless, it is generally assumed that the Convention's definition of a work implies a requirement of human intellectual effort or creativity.[294] This prerequisite, however, does not exclude intellectual productions made with the aid of machines. For example, among the many work types mentioned in art. 2(1) of the Berne Convention are photographic works as well as "cinematographic works to which are assimilated works expressed by a process analogous to cinematography", in other words audio-visual works (films).

The Berne Convention is largely built on the continental-European system of authors' rights, as is apparent inter alia in its art. 2(6): "protection shall operate for the benefit of the author and his successors in title".[295] The Berne Convention does not define the "author" of a work, leaving this to the contracting parties, but its text and historical context strongly suggest that "author" and "authorship" for purposes of the Convention refer to the natural person who created the work.[296] This implies that copyright protection initially vests in *human* authors.[297] This is confirmed by the Convention's ubiquitous references to the "author" as the originator of works and the beneficiary of protection. The Convention's provision on moral rights (art. 6*bis*) that are expressly granted to "authors" underscore that its minimum standards of copyright protection are triggered only by acts of human creation. Consequently, the Convention does not in any way impose upon contracting states an obligation to accord copyright protection to AI-assisted outputs that are not the result of any act of human authorship.[298]

Human rights provide additional arguments in support of the proposition that copyright presupposes human authorship.[299] For example, the Universal Declaration on Human Rights (UDHR) protects the moral and material interests of authors resulting from scientific, literary or artistic production. Given that human rights by definition vest in human beings, the concept of authorship under the UDHR necessarily refers to human authorship.[300]

The Berne Convention does not deal with authorship of joint or collective works, nor does it prescribe who qualifies as (co)owner. For film works, there are special ownership provisions in art. 14*bis*, which were introduced at the 1967 Stockholm revision conference.

3.3. The EU copyright framework

The current EU copyright framework is mostly silent on questions of copyright subject matter and authorship. Despite extensive copyright harmonisation, no single directive harmonises the concept of the work of authorship

(290) *See* art. 9(1) TRIPS and art. 3 WCT.

(291) *See e.g.* Case C-277/10 Martin Luksan v Petrus van der Let (2012) ECLI:EU:C:2012:65 (*Luksan*), para. 59, and Case C-310/17 Levola Hengelo BV v Smilde Foods BV (2018) ECLI:EU:C:2018:899 (*Levola Hengelo*), para. 38.

(292) In art. 2(3) on adaptations, and in Art. 14*bis* on film.

(293) On the (implicit) requirement of originality in international copyright law, see for example Daniel J. Gervais, "The Protection of Databases," *Chicago-Kent Law Review* 82, no. 3 (2007), https://papers.ssrn.com/abstract=1116643.

(294) Goldstein/Hugenholtz, 4th ed., p. 176

(295) Dreier, in Dreier/Hugenholtz, Concise Copyright, Berne Convention, art. 2, note 7(b).

(296) Goldstein/Hugenholtz, 4th ed., p. 229; Adolf Dietz, "Le Concept d'auteur Selon Le Droit de La Convention de Berne," *RIDA* 155 (1993): 11–13; Ricketson, "The 1992 Horace S. Manges Lecture - People or Machines," 8, 11; Jane Ginsburg, "The Concept of Authorship in Comparative Copyright Law," *DePaul Law Review* 52, no. 4 (June 1, 2003): 1069.

(297) Art. 2(6) Berne Convention ("protection shall operate for the benefit of the author").

(298) Ginsburg, "People Not Machines," 135.

(299) Art. 27 UDHR. See Gervais, "The Machine As Author," 30.

(300) Art. 27 (2) UDHR.

in general terms. Art. 1 of the Term Directive[301] comes fairly close to a general definition by referring to copyright subject matter as "a literary or artistic work within the meaning of the Berne Convention".[302]

Moreover, in its jurisprudence, the CJEU regularly seeks guidance from the Berne Convention's concept of work.[303] The EU concept of a "work" is therefore intrinsically linked to that of art. 2(1) of the Berne Convention.

The EU *acquis* expressly harmonises three or possibly four specific categories of copyright-protected subject matter: computer programmes, databases, photographs, and – possibly – works of visual art.[304] Such works are protected if they are "original in the sense that they are the author's own intellectual creation". In its landmark *Infopaq* judgment of 2009, the CJEU extrapolated from this piecemeal harmonisation a general, autonomous concept of EU law of the work as "the author's own intellectual creation".[305] This has been confirmed in later decisions by the Court, most recently in *Levola*, *Funke Medien* and *Cofemel*.[306]

In the following, we will discuss the main prerequisites for protection of subject matter under EU copyright law, in the light of AI-assisted outputs. Although we will occasionally illustrate our findings by referring to (examples of) AI productions, we will leave for section 3.4 the core question of whether, and under which conditions, AI outputs qualify as "works" protected under harmonised copyright standards.

3.3.1. Production in literary, scientific or artistic domain

From the definition of "work" in the Berne Convention follows **a general requirement that works be produced within the "literary, scientific or artistic domain"**. Whereas some scholars have given normative meaning to this categoric notion,[307] the CJEU has not clearly embraced this "domain test" as a separate criterion. In *Premier League* the Court denied copyright to sporting events for reason that they "cannot be regarded as intellectual creations classifiable as works"[308], which possibly suggests an application of the domain test. However, elsewhere in the judgement it transpires that the court's exclusion of sporting events is based on the lack of originality.[309] Similarly, in *Levola Hengelo* the Court could have relied on this test to deny work status to the taste of a food product. Instead, the Court formulated a criterion of "identifiable expression" to achieve the same result.[310] It is therefore **not clear whether this general test is incorporated into EU copyright law**.

3.3.2. Human intellectual effort

Both the Berne Convention and the EU copyright *acquis* are primarily grounded in the tradition of author's rights (*droit d'auteur*): copyright protects original expression directly emanating from a human creator. A second requirement for copyright protection, therefore, is that a production be the result of human intellectual effort.

Although EU copyright law nowhere expressly states that copyright requires a human creator, its "anthropocentric" focus (on human authorship) is self-evident in many aspects of the law.[311] For one thing, the CJEU's case law on originality, discussed below, completely relies on the notion of a human being engaging in creative acts – reflecting

(301) Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (codified version) (Term Directive).

(302) See also Recital 16 Term Directive: "A photographic work within the meaning of the Berne Convention is to be considered original if it is the author's own intellectual creation reflecting his personality, no other criteria such as merit or purpose being taken into account."

(303) See, e.g., CJEU Levola Hengelo, para. 39.

(304) Art. 1(3) Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Codified version) (Computer Programs Directive); Art. 3(1) Database Directive; Art. 6 Term Directive; Art. 14 CDSM Directive (on works of visual art in the public domain).

(305) Case C-05/08 Infopaq International v Danske Dagblades Forening (2009) ECLI:EU:C:2009:465 (*Infopaq*),.

(306) CJEU *Levola Hengelo*; Case C-469/17 Funke Medien NRW GmbH v Bundesrepublik Deutschland (2019) ECLI:EU:C:2019:623 (*Funke Medien*); Case C-683/17 Cofemel – Sociedade de Vestuário SA v G-Star Raw CV (2019) ECLI:EU:C:2019:721 (*Cofemel*).

(307) See e.g. Dietz, "Le Concept d'auteur Selon Le Droit de La Convention de Berne."

(308) Joined Cases Football Association Premier League Ltd and Others v QC Leisure and Others (C-403/08) and Karen Murphy v Media Protection Services Ltd (C-429/08) (2011) ECLI:EU:C:2011:631 (*Premier League*), para. 98.

(309) Stef van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright," in *The Work of Authorship* (Amsterdam: Amsterdam University Press, 2014), 106.

(310) CJEU Levola Hengelo.

(311) See Ana Quintela Ribeiro Neves Ramalho, "Originality Redux: An Analysis of the Originality Requirement in AI-Generated Works," *AIDA*, 2019, 11.

"creative choice". As the Court considered in *Painer*, "[b]y making those various choices, the author of a portrait photograph can stamp the work created with his 'personal touch'."[312]

For example, in its recent *Cofemel* judgement the CJEU states that "if a subject matter is to be capable of being regarded as original, it is both necessary and sufficient that the subject matter reflects the personality of its author, as an expression of his free and creative choices".[313]

Also, according to the CJEU the rights of reproduction, communication to the public and distribution accorded to the author in the InfoSoc Directive necessarily attach to a human creator, not a legal entity such as a film producer or publisher.[314]

In her opinion in *Painer*, Advocate General Trstenjak and the Court concluded from the wording of art. 6 of the Term Directive, that "*only human creations are therefore protected*, which can also include those for which the person employs a technical aid, such as a camera".[315]

In sum, the requirement of human intellectual effort excludes from copyright protection outputs that are produced without any human intervention.[316] For example, the aesthetically pleasing flowers of a rose or wings of a butterfly cannot be qualified as works. Likewise, a production that is wholly generated by an AI system without any human intellectual effort is excluded from copyright protection. This requirement does not rule out, however, creations by human authors made with the aid of machines, provided that the human contribution to the output meets the legal standard of originality/creativity to be discussed below.[317]

3.3.3. Originality/creativity

The concept of "the author's own intellectual creation" implies, in the first place, that the subject matter must be "the author's own", i.e. not copied. In the second place, and more importantly, it must constitute an "intellectual creation" by the author. [318] This twofold requirement usually goes by the name of "originality". As the CJEU, rather circularly, held in *Levola* and *Cofemel*: "the subject matter concerned must be original in the sense that it is the author's own intellectual creation".[319]

In *Painer* and *Funke Medien*, the Court clarified that intellectual creation implies originality which in turn implies making personal,[320] creative choices.[321] This was more recently confirmed in *Cofemel*.[322]

Originality in copyright law is not to be confused with novelty in design or patent law. A production that is the result of creative choices may qualify as a protected "work" even if it is identical to an earlier work, provided the second author did not copy the first. Whereas novelty is sometimes used as a proxy to determine originality, this is not an approach followed by the CJEU in its extensive case law on originality.[323]

(312) CJEU *Painer*, para. 92.
(313) CJEU *Cofemel*, para. 30.
(314) CJEU *Luksan*; Case C-572/13 Hewlett-Packard Belgium SPRL v Reprobel SCRL (2015) ECLI:EU:C:2015:750 (*Reprobel*). *See later* discussion of authorship and ownership at .
(315) Opinion AG Trstenjak in *Painer*, para. 121 (our emphasis).
(316) The Copyright Office of the U.S.A. refuses to register copyright claims in respect of "works produced by a machine or a mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author" if it determines that a human being did not create the work. *See* USPTO, Compendium of U.S. Copyright Office Practices § 101 (3d ed. 2017). Online: https://www.copyright.gov/comp3/., arts. 306 and 313 (2).
(317) See Ginsburg, "The Concept of Authorship in Comparative Copyright Law," 1074.: the participation of a machine or device, such as a camera or a computer, in the creation of a work need not deprive its creator of authorship status, but the greater the machine's role in the work's production, the more the "author" must show how her role determined the work's form and content."
(318) See on the distinction between "objective" and "subjective" originality: Mireille van Eechoud, "Along the Road to Uniformity - Diverse Readings of the Court of Justice Judgments on Copyright Work," *JIPITEC* 3, no. 1 (May 15, 2012): 70, http://www.jipitec.eu/issues/jipitec-3-1-2012/3322. Tatiana-Eleni Synodinou, The Foundations of the Concept of Work in European Copyright Law, in: Synodinou (ed.), Codification of European Copyright Law, p. 94 ff.
(319) CJEU *Levola Hengelo*, para. 36. *See also* CJEU *Cofemel*, para. 29.
(320) See Term Directive (version 93/98), recital 17: "an intellectual creation is an author's own if it reflects the author's personality".
(321) Note that US courts and doctrine also focus on "creative choice". See Gervais, "The Machine As Author," 41–42.
(322) CJEU *Cofemel*, para. 30.
(323) It is therefore surprising, and in our view incorrect, that the aforementioned (at 2.2.1) Draft Report on "IPR for the development of AI Technologies" identifies a "general trend with regard to that condition [*of originality*] … towards an objective concept of relative novelty, making it possible to distinguish a protected work from works already created". See DRAFT REPORT on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)), Committee on Legal Affairs of the EU Parliament, (Rapporteur: Stéphane Séjourné), 24.4.2020, https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.html?redirect., p. 9

The *Painer* decision is particularly instructive, as it concerns subject matter created with the aid of a machine – notably, a portrait photograph. According to CJEU, a portrait photographer

> *"can make free and creative choices in several ways and at various points in its production. [...] By making those various choices, the author of a portrait photograph can stamp the work created with his 'personal touch'. Consequently, as regards a portrait photograph, the freedom available to the author to exercise his creative abilities will not necessarily be minor or even non-existent".[324]*

Similarly, with regard to databases, a category of subject matter governed by the Database Directive, the Court held that "[*the*] criterion of originality is satisfied when, through the selection or arrangement of the data which it contains, its author expresses his creative ability in an original manner by making free and creative choices."[325]

In *Infopaq*, a case regarding the protection of newspaper articles against unauthorised scanning, the focus of the Court's originality enquiry is on "the form, the manner in which the subject is presented and the linguistic expression."[326] The CJEU clarified that, for literary works, the author's "free and creative choices" pertain to the selection, sequence and combination of words.[327] While admitting that words in isolation do not amount to intellectual creation, the Court adds that "it is only through the choice, sequence and combination of those words that the author may express his creativity in an original manner and achieve a result which is an intellectual creation." [328]

Originality or creativity do not, however, imply a requirement of artistic merit or aesthetic quality.[329] EU copyright law protects works of high art as much as it protects more mundane intellectual productions, such as simple photographs, industrial design, databases or computer software. Conversely, as the CJEU has clarified in *Cofemel*, "the circumstance that a design may generate an aesthetic effect does not, in itself, make it possible to determine whether that design constitutes an intellectual creation reflecting the freedom of choice and personality of its author, thereby meeting the requirement of originality."[330] In other words, the fact that a production may be aesthetically pleasing is no reason to qualify it as subject matter protected under EU copyright law. **This is an important observation in relation to AI-assisted outputs, many of which are undeniably of aesthetic value**.

EU copyright law's focus on the act of creation in terms of making free and creative choices necessarily implies that economic investment cannot, as such, justify protection. In *Football Dataco*, the CJEU squarely rejected "significant skill and labour" on the part of the producer of football fixtures lists as a relevant factor in assessing originality.[331] Similarly, in *Funke Medien* the Court considered "the mere intellectual effort and skill of creating [*military status*] reports are not relevant in that regard."[332]

In the following, we will more closely scrutinise the notion of "creative choice" and query whether EU law sets any minimum threshold of creativity.

(324) CJEU *Painer*, paras 90-93.
(325) CJEU *Football Dataco*, paras. 38-39.
(326) CJEU *Infopaq*, para. 44.
(327) CJEU *Infopaq*, para. 45.
(328) See also CJEU Funke Medien, para. 23.
(329) van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright," 99. *See, e.g.,* Computer Programs Directive, recital 8: "In respect of the criteria to be applied in determining whether or not a computer program is an original work, no tests as to the qualitative or aesthetic merits of the program should be applied." *See also* Ramalho, "Originality Redux," 7.
(330) CJEU *Cofemel*, para. 54.
(331) CJEU *Football Dataco*, para. 42: "On the other hand, the fact that the setting up of the database required, irrespective of the creation of the data which it contains, significant labour and skill of its author, as mentioned in section (c) of that same question, cannot as such justify the protection of it by copyright under Directive 96/9, if that labour and that skill do not express any originality in the selection or arrangement of that data."
(332) CJEU *Funke Medien*, para. 23.

Creative choice and external constraints

In essence, EU law's requirement of originality is met "if the author was able to express his creative abilities in the production of the work by making free and creative choices."[333] This raises the question of what the relevant parameters are of such creative choices.

We can distinguish between **various types of external constraints**: rule-based, technical, functional, and informational. All these may play a role in the legal assessment of AI-assisted outputs.

A **first important constraint relates to the rules of the genre in the relevant area**: literature, film, visual arts, music, etc. For example, painting in a cubist style necessarily entails certain standardised expressive features, e.g. using "cubist" form elements.[334] Similarly, composing a song in the blues genre involves certain standardised harmonies and rhythms.[335] Copyright law recognises constraints imposed by rules of the genre by excluding "style" from protection.[336] The exclusion of style follows from the so-called *idea/expression dichotomy*, i.e. the principle that "[c]opyright protection shall extend to expressions and not to ideas, procedures, methods of operation or mathematical concepts as such."[337]

Similarly, the **rules of a game** impose external constraints. In *Premier League,* the CJEU considered that "sporting events cannot be regarded as intellectual creations classifiable as works within the meaning of the Copyright [*InfoSoc*] Directive. That applies in particular to football matches, which are subject to the rules of the game, leaving no room for creative freedom for the purposes of copyright."[338] In *Football Dataco*, the CJEU held that the criterion of originality "is not satisfied when the setting up of the database is dictated by technical considerations, rules or constraints which leave no room for creative freedom."[339]

In *BSA*, a case concerning the graphic user interface of a computer program, the Court held that where the expression of components of a subject matter "is dictated by their **technical function**, the criterion of originality is not met, since the different methods of implementing an idea are so limited that the idea and the expression become indissociable."[340] That situation does not permit "the author to express his creativity in an original manner and achieve a result which is an intellectual creation of that author".[341] In other words, if an unprotected idea can be expressed in only one way, or in a very limited number of ways, idea and expression merge, and no copyright protection will arise.

In its recent *Brompton Bicycle* judgement, the CJEU reiterates that the originality test will not be met if a potential author's creative space is so narrow that idea and expression become "indissociable".[342] Nevertheless, the Court does see room for a finding of originality in cases where an author has sufficient creative freedom to express their personality in design choices that ultimately have a technical function.[343]

(333) CJEU *Funke Medien*, para. 19; CJEU *Painer*, paras 87-88.

(334) *See, e.g.,* Tate, Look Closer, All About Cubism, *What is cubism and why was it so radical?* https://www.tate.org.uk/art/art-terms/c/cubism/all-about-cubism.

(335) *See* Wikipedia, entry 'Twelve-bar blues', https://en.wikipedia.org/wiki/Twelve-bar_blues.

(336) van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright," 112.

(337) Art. 9(2) TRIPs Agreement; art. 2 WCT; see also art. 1(2) Computer Programs Directive.

(338) CJEU *Premier League*, para. 98.

(339) CJEU *Football Dataco*, para. 39. *See also* CJEU *Cofemel*, para. 31. Remarkably, in *Paine*r the CJEU did not contemplate external constraints to the making of school portrait photographs, such as the standardised format of such photos, and the obligatory poses. See van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright," 124.

(340) Case C-393/09 Bezpečnostní softwarová asociace - Svaz softwarové ochrany v Ministerstvo kultury (2010) ECLI:EU:C:2010:816 (*BSA*), paras. 49-50. For a comment on this case, *see* Stephen Vousden, "Protecting GUIs in EU Law: Bezpečnostní Softwarová Asociace," *Journal of Intellectual Property Law & Practice* 6, no. 10 (October 1, 2011): 728–37, https://doi.org/10.1093/jiplp/jpr111.

(341) CJEU *BSA*, paras. 49-50.

(342) *See* Case C-833/18 SI and Brompton Bicycle Ltd v Chedech / Get2Get (2020) ECLI:EU:C:2020:461 (*Brompton Bicycle*), para. 27: "the criterion of originality cannot be met by the components of a subject matter which are differentiated only by their technical function. It follows in particular from Article 2 of the WIPO Copyright Treaty that copyright protection does not extend to ideas. [...] Where the expression of those components is dictated by their technical function, the different methods of implementing an idea are so limited that the idea and the expression become indissociable." In U.S. copyright law this rule is embodied in the "merger doctrine". *See, e.g.,* Pamela Samuelson, "Reconceptualizing Copyright's Merger Doctrine," *Journal of the Copyright Society of the U.S.A.* 63 (2016).

(343) CJEU *Brompton Bicycle*, para. 26. For an analysis of the case, see Estelle Derclaye, "The CJEU Decision in Brompton Bicycle – A Welcome Double Rejection of the Multiplicity of Shapes and Causality Theories in Copyright Law," *Kluwer Copyright Blog* (blog), June 25, 2020, http://copyrightblog.kluweriplaw.com/2020/06/25/the-cjeu-decision-in-brompton-bicycle-a-welcome-double-rejection-of-the-multiplicity-of-shapes-and-causality-theories-in-copyright-law/.

Another external constraint is the **informational purpose** of the subject matter at issue. In *Funke Medien*, a case concerning copyright protection of highly descriptive military status reports, the CJEU suggested that the reports at issue might "constitute purely informative documents, the content of which is essentially determined by the information which they contain, so that such information and the expression of those reports become indissociable and that those reports are thus entirely characterised by their technical function, precluding all originality."[344] The Court's reluctance to extend copyright protection to factual reports reflects the general principle that copyright does not extend to facts or information *per se*.

In sum, the existence of rule-based, technical, functional or informational external constraints does not rule out creative choice, as long as these external factors leave sufficient space for creative choice.

Assuming external constraints leave a potential creator sufficient space for personal creative choices, the question arises: what kind of choices may be regarded as creative justifying protection under copyright. Clearly, the choice of a single general idea underlying a work (e.g., the idea of creating a Rembrandt portrait with the aid of a self-learning algorithm) is not sufficient, since ideas as such are not protected – even if they are original to the author.

However, **a combination of unprotected ideas may indeed reflect "creative choice"**. This follows clearly from the *Painer* decision, where the Court (non-exhaustively) enumerated a number of parameters that – when considered in isolation – would qualify as unprotected ideas, but taken together could lead to a finding of originality. According to the CJEU, a portrait photographer

> *"can make free and creative choices in several ways and at various points in its production. In the* ***preparation phase****, the photographer can choose the background, the subject's pose and the lighting.* ***When taking*** *a portrait photograph, he can choose the framing, the angle of view and the atmosphere created. Finally,* ***when selecting*** *the snapshot, the photographer may choose from a variety of developing techniques the one he wishes to adopt or, where appropriate, use computer software."*[345]

As *Painer* illustrates, a *creative* combination of ideas expressed in an intellectual production, might be enough for the result to qualify as a "work" protected under EU copyright.[346] These **creative choices may occur at various levels and in various phases of the creative process: conception/preparation, execution, and finalisation.**[347] Even creativity occurring solely at the conceptual or preparatory stage of a work might suffice for a finding of originality. This is in line with the Computer Programmes Directive's prescient inclusion of "preparatory design material" in the definition of a "computer programme",[348] which clarifies that computer-generated computer code will qualify as a copyright-protected computer programme if it is based on preparatory design work reflecting creative choices by a human author.

By the same token, in conceptual art, many, if not most, creative choices will occur at the conceptual level. Conceptual artists operate "by providing ideas in written form, as sets of instructions, leaving their execution or instantiation to others".[349]

The CJEU does not seem to require that the author's creativity or personality ("personal stamp") be objectively discernible in the resulting expression (the output). The CJEU decisions recited above do not mention such

(344) CJEU *Funke Medien*, para. 24. See M. Leistner, Das Urteil des EuGH in Sachen »Funke Medien NRW/Deutschland« gute Nachrichten über ein urheberrechtliches Tagesereignis, Anmerkung zu EuGH, Urteil vom 29.7.2019 – C-469/17 – Funke Medien NRW /Deutschland (ZUM 2019, 751), p. 720 (pointing out that the CJEU's decision might affect copyright protection of AI-generated news reports).

(345) CJEU *Painer*, paras 90-91 (our emphasis).

(346) The Painer court's reasoning that a combination of (genre-related, functional or technical) ideas might suffice for a finding of originality, has been criticised in the legal literature. See van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright," 121.

(347) Ramalho, "Originality Redux," 7.

(348) Art. 1(1), second sentence Computer Programs Directive. See recital 7 of the same directive: "Whereas, for the purpose of this Directive, the term 'computer program` shall include programs in any form, including those which are incorporated into hardware; whereas this term also includes preparatory design work leading to the development of a computer program provided that the nature of the preparatory work is such that a computer program can result from it at a later stage".

(349) L. Bently & L. Byron, in: In M. van Eechoud (ed.), The Work of Authorship. Amsterdam: Amsterdam University Press, 2014., p. 244.

a requirement. What appears to be sufficient is that a prospective author exercises their "free and creative choices" and *thereby* expresses their personality.[350]

We may conclude from the jurisprudence of the CJEU that for an AI-assisted output to pass the test of originality/creativity it is sufficient that the output be the result of creative choices. These choices may occur at several stages of the creative process: conception, execution, and/or finalisation (or redaction). As we will discuss in some depth later, the Court's case law therefore does not rule out a finding of originality in cases where the (human-authored) design is executed by an AI system without any further human intervention.

Minimum level of originality/creativity?

As the Court's case law CJEU suggests, the availability of sufficient "creative space" for the creator is a strong indicator of originality. But is that enough – or does the law additionally require that the "creative space" be creatively used as the ideas are being expressed in the final production? On the face of it, the case law does indeed suggest the latter. The Court speaks of choices that must be "creative", and that "by making those various choices, the author of a portrait photograph can stamp the work created with their 'personal touch'".[351] The Court's language might suggest that exercising creative freedom in a non-creative way, e.g. by making only obvious choices, would not result in a protected work. On the other hand, as we have seen before, the requirement of originality or creativity does not entail a test of artistic merit or aesthetic quality, or that the work is novel (new).

At the national level, courts have dealt with this problem in different ways. For example, the Dutch Supreme Court expressly denies copyright protection to "trivial" or "banal" expression, even under conditions of broad creative freedom.[352] The copyright cases so far decided by the CJEU do not give much guidance on how to assess the "creativeness" of the act of creating, if at all, nor do they define a minimum standard of creativity.

Early CJEU decisions suggest that if the external constraints allow an author sufficient creative freedom, then the level of creativity actually required by the Court is fairly low.[353] In *Infopaq* the CJEU suggested that even a short, 11-word text fragment might qualify.[354] Judging from the reasoning in *Painer*, the originality of a photographic work is practically a given.[355] Even in a case concerning run-of-the-mill school portrait photographs, "the freedom available to the author to exercise his creative abilities will not necessarily be minor or even non-existent".[356] **This suggests that even a combination of fairly obvious choices in the design, execution and editing of an AI-assisted output could suffice**.

By contrast, in *Funke Medien,* the Court (in line with Advocate General Szpunar) expressed serious doubts over whether the military status reports at issue could qualify as "works", since the standard format of these reports and their purely informational purpose left (too) little room for creative choices.[357] Even though the Court's reasoning in *Painer* and *Funke Medien* point in different directions, the focus of the CJEU's originality analysis in both cases is on the availability of creative choice.

In sum, if the creative space left to the prospective author is minimal or non-existent, because of external (technical, functional or informational) constraints, the production will not qualify as the author's own intellectual creation. In the end, the CJEU leaves this determination to the national courts.[358]

(350) See CJEU Painer, para. 92: "By making those various choices, the author of a portrait photograph can stamp the work created with his 'personal touch'."
(351) CJEU *Painer*, para. 93.
(352) Zonen Endstra v. Nieuw Amsterdam (2008) ECLI:NL:HR:2008:BC2153, HR 30.05.2008, NJ 2008, 556. See van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright."
(353) van Gompel, 100.
(354) CJEU, *Infopaq I*.
(355) For a critique of this approach, *see* van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright," 121.: "Still, while the existence of creative space is a prerequisite for an author to make an 'original' creation, this does not of itself lead to the conclusion that the creation must therefore also be original. Instead, as will be argued below, courts should distinguish the presence of creative space from how it is used. An individual assessment of the use of space in expressive form should thus be the ultimate test for establishing originality."
(356) CJEU *Painer*, para. 93.
(357) CJEU *Funke Medien*, para. 23.
(358) In its follow-up decision in *Funke Medien*, however, the German Federal Supreme Court still leaves the question of the "work" status of the military reports unanswered. *See* BGH 30 April 2020 - I ZR 139/15 - Afghanistan Papiere II. For early commentary, *see*: Viktoria Kraetzig, "Does Copyright Law Have to Balance Fundamental Rights beyond the Written Exceptions? Unfortunately, the German Federal Supreme Court Has Recently Left This Question Open.," *Kluwer Copyright Blog* (blog), May 28, 2020, http://copyrightblog.kluweriplaw.com/2020/05/28/does-copyright-law-have-to-balance-fundamental-rights-beyond-the-written-exceptions-unfortunately-the-german-federal-supreme-court-has-recently-left-this-question-open/;

3.3.4. Expression

A fourth prerequisite for copyright protection is that the human creator's creativity be "expressed" in the final production. The use by the author of their creative freedom must be somehow perceptible in the author's expression. Ideas that are not given shape or form cannot qualify as "works". The CJEU has on several occasions confirmed that expression is a *sine qua non* for copyright protection. Both in *Infopaq* and in *BSA,* the Court states that the author must have "express[ed] his creativity in an original manner".[359] In *Painer,* the CJEU observes that, for a work to be original, the author must be able to "express his creative abilities in the production of the work by making free and creative choices."[360] Similarly, in *Funke Medien* the Court opines that "only something which is the expression of the author's own intellectual creation may be classified as a 'work' within the meaning of Directive 2001/29".[361]

This requirement of expression implies a **causal link** between an author's creative act (the exercising of their creative freedom) and the expression thereof in the form of the work produced. What, however, remains unclear from the Court's case law, is whether and to what extent the original features of the work should (all) be preconceived or premeditated by the author. Indeed, it is fair to assume that **the concept of a work as "the author's own intellectual creation" not merely requires human agency or intervention, but also some degree of authorial intent**.[362] This follows logically from the prerequisite that the work is "created" by a human "author" and subsequently "expressed" – terminology that excludes wholly haphazard acts of nature, such as the shape of a flower or of a solidified lava stream.

The question that remains is whether copyright law requires precise intent of every original feature of the work, or whether overall authorial intent suffices. Assuming that human authorship goes hand in hand with – and often partly relies on – fortuitous expression, such as slapdash paint drippings in a work of art or spontaneous facial expressions in portrait photography, a requirement that all expressive features of the work be preconceived would be too strict – and not supported by existing law and practice. Instead, what is probably enough is ***general authorial intent***. That is to say, **it is sufficient that the author has a general conception of the work before it is expressed, while leaving room for unintended expressive features**.[363]

What is clear from the CJEU's jurisprudence is that the author's creative choices must be sufficiently clearly expressed, in the interest of legal certainty. In *Levola Hengelo,* the CJEU was asked to rule on whether the taste of a food product (a cheese spread) could qualify as a "work". Defendants had argued that the taste of food depends on too many subjective factors to justify IP protection, such as the freshness of the product, the palate of the taster, etc. The Court agreed. According to the CJEU the requirement of expression implies that a production be expressed "in a manner which makes it identifiable with sufficient precision and objectivity, even though that expression is not necessarily in permanent form."[364] For this reason, copyright protection for the taste of a food product is precluded.[365]

In the end, the CJEU's focus on creative *choice* as the hallmark of intellectual creation suggests that it is the process of creating rather than the ensuing act of expression that is truly decisive for copyright protection provided there is an attributable connection between the creative process and the expression.[366] This is in line with EU law's rejection of artistic or qualitative merit as a relevant criterion.

Julia Reda, "German Federal Supreme Court Defends Press Freedom in Two High-Profile Copyright Cases, No Resolution of Sampling Dispute," *Kluwer Copyright Blog* (blog), May 1, 2020, http://copyrightblog.kluweriplaw.com/2020/05/01/german-federal-supreme-court-defends-press-freedom-in-two-high-profile-copyright-cases-no-resolution-of-sampling-dispute/..

(359) CJEU *Infopaq I*, para. 45; CJEU *BSA*, para. 50.

(360) CJEU *Painer*, para. 89.

(361) CJEU *Funke Medien*, para. 20.

(362) But s*ee* Zonen Endstra v. Nieuw Amsterdam (2008) ECLI:NL:HR:2008:BC2153, HR 30.05.2008, NJ 2008, 556 (creative intent not required, whether or not a work is the result of a creative act should be judged merely on the basis of the creative production as such).

(363) Burk, Dan L., Thirty-Six Views of Copyright Authorship, By Jackson Pollock (April 7, 2020). Houston Law Review, Vol. 58, 2020, UC Irvine School of Law Research Paper No. 2020-40, Available at SSRN: https://ssrn.com/abstract=3570225. Ginsburg and Budiardjo, "Authors and Machines," 363.

(364) CJEU *Levola Hengelo*, para. 40.

(365) For commentary on this case, see e.g. Caterina Sganga, "Say Nay to a Tastier Copyright: Why the CJEU Should Deny Copyright Protection for Taste (and Smells)," *Journal of Intellectual Property Law & Practice* 14, no. 3 (March 1, 2019): 187–96, https://doi.org/10.1093/jiplp/jpy161; Eugenie Coche, "Heks'nkaas or the 'Fifty Shades of Taste' Explained by the CJEU through EU Copyright Law," *EIPR*, 2019, 173–80.

(366) See Martin Senftleben; Laurens Buijtelaar, "Robot Creativity: An Incentive-Based Neighbouring Rights Approach," *European Intellectual Property Review*, 2020.

We conclude from the preceding analysis that current EU copyright law, as interpreted by the CJEU, leaves room for the protection of AI-assisted outputs in a wide range of creative fields. **As long as the output reflects creative choices by a human being at any stage of the production process, an AI-assisted output is likely to qualify for copyright protection**. This is true even if the AI system has played a significant or even predominant role in the entire creative process.

We will apply the findings of our analysis to the AI output in section 3.4 below.

3.3.5. Authorship and ownership

"Work" and "author" are two sides of the same coin. In copyright law, no work exists without an author. Conversely, absent a work, there will be no author; the question of authorship will only arise if it has been established that there is a work – an intellectual creation – to which authorship can be attributed. However, the EU copyright *acquis* is not very instructive on the notion of authorship, and therefore relies largely on the Berne Convention.[367] While the InfoSoc Directive requires Member States to provide rights of reproduction, communication to the public and distribution to "authors", it does not define this notion. Nevertheless, the CJEU has on various occasions suggested that **the notion of "author" is reserved for a human creator**, not a legal entity such as a film producer or publisher,[368] let alone an AI system or a robot.

Only few provisions in EU copyright law directly address the issue of authorship.[369] The Computer Programmes Directive enshrines the general principle that the author shall be the natural person, that is, the human being, who has created a work. However, it leaves the Member States considerable discretion to deviate from this principle in their national laws relative to computer software, by allowing States to designate a "legal person" (e.g. a company) as the "author" of a computer programme.[370] Various other directives contain rules on authorship of audio-visual (film) works, the most important of which is the Term Directive that designates the principal director as (co-)-author of the work.[371]

In addition, the *acquis* comprises a number of special rules on copyright ownership. The Computer Programmes Directive provides that employers are exclusively entitled to the economic rights in software that their employees create in the course of their duties, or when following specific instructions. The employer and employee may agree otherwise.[372] For joint works, the Computer Programmes Directive stipulates that the exclusive rights are owned jointly.[373] A similar provision is laid down for copyright-protected databases.[374]

Most authorship issues are dealt with by national law. If two or more authors collaborate on creating a work and their individual creative contributions cannot be separated, the ensuing production will be a joint work, with each contributor qualifying as (joint) co-author.[375] Additionally, most national laws will require that the co-authors work according to a common design, making the joint work a "concerted creative effort". [376]

If only one of the collaborators engages in creative choices, with the other collaborator reduced to the role of "amanuensis", only the creatively acting person will qualify as author, and no joint authorship will ensue. Some national laws in the EU provide for special rules of authorship allocation in the case of works created following the

(367) *See* Dietz, "Le Concept d'auteur Selon Le Droit de La Convention de Berne."

(368) CJEU *Luksan*; CJEU *Reprobel*.

(369) *See* P. Bernt Hugenholtz et al., "The Recasting of Copyright & Related Rights for the Knowledge Economy," Report to the European Commission, DG Internal Market (Amsterdam: IViR, 2006), 159–78, https://papers.ssrn.com/abstract=2018238.; A. Quaedvlieg, Authorship and ownership in Synodinou, Codification of European Copyright Law. *See also* Ramalho, "Originality Redux." p. 10.

(370) Art. 2(1) Computer Programs Directive.

(371) Art. 2(1) Term Directive ("The principal director of a cinematographic or audio-visual work shall be considered as its author or one of its authors. Member States shall be free to designate other co-authors.").

(372) Art. 2(3) Computer Programs Directive ("Where a computer program is created by an employee in the execution of his duties or following the instructions given by his employer, the employer exclusively shall be entitled to exercise all economic rights in the program so created, unless otherwise provided by contract").

(373) Art. 2(2) Computer Programs Directive ("In respect of a computer program created by a group of natural persons jointly, the exclusive rights shall be owned jointly.").

(374) Art. 4(3) Database Directive ("In respect of a database created by a group of natural persons jointly, the exclusive rights shall be owned jointly.").

(375) P. Goldstein and P.B. Hugenholtz, International Copyright Law: Principles, Law, and Practice (4th Ed. 2019, OUP), p. 233.

(376) Ibid. S*ee* French Cour de Cassation, 1st Civ. Chamber, Oct. 18, 1994, 164 R.I.D.A. 304, 308 (1995).

design and under the supervision of an author.[377] The Dutch Copyright Act goes a step further by providing that when a legal person "discloses a work to the public as its own without indicating any natural person as the author, it is then considered to be the author of that work unless it is proved that, in the circumstances, the disclosure to the public of the work was unlawful."[378] Nevertheless, national rules like these that reallocate authorship to a legal person or other non-creating entity are rare and controversial, since they deviate from the fundamental principle, enshrined in the Berne Convention, that the copyright work results from an act of human creation.[379] Indeed, the French Court of Cassation has expressly rejected the possibility of according authorship status to any other entity than a natural person.[380]

On the other hand, national rules that allocate copyright ownership to non-authors are quite common. For example, most Member States have special rules on copyright ownership of audio-visual works. In addition, some Member States maintain rules that directly vest the copyright in works created under employment to the employers, usually legal persons.[381]

Proving or enforcing authorship or copyright ownership of a work is sometimes difficult in practice. For this reason, many Member States provide for rules that establish a (rebuttable) presumption of authorship or copyright ownership, in that the person indicated on or with the published work as the author is deemed to be the author, unless proven otherwise. The Berne Convention and Enforcement Directive validate such legal presumptions, and allow the person whose name "appear[s] on the work in the usual manner" to instigate infringement procedures.[382] Evidently, rules like these might be abused in cases where AI-assisted outputs that do not meet the standards of copyright protection are published.[383]

> In sum, rules on authorship and copyright ownership are largely unharmonised in the EU. This allows Member States discretion to devise tailor-made rules on attribution of authorship and allocation of ownership in borderline cases.

As we will further discuss in section 3.5, this may lead to divergent outcomes across the EU in respect of AI-assisted outputs.

3.4. Are AI-assisted outputs "works"? A four-step test

In light of the preceding analysis, we will now assess whether AI-assisted outputs can qualify as "works" protected under current EU copyright law. In answering this pivotal question we will concentrate on outputs produced by or with the aid of a pre-existing AI system. As explained in the state of the art review (Section 2 of this Report), the use of "artificial intelligence as a service" (AIaaS) allows users access to powerful AI systems capabilities without the need to develop the technology themselves. Nevertheless, we will occasionally also refer to bespoke (tailor-made) AI systems.

In the following we will assume a "user" of an AI system as a company or individual having access to an AI system that he, she or it has not developed, who produces an artefact with the aid of the system – the AI-assisted output in our terminology. It is this user, and this artefact, that will be central to our copyright law analysis.

(377) For example, art. 6 Dutch Copyright Act. For French law, see Ginsburg, "The Concept of Authorship in Comparative Copyright Law," 1072.

(378) Art. 8(1) Dutch Copyright Act reads: "Where a public institution, an association, a foundation or a company discloses a work to the public as its own without indicating any natural person as the author, it is then considered to be the author of that work unless it is proved that, in the circumstances, the disclosure to the public of the work was unlawful."

(379) See Quaedvlieg, in: Sinodinou (ed.), Codification of European Copyright Law, Alphen aan den Rijn (2012), p.215.

(380) Cass. 1re civ., 15 janv. 2015, n° 13-23.566, D. 2015, p. 206 et p. 2215. See CLSPA, Bensamoun, and Farchy, "Mission du CSPLA sur les enjeux juridiques et économiques de l'intelligence artificielle dans les secteurs de la création culturelle," 33.

(381) For example: art. 7 Dutch Copyright Act; art. 51(1)-(3) Spanish Intellectual Property Act.

(382) Art. 15(1) Berne Convention; Art. 5 Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights (as corrected in OJ L 157, 30.4.2004) (Enforcement Directive).

(383) See infra discussion at 3.5.

As our inquiry into EU copyright law has revealed[384], four interrelated criteria – comprising a "**four-step test**" – need to be met **for an AI production to qualify as a "work"**:

- Step 1 – Production in literary, scientific or artistic domain;
- Step 2 – Human intellectual effort;
- Step 3 – Originality/creativity (creative choice);
- Step 4 – Expression.

Step 1: Production in literary, scientific or artistic domain

According to art. 2(1) of the Berne Convention, copyright protects productions "in the literary, scientific or artistic domain". Whether this translates into a substantive requirement under EU copyright law however remains unclear. The CJEU has yet to apply this test in its case law on the concept of "work".

Assuming the words "in the literary, scientific or artistic domain" would indeed constitute a hard-core preliminary test, **it will probably not stand in the way of copyright protection for many categories of AI-assisted outputs**. As noted previously in this Report, many AI productions resemble archetypal works, and belong to "the literary, scientific or artistic domain" without any difficulty. AI systems are capable of generating almost the entire spectrum of work types mentioned in art. 2(1) of the Berne Convention, including news articles, poems, musical compositions, paintings, maps, industrial designs, geographical maps, photographs, films, et cetera. For these kinds of outputs, passing this initial test will therefore be unproblematic, assuming the domain requirement is a material prerequisite under EU law at all.

Step 2: Human intellectual effort

The second condition for an AI-assisted output to qualify as "work" is that it is the result of human intellectual effort. *Prima facie*, this requirement seems to present an obstacle, since many AI outputs appear to be automatically generated. The criterion of human intervention does not however rule out AI productions as a matter of course. As the Court clarified in *Painer*, **it is entirely possible to create works of authorship with the aid of a machine or device**.

Moreover, leaving aside the futuristic scenario of a completely autonomous creative robot, AI-assisted outputs will always go hand in hand with some form of human intervention, be it the development of the AI software, the gathering and choice of training data, the drawing up of functional specifications, supervising the creative process, editing, curation, post-production, etc. Even if the connection between the human intervention and the AI-assisted output is increasingly remote, at this point in time, **it is hard to conceive of content that is generated through AI that involves no human agency whatsoever**. What is problematic today and for the immediate future is whether, and to what extent, a natural person's involvement with the AI-assisted output – however remote – is sufficient for it to qualify as an intellectual *creation*. This brings us to the third criterion.

Step 3: Originality/creativity (creative choice)

The third and most crucial criterion is originality or creativity. In the words of the CJEU, this test is met "if the author was able to express his creative abilities in the production of the work by making free and creative choices."[385] As we have seen, the emphasis here is on the existence (a priori) of sufficient creative space, rather than on the creativity of the production as such. As long as externally imposed rule-based, technical, functional or informational constraints do not rule out creative freedom, the originality of the production is basically a given since the minimum level of required originality/creativity is rather low.

(384) See the previous analysis at 3.3.

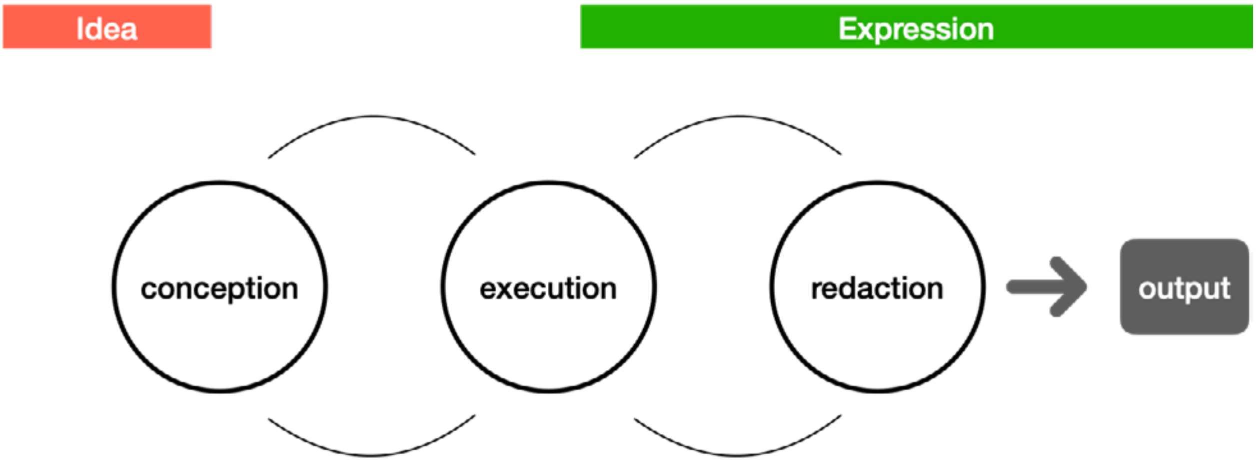
(385) CJEU *Funke Medien*, para. 19 ; CJEU *Painer*, paras. 87-8.

Creative choices may occur at various levels and in different phases of the creative process: preparation, execution, and finalisation.[386] As the *Painer* case illustrates, a creative combination of ideas at distinct stages in the production process might be enough for the result to qualify as a "work" protected under EU copyright. Even creativity occurring solely at the preparatory stage of a work could potentially suffice. This, however, must be assessed on a case-by-case basis.

The originality of an AI-assisted output will therefore depend on whether the production process involved creative choices by human authors that are reflected in the final result.

Inspired by the Court's decision in *Painer*, it seems useful at this stage of our analysis to have a closer look at the process of creating works with the aid of intelligent machines. As the *Painer* Court has well understood, **creativity in machine-aided production may occur at three distinct phases of the creative process: "conception", "execution" and "redaction"**.[387] Although these phases are to be distinguished for analytical purposes, in practice the **creative process** is usually **iterative**, involving multiple cycles of conception, execution and redaction. Figure 3.1 provides a simplified diagram of this iterative creative process.

Figure 3.1 Diagram of an iterative creative process



The conception phase involves creating and elaborating the design or plan of a work. This phase goes beyond merely formulating the general idea for a work.388 It requires a series of fairly detailed design choices on the part of the creator: choice of genre, style, technique, materials, medium, format, et cetera. It also involves conceptual choices relating to the substance of the work: subject matter (news article, portrait), plot (novel, film), melodic idea (musical work), functional specifications (software, databases), etc.[389] As the CJEU clearly stated in *Painer*, creative choices at this pre-production stage are important factors in a finding of originality of the final production.

In the case of **productions created with the aid of DL algorithms**, above and beyond the design choices identified above, additional conceptual choices may involve the choice of AI system (e.g. the type and characteristics of the models used), as well as the selection and "curation" of input data (e.g. in the labelling of training data) and other parameters.[390] With AI-assisted outputs most of these conceptual choices will be exercised by human actors. **The AI system at this stage has no role in the creative process,** other than acting as an external constraint limiting the designer's creative possibilities.

(386) Ramalho, "Originality Redux," 7.

(387) *See* Ginsburg and Budiardjo, "Authors and Machines." (discussing "detailed conception" and controlled execution). *See also* Ramalho, "Originality Redux," 7. (distinguishing "preparation", execution" and "final' phases in analysis of *Painer* judgment).

(388) Ginsburg and Budiardjo, "Authors and Machines," 347–48.

(389) van Gompel, "Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU's Originality Test for Copyright," 112ff. *See, e.g.,* Court of Brussel, 23 May 2017 (Diplomatic card v. Forax), I.R.D.O., 2017, 204 (functional and technical specifications protected as part of computer program). *But see* Dutch Supreme Court (Hoge Raad) 19 January 2018, ECLI:NL:HR:2018:56, NJ 2018/237 (Diplomatic Card v. Forax) (technical specifications do not qualify as "preparatory design material" if programming requires subsequent "creative steps").

(390) See Th. Dreier, 'Creation and Investment: Artistic and Legal Implications of Computer-generated Works', in: H.G. Leser & T. Isomura (eds.), Wege zum japanischen Recht. Festschrift für Zentaro Kitagawa, Berlin (1992), p. 869-888.

The ***execution* phase** involves, in simple terms, **converting the design or plan into what could be considered (rough) draft versions of the final work**. This phase involves the producing of text, the painting of art work, the notation or first recording of music, the "shooting" of photographs or video, the "coding" of software, etc. In the context of a supervised ML system, this phase would also include the optimisation of the model or models within the AI operational logic to match the pre-defined objectives or goals of the system.

With traditional forms of creation, the role of a human author at this execution stage is crucial. The novelist converts their plot for a novel into words, the composer translates their musical ideas into notes. From the 19th century onwards, machines have played an increasingly important auxiliary role in this creative phase. Photographs and films cannot be made without cameras, music not recorded without recording devices, etc. Even so, the human author has always stayed in control of the execution phase of creation. That is to say, the role performed by the machine was that of a tool or aid in converting executing the conception of the human author into towards an understandable or explainable outcome.

With AI technologies part of this relationship has arguably changed, at least in degree if not in nature. ML systems can be instructed and trained to perform complex tasks and produce sophisticated output in ways that the user of the system will not be able to (precisely) preconceive, understand or explain. From the user's perspective, this creates the impression of an autonomously operating system; one that they do not fully control and that strains the classification as a "tool". This is particularly true for DL systems, where the architecture based on several layers of neural networks places a greater distance between the user and the machine during the execution phase. This is even more so where the system relies on unstructured data and follows a paradigm of unsupervised or reinforced learning, which rely less (or sometimes not at all) on human intervention during the execution phase.

Whereas some AI systems are capable of generating highly sophisticated, "work-like" content at this stage of the creative process, the quality of the output should not be mistaken for proof of "creativity". What is relevant under a copyright analysis is *human* creativity – whether a machine is capable of autonomous creative behaviour is, *per se*, irrelevant.

While the AI system has taken over much of the human author's role in the execution stage, this does not imply that the user remains entirely passive. Especially in the case of supervised DL systems, the user's role is vital in constantly monitoring the output of the process and giving feedback to the AI system, by adjusting weights and parameters to better match pre-defined objectives, as noted previously in this Report.

> Nevertheless, at this execution stage, the AI system will normally play a dominant role in the creative process, while the role of the user will be mostly operational, by incrementally guiding the AI system towards the desired output. Even so, this may involve (additional) creative choices on the part of the (human) user.

Finally, the *redaction* phase involves processing and reworking the draft versions produced in the execution phase into a finalised cultural product or output ready to be delivered to a publisher or other intermediary, or directly to the market. This final phase *might* (and often does) involve a wide range of activities, depending on the genre and medium of the production. These include extensive rewriting, editing, correction, formatting, framing, cropping, colour correction, refinement and all sorts of (other) "post-production" activities that are necessary to give the final touch to the production before it is published and marketed.

Redaction is an often underestimated but essential, final stage in the creative process, allowing the human author many additional creative choices. As the *Painer* Court has explained, this final phase of the creative process may involve a variety of creative choices.[391] Indeed, depending on the circumstances, creative choice at the redaction phase may even suffice for a finding of originality of the entire production. For example, in a case involving geographical maps directly created on the basis of unprotected satellite photographs, the French Court of Cassation accepted that the maps qualified for copyright protection because they were "the result of a personalised implementation of a complex technology by a process of transformation and improvement of choices, in particular colours, contrasts and of luminosity".[392]

(391) CJEU *Painer*.
(392) Cass. Civ. I, 8 jan, 2002, R.I.D.A. 2002, no. 193, 321 (Père-Lachaise cemetery map).

With AI-assisted outputs, this has not fundamentally changed. Even a largely autonomously operating AI system will normally not deliver output that is immediately ready for publication or commercial use. More likely, the output produced by the AI system in the execution phase will often require redaction and selection by human actors. For example, a professional musician using an AI music composer such as AIVA or MuseNet would probably rework and edit output generated by the AI system before finalising the composition.[393]

Even so, not all AI-assisted productions will call for extensive redaction. For example, DL translation machines such as DeepL and Google Translate generate output that is almost ready to use. Nevertheless, here too some measure of human redaction will be required to convert the output into a useful and potentially marketable professional translation. Indeed, DeepL allows its users virtually endless creative freedom in selecting and rephrasing the wording and ordering of each (part of) the translated text.

In practice, producing content with the aid of AI systems will often be an iterative process where the execution and redaction phases are constantly repeated until the desired outcome is achieved. This will most likely be the case where users employ customisable DL systems that are trained to produce a specified type and quality of output. When the system has eventually "learned" to generate output in conformity with the user's specifications, ideally even the redaction stage can be (largely) performed by the AI system itself. This is the case, for example, with the football match reports generated by Fussball.De introduced in Section 2 and discussed below.[394]

In some cases, the redaction role of the human user will be reduced to that of selecting or refusing ready-made output generated by the AI system. This raises interesting questions from a copyright perspective. Clearly, the mere act of selecting may be one of many factors contributing to a finding of originality. But what if selecting one from multiple AI outputs is the *only* creative choice left to the user? Like many other questions raised by AI, this is not really a novel issue.

In the past, the emergence of non-traditional art forms such as the "ready-mades" created by conceptualist artists, have triggered similar questions. What is it that elevates a pre-existing artefact, such as a prefabricated urinal[395] or a bicycle wheel[396] to a work of art – and, by implication, to a work of authorship? According to Swiss copyright scholar Kummer, the decisive creative act here is converting the (in itself unprotectable) idea of a "ready-made" into copyright protected expression by *presenting* the artefact (the "objet trouvé") as a work of art.[397] Kummer's "presentation theory" implies that the mere act of selecting a pre-existing object suffices to convert the object into a work. While Kummer's theory has been embraced by some copyright scholars, it remains somewhat controversial.[398] In any case, **personal selection undoubtedly contributes to a finding of originality of an AI-assisted output**.

As the preceding discussion reveals, the use of highly advanced AI systems in the production of cultural goods does not imply that human beings have totally surrendered their vital role in the creative process to machines. Whereas the human creator has been partly or largely replaced by the machine in the execution stage of creative production, their role in the conception stage remains essential, while their role in the redaction stage may have become even more important than before – given that many AI-assisted outputs will probably require more redactional work than rough drafts produced by human beings. **This leaves both the design choices in the *conception* phase, some calibration in the *execution* phase, and the editing and post-production choices at the *redaction* phase for human authors**.

Moreover, it is important to realise that the three-phase creative process described above is simply a model to analyse and explain the authorial choices that contribute to a finding of originality. In reality, as noted, the creative process will be iterative; the execution phase will often yield unexpected results that inspire conceptual changes. Redaction as well may inspire new ideas that feedback to the conceptual level. In light of the *Painer* Court's reasoning regarding machine-aided creation, which designates both conceptual choices and post-production decisions

(393) See AIVA, https://www.aiva.ai/, and OpenAI, Musenet, https://openai.com/blog/musenet/.
(394) See infra at 3.7.
(395) Tate, Marcel Duchamp, Fountain, https://www.tate.org.uk/art/artworks/duchamp-fountain-t07573.
(396) Tate, Marcel Duchamp, Bicycle wheel, https://www.moma.org/learn/moma_learning/marcel-duchamp-bicycle-wheel-new-york-1951-third-version-after-lost-original-of-1913/.
(397) Max Kummer, *Das Urheberrechtlich Schützbare Werk* (Stämpfli, 1968), 193ff.
(398) *See* [AIPPI report Germany. Pro: Dreier, Fs Kitagawa. Contra Lauber-Rönsberg GRUR 2019, 244 at 247.]

as relevant factors in the originality analysis, these choices should in many cases be sufficient for a finding of originality in AI-produced content.[399] Applying the precedent set in *Painer*, "The Next Rembrandt" will most likely qualify as a work protected according to EU copyright.

Copyright doctrine and case law lend support to our conclusion that the production of an artefact executed by a largely autonomous AI system *could* qualify as a work protected under EU copyright law on condition that a human being initiated and conceived the work and subsequently redacted the AI-assisted output in a creative manner. That is to say, mere human intervention at the conception and redaction stages could suffice for copyright protection.

This conclusion is in line with copyright rules in many national laws that allocate authorship to the person that "masterminds" (conceives) and closely supervises the execution of a work by others, without that person materially contributing to the execution phase of creation.[400] In the words of Professor Ginsburg, "authorship places mind over muscle: the person who conceptualises and directs the development of the work is the author, rather than the person who simply follows orders to execute the work. Most national copyright laws agree that mere execution does not make one an author. An 'author' conceives of the work and supervises or otherwise exercises control over its execution."[401] For example, a marble sculpture designed in some detail by an artist who instructs and closely supervises a craftsman to execute the sculpture would qualify as the sole author of the sculpture, even if they never touched the marble stone.

While the CJEU has not pronounced itself on the issue of computer-generated productions, there is some case law at the national level that supports our general conclusion. For example, the Paris Court of First Instance has held that "computer-assisted musical composition, when it involves human intervention, the choice of the author [...] leads to the creation of original works".[402] In the same vein, the Bordeaux Court of Appeal opined "that a work of the mind created by a computer system can benefit from the rules protecting copyright, provided that it reveals even in a minimal way the originality that its creator wanted to bring."[403]

The preceding conclusion is consistent with the European Commission's Explanatory Memorandum that accompanied the initial proposal for the Computer Programmes Directive, which included a brief discussion of machine-generated software:

> *"The human input as regards the creation of machine generated programs may be relatively modest, and will be increasingly modest in the future. Nevertheless, a human 'author' in the widest sense is always present, and must have the right to claim 'authorship' of the program".*[404]

Step 4: Expression

The fourth part of our four-step test of copyright protection is that the human creator's creativity be "expressed" in the final production. As we have seen, this requirement rules out largely subjective subject matter that cannot be expressed with sufficient precision, such as the taste of a food product. In addition, we have derived from this criterion a prerequisite of *general authorial intent*: the human author must have a general conception of the work before it is expressed, while leaving room for unintended expressive features.

Prima facie, this requirement of general intent might present an obstacle for AI-assisted outputs. Due to the **"black box" characteristic of ML systems**, the human author in charge of designing the production at the conception

(399) *See also* Dreier, p. 882.

(400) Ginsburg & Budiardjo , p. 360. *See also* Ginsburg, "The Concept of Authorship in Comparative Copyright Law," 1072. For example, art. 6 Dutch Copyright Act provides: "Where a work has been made according to the design by and under the direction and supervision of another person, that person is considered to be the author of the work." See J. Seignette, in Hugenholtz, Quaedvlieg & Visser (eds.), A Century of Dutch Copyright Law, Amsterdam 2012, p. 123. Spoor Verkade Visser, p. 31 ff.

(401) Ginsburg, 1072.

(402) TGI Paris 5 July 2000, No. 97/24872 (Matt Cooper v. Ogilvy and Mather),

(403) Cour d'Appel Bordeaux, 31 January 2005, No. 03/05512.

(404) Explanatory Memorandum to Directive Draft L 1989 OJ. (C 91) 4, p. 21. *See* Ramalho, "Originality Redux," 11.

phase will not be able to precisely predict or explain the outcome of the execution phase. This, however, need not present an obstacle to the "work" status of the final output, **assuming that such output stays within the ambit of the author's general authorial intent**. Moreover, even completely unpredicted, non-explainable, quasi-random AI-assisted output might still be converted into a protected "work" at the redaction phase.

What "expression" does not require is that courts engage in an assessment of the production's creative merit, aesthetic value or cultural importance. As the case law of the CJEU suggests, it is sufficient for a production to be the expression of free creative choices.

As regards AI-assisted outputs, none of these expression-related conditions seem to pose insurmountable obstacles to copyright protection.

Borderline cases

The preceding analysis does not, however, imply that all AI-assisted outputs will unconditionally qualify as copyright-protected works under EU copyright standards. As the examples given and the case studies reveal, much will depend on the facts and circumstances of a case. While sophisticated art created with the aid of an AI system, such as *The Next Rembrandt* portrait, will necessarily involve important human creative input at several stages of the creative process, this may not be the case for more mundane AI-assisted output such as weather forecasts and news reports (see the cases discussed in section 3.7).

For example, in the case of a **weather report**, most design choices at the conception level (phase 1) will be predetermined; the specifications of the report are largely dictated by the utilitarian function of the weather forecast, and leave only limited space for creative choices by a human being. Possibly the designing of the format of the weather report (involving, for example, the template of a weather map) may involve some human creativity, but whether this will be enough for a finding of originality is questionable.

Assuming weather reporting entails insufficient creative choice at the conception level, we will have to look for originality elsewhere in the process. This may be difficult. At the execution phase, AI production of weather forecasts leaves hardly any room for human creativity. Whether the report as published will eventually qualify as a "work" will therefore largely, if not entirely, depend on the final stage of redaction. Again, this will hinge on the circumstances of the case. If the AI-produced weather reports are edited and formatted by human editors before being submitted to the media or other clients for publication and marketing, there may still be sufficient room for human creativity justifying copyright protection.[405] If, by contrast – as occurs in our case studies – the AI-produced reports are marketed "as is" with no or only limited human redaction, the reports will most likely not qualify as copyright-protected works.

While the latter scenario may be disappointing for the companies producing the news or weather reports with the aid of AI, it is not entirely surprising. More than a century before AI arrived on the creative scene, the drafters of the Berne Convention agreed that "[t]he protection of this convention shall not apply to news of the day or to miscellaneous facts having the character of mere items of press information."[406] In other words, simple news or weather reports "having the character of mere items of press information" will not be deemed literary artistic works under the Convention, even absent the use of any AI system. Evidently, the drafters of the Convention assumed that such productions do not reflect sufficient *human* creative activity.

In extreme cases, the AI system will not leave its users any meaningful choice beyond pushing a few buttons. Such cases are evident in the domain of natural language generation (relying on unsupervised learning), such as the GP-T2 and GP-T3 text generator from OpenAI discussed above.[407] One famous illustration is *Talk to Transformer* (now InferKit), which automatically completes a text based on a text fragment (prompt) supplied by the user.[408]

(405) See, for example, Cass. Civ. I, 8 Jan, 2002, R.I.D.A. 2002, no. 193, 321 (Père-Lachaise cemetery map); see section 5.4.1.3 above,

(406) Art. 2(8) Berne Convention.

(407) See: Open AI, GPT-2: 1.5B Release (5 November 2019), https://openai.com/blog/gpt-2-1-5b-release/; Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. "Language Models Are Few-Shot Learners." ArXiv:2005.14165 [Cs], July 22, 2020. http://arxiv.org/abs/2005.14165. For further references on GPT-3, see supra at 1.3.

(408) *See* Talk to Transformer, https://talktotransformer.com. Talk to Transformer has recently been turned into the paid product InferKit. According to its website, "InferKit offers an interface and API for custom AI–based text generators. Whether you're a novelist looking for inspiration, or an app developer, there's

Somewhat similar tools are Deep AI's Text Generation API[409] and StoryAI[410]. Recently, OpenAI has begun experimenting with applying the "transformer" model previously used on text to images, by training it with pixels. Early findings are that "sequences can generate coherent image completions and samples", meaning that automatic image generation through unsupervised learning may develop significantly in the foreseeable future.[411] In other words, the legal challenges raised here for text may soon apply also to image generation. In both cases, however, it will difficult to identify any creative choice by the human user either at the conception, the execution or the redaction phases. Consequently, the AI-assisted output generated by such systems would not qualify as a "work".

3.5. Authorship and Ownership of AI outputs

3.5.1. Authorship

As discussed, copyright law generally requires human authorship. The law of copyright allocates authorship to those person(s) that have, individually or jointly, creatively contributed to the production. In this section we query to whom authorship and copyright ownership in AI-assisted outputs is to be allocated. This analysis assumes that the AI-assisted output is a protected work in the first place. **In cases of AI-assisted outputs that do not qualify as works, no authorship can exist**. Such **"authorless" productions might** however still **enjoy protection under related rights**, as discussed below.[412]

Our analysis will focus on situations where *multiple* persons have a potential claim to authorship. These can involve for example the developer or programmer of the system, its owner, or its users. In light of the near absence of harmonised EU rules on authorship, and the lack of uniformity of national rules on authorship throughout the EU, the enquiry will necessarily remain very general.

For our authorship analysis of AI-assisted outputs, we follow the three-phase model of creativity previously developed: conception, execution and redaction.[413] As we have seen, in the case of artefacts produced with the aid of DL systems, at least the conception phase and (often) the redaction phase will normally entail creative choices by human persons to justify a finding of copyright protection of the AI-assisted output. Authorship in such cases is to be attributed to the person or persons individually or collectively engaging in these creative choices. **If more than a single author is involved in the process, and the authors collaborate, this will lead to co-authorship, even if the creative contributions occur at different stages of the creative process**. For example, the Dutch Supreme Court has held that a stylist who creatively arranged needlework pieces to be photographed for a magazine was a co-author with the photographer of the resulting photographs.[414]

We have also seen that the AI system usually plays a dominant role at the execution stage, despite some level of human intervention in systems relying on supervised learning. Following the design conceived by a (team of) human authors, the AI system (with the help of actuators) will paint, compose, photograph, 3D-print or otherwise produce a draft version of the production. Even if the AI system, depending on its capabilities, has added all sorts of expressive features to the human-conceived design, this will not justify an additional claim to authorship, since the AI system is not a human author and does not have legal personality.

something for you." *See* InferKit, https://inferkit.com/. For a basic explanation of how InferKit's text generation tool works, *see* InferKit, Docs, Text generation https://inferkit.com/docs/generation ("InferKit's text generation tool creates continuations of any text you give it, using a state-of-the-art neural network. It's configurable and can produce any length of text on practically any topic. You can also create custom generators for specific kinds of content.").

(409) DeepAI, Text Generation API, https://deepai.org/machine-learning-model/text-generator ("The text generation API is backed by a large-scale *unsupervised* language model that can generate paragraphs of text. This transformer-based language model, based on the GPT-2 model by OpenAI, intakes a sentence or partial sentence and predicts subsequent text from that input") (our emphasis).

(410) *See* Storyai, About – Q&A, https://storyai.botsociety.io/about ("Write the story you couldn't quite find the words to complete with this easy to use OpenAI model. Input 40 words to start, and watch what the model comes up with. It's powered by the GPT-2 774M model released on August 20th 2019 by OpenAI.")

(411) *See* OpenAI, Image GPT (17 June 2020), https://openai.com/blog/image-gpt/ ("By establishing a correlation between sample quality and image classification accuracy, we show that our best generative model also contains features competitive with top convolutional nets in the unsupervised setting").

(412) See infra at 3.6.

(413) *See* 3.4.

(414) Kluwer v. Lamoth, Supreme Court of the Netherlands, June 1, 1990, Nederlandse Jurisprudentie 1991, 377. *See also* TGI Paris, 6 July 1970, RIDA 190 (1970) (affaire Paris Match). *See* Ginsburg, "The Concept of Authorship in Comparative Copyright Law," 1070 (n.24).

Much of the literature on AI and copyright focuses on another scenario, that of the AI system producing content with only limited input on the part of the user of the system.[415] **If the role of the user of the system is so constricted that he cannot exercise free choices at any stage of the creative process, the user will not qualify as author of the ensuing production**. In such cases, the role of the user is essentially reduced to an initial prompt (e.g. writing an initial sentence) and/or "pushing buttons", as in the case of the AI text generation tools discussed above. Here the user's role is somewhat comparable to that of a person playing a computer game.[416] For example, authorship of film footage generated by a person playing the popular video game *Grand Theft Auto* most likely vests in the developers and animators of the video game – not in the player of the game. Even if the player feels empowered and "in control" of whatever transpires on the computer screen, they have no control over the creative process, and their choices do not amount to creative acts justifying a claim of authorship.[417]

As regards such AI systems, where users are effectively no more than passive "players", the user clearly does not have a valid claim to authorship in the AI-assisted output (i.e. in anything beyond its initial prompt) – leaving the **developer of the AI system** as the only candidate for authorship of the AI-assisted output.[418] Note, however, that a valid authorship claim may only arise if it is established that the output qualifies as a "work" in the first place. In the case of AI text generation tools such a finding, however, seems unlikely. The text generated by the AI system was not preconceived by the designer of the system, nor is it creatively redacted. At best one could argue that the output text is an adaptation (transformation) of the text the user input, of which the user (not the developer) is the author.

Valid authorship or co-authorship claims by developers of AI systems are likely to arise primarily in situations where developers and users collaborate on an AI production. Again, *The Next Rembrandt* project is a good example; the painting that the project eventually produced is the result of a team of AI developers, engineers, art historians, and others closely collaborating and jointly creating a work of authorship.[419] If the AI system developer played a creative role in the process, he or she clearly deserves co-authorship status.

In many if not most **cases**, however, the **developers of AI systems will not collaborate in a material way with the users on generating *specific* outputs**.[420] For example, providers of AIaaS (AI as a service) will provide customers (users) access to their ML systems (usually in B2B relationships) and customise it to the user's particular needs. As Section 2 notes, this is an emerging model in the field of automated journalism.[421] This may involve assistance in adjusting the system's goals to the needs of the users and relying on user data as input to the learning algorithm. The customised AI solution will then generate outputs without any further knowledge or direct intervention of the AI developers. In such cases, instances of (co-)authorship by AI systems developers are unlikely to materialise, since under prevailing (national) copyright law co-authorship can only arise if the work is the result of a "concerted creative effort", i.e. if multiple authors collaborate according to a common plan to create a specific work. Note that the developers of such general-purpose AI systems will usually be ignorant of the specific productions created with the aid of their systems.

Moreover, **co-authorship claims will also be unlikely for obvious commercial reasons**. An AI developer that would claim authorship (or even copyright co-ownership) in outputs generated with the aid of its system would probably not attract many customers. Assuming, as does this report, that AI systems will eventually become standard services or tools in the hands of business or commercial users and individual creators (similar to, e.g. Photoshop or Garage Band), the contractual terms of use of the AI system will probably resolve – and preclude – any such (co)authorship claims.[422]

(415) *See* e.g. Gervais, "The Machine As Author"

(416) *See* Cass. Ass. Plen., 7 March 1986: two cases: Atari and Williams Electronics, R.I.D.A. 1986, no. 129, 136.

(417) *See* [Ginsburg a.o.] with reference to U.S. cases on computer games.

(418) *See*, *e.g.*, Express Newspapers v Liverpool Daily Post [1985] 1 W.L.R. 1089, 1093 (computer programmer considered author of output generate with tailor-made program). *See also*: [AIPPI UK report, p. 5]; Ginsburg, "The Concept of Authorship in Comparative Copyright Law," 1074. Ginsburg, The Concept of Authorship, p. 1074.

(419) *See* Microsoft reporter, "The Next Rembrandt: Recreating the Work of a Master with AI," *Blurring the Lines between Art, Technology and Emotion: The Next Rembrandt* (blog), 04 2016, https://news.microsoft.com/europe/features/next-rembrandt/.

(420) See Samuelson (1986), p. 1223-1224.

(421) See supra at 2.5 (namely the future trends identified at 2.5.4.3).

(422) For a brief overview of the many legal and practical complexities that authorship/ownership claims by AI developers would entail, *see* CLSPA, Bensamoun, and Farchy, "Mission du CSPLA sur les enjeux juridiques et économiques de l'intelligence artificielle dans les secteurs de la création culturelle," 39.

For example, the popular DeepL AI-powered translation service does not claim any authorship or copyright in relation to content produced by its users with the aid of DeepL. Article 7.5 of DeepL Pro's terms and conditions provides:

> *DeepL does not assume any copyrights to the translations made by Customer using the Products. In the event that the translations made by Customer using the Products are deemed to be protected under copyright laws to the benefit of DeepL, DeepL grants to Customer, upon creation of such translations, all excusive, transferable, sublicensable, worldwide perpetual rights to use the translations without limitation and for any existing or future types of use, including without limitation the right to modify the translations and to create derivative works.*[423]

Another example can be found in AIVA's end-user license agreement, which distinguishes between different "types of licenses" that apply depending on the plan the user chooses. A "free plan" entitles users to a broad "non-commercial" license to use the output (musical composition) created with the assistance of the AIVA system, but ownership remains with AIVA and users must give a "Copyright Buyout" clause applies, according to which the "Licensor [AIVA]... assigns, grants and conveys all copyrights of the MIDI and/or Audio Composition to Licensee [the user]".[424]

Issues of co-authorship will also not arise in the scenario that an AI-assisted output incorporates (parts of) a pre-existing work. For example, a ML system trained to produce poetry will be "fed" vast amounts of poems, some of which are probably still in copyright. If the system inadvertently copies a part of a poem to the extent that it reproduces "the author's own intellectual creation", this will be deemed an infringement of the reproduction right.[425] As the CJEU has stated in *Infopaq I*, even reproducing short fragments of text might amount to infringement in so far as the reproduced fragment reflects the author's creative choices. By the same token, such a taking might be deemed an *adaptation* or derivative work (under national copyright law) in the case that the final AI-assisted output qualifies as an original work in its own right.[426] Although in such cases the author or copyright owner of the original work may have a valid claim of copyright infringement against the producer of the output, (s)he cannot claim (co-)authorship since the AI-assisted output is not the result of a concerted creative effort.

In the EU, allocating authorship to developers of AI systems may be further complicated by the divergent treatment of computer programmes, databases and other creative content.[427] Like computer games, AI systems that generate audio-visual content are a mix of computer software, databases and (in some cases) audio-visual works. Since the authorship of the component parts (software, databases, other works) will rarely coalesce in a single author, it may be problematic to establish (co-)authorship of the output generated by the system in those cases where the AI developer have a valid claim to (co-)authorship, that is, when the developer and the user of the AI system collaborate in producing creative output.

This point is illustrated by the *BSA* judgement of the CJEU. According to the Court, the copyright in a computer programme does not extend to the graphic user interface (GUI) generated by the computer programme on the screens of computer users, since the visual characteristics of the GUI are not "expression" of the computer code protected under the Computer Programmes Directive.[428] Nevertheless, the Court left open the possibility that the GUI is protected according to the general rules of the InfoSoc Directive.[429] Since the authorship and ownership rules of the Computer Programmes Directive are not repeated in the InfoSoc Directive, this may lead to inconsistent authorship and ownership allocation for AI-assisted outputs.

(423) DeepL Pro Terms and Conditions, available at https://www.deepl.com/pro-license/.

(424) See: AIVA, AIVA End-User License Agreement, https://www.aiva.ai/legal/1 (clauses 2 and 3); AIVA,Pricing, https://www.aiva.ai/; AIVA, Frquently Asked Questions – I don't understand the terms of License, https://aiva.crisp.help/en/article/i-dont-understand-the-terms-of-license-1wqvh5v/.

(425) Case C-05/08 Infopaq International v Danske Dagblades Forening (2009). Note that the copyright issues relating to the input process are outside the terms of reference of the present study.

(426) Note that the right of adaptation has not yet been (fully) harmonized at EU level.

(427) See BGH 06.10.2016 - I ZR 25/15, World of Warcraft I, GRUR 2017, 266 ('2nd world' game World of Warcraft contains distinct elements (software, graphics, sound) protected by different IP regimes). *See also* Case C-355/12 Nintendo Co. Ltd and Others v PC Box Srl and 9Net Srl (2014) ECLI:EU:C:2014:25 (***Nintendo***).

(428) In the United States, the 9th Circ. Court of Appeal has similarly held that the copyright in a computer-aided design program does not extend to the output produced with the aid of the program. Design Data Corp. v. Unigate Enterprise, No. 14-16701 (9th Cir. 2017).

(429) CJEU *BSA*. See also Court of Appeal (OLG) Karlsruhe 14 April 2010, case 6 U 46/09, GRUR-RR 2010, 234.

This is not to say that all such systems necessarily generate copyright-protected output in the first place. If an AI system is programmed or instructed to automatically produce content without the content being conceived (phase 1) or redacted (phase 3) by one or more persons exercising creative choices, there will be no work, and no authorship.

3.5.2. Copyright ownership

Copyright ownership follows authorship. It is a universal rule of copyright law that, by default, copyright vests in the person having created the work. This rule, however, is subject to numerous exceptions.[430] As we have seen, the EU copyright *acquis* regarding copyright ownership is limited to a small number of rules on computer programmes and databases created under employment, and on joint ownership of the same.[431] This has left room for divergent approaches to copyright ownership at the national level.

These and other rules on copyright ownership obviously apply *mutatis mutandis* to works created with the aid of AI systems. For example, copyright ownership in an audio-visual works created with the aid of AI will presumptively vest in the film's producer, i.e. the (legal) person initiating, organising and/or financing the film production. Given the flexibility of the notion of "audio-visual work", which according to some national case law includes computer games and the like,[432] this rule might become quite relevant for numerous AI productions that can be qualified as such.

Presumption of authorship and ownership

As discussed above[433], the Berne Convention and many national copyright laws provide for legal presumptions of authorship and/or copyright ownership in favour of the person whose name "appear[s] on the work in the usual manner". While these rules are intended to facilitate proof of authorship and ownership, they might in practice be abused to disguise the absence of copyright protection of an AI-assisted output by falsely attributing it to a natural or legal person. The issue has been flagged in the literature,[434] but it remains unclear whether it will amount to a serious problem in practice.

Note that falsely claiming copyright protection – also known as "copyfraud"[435] – is already a well-known, and growing, problem outside the domain of AI.[436] The problem is exacerbated by the rise of "copyright trolls" that extort content providers on platforms such as YouTube by threatening to trigger the notice and take-down procedures that these platforms (automatically) apply.[437] In the United States, the fraudulent use of copyright notice is criminally punishable under the U.S. Copyright Act.[438] In most EU Member States, no similar provisions exist. Nevertheless, falsely claiming authorship will probably be deemed unlawful and punishable under general criminal statutes. Nevertheless, **it might be useful to consider revision of the legal presumption of authorship enshrined in art. 5 of the Enforcement Directive in the light of possible abuses**.

British and Irish rules on copyright protection of computer-generated works

In some copyright laws of the British tradition – including the UK, Ireland, New Zealand, and South Africa – the requirement of human authorship has been circumvented by establishing authorship of "**computer-generated works**" in cases where no human authorship can be established.[439] Under these regimes, **authorship – and by implication copyright ownership – is accorded to the person who undertook the arrangements necessary for its creation**.

(430) See previously in this Report 3.3.5.
(431) Ibid
(432) Österreichischen Obersten Gerichtshofs vom 6. Juli 2004 – 4 Ob 133/04v , ZUM-RD 2005, 11.
(433) *See* previously in this Report at 3.3.5.
(434) CLSPA, Bensamoun, and Farchy, "Mission du CSPLA sur les enjeux juridiques et économiques de l'intelligence artificielle dans les secteurs de la création culturelle," 31.
(435) See Wikipedia entry on "Copyfraud".
(436) See Matthew Sag, 'Copyright Trolling, An Empirical Study', 100 Iowa L. Rev. 1105 (2015).
(437) See e.g. W. Worrall, 'YouTube Has a Massive False Copyright Claim Problem', CNN, 13 January 2020, available at https://www.ccn.com/youtube-has-massive-false-copyright-claim-problem/.
(438) Section 506(c) and 506(e) U.S. Copyright Act.
(439) See Guadamuz, "Do Androids Dream of Electric Copyright?" (including a survey of these national laws).

For example, the Copyright and Related Rights Act 2000 of Ireland defines “computer-generated”, in relation to a work, as meaning “that the work is generated by computer in circumstances where the author of the work is not an individual”.[440] The Irish Act proceeds to define as “author” “(f) in the case of a work which is computer-generated, the person by whom the arrangements necessary for the creation of the work are undertaken”.[441]

The UK provisions that inspired the Irish regime are similar, but not identical.[442] If the existence of a “work” is conditional upon human authorship, this statutory language seems to suggest that the Irish and British regimes allocate authorship to productions that would not qualify as “works” according to EU copyright law standards. Whether that is, indeed, the correct reading of these provisions, is however still unclear. Since the introduction of the regime on computer-generated works in UK law in 1988, it has led to just a single court decision, which has not clarified this issue.[443]

If the British regime indeed protects “authorless” computer-generated works, this would imply that an AI-assisted output that does not meet the standard of originality (and therefore is without an “author”) could nonetheless be accorded copyright protection under Irish and UK law, with the producer (“the person by whom the arrangements necessary for the creation of the work are undertaken”) as its author and copyright owner. Not surprisingly, the British and Irish regimes have been criticised as being incompatible with EU copyright standards.[444] Indeed, a national rule that accords copyright protection to subject matter that does not meet the standard of “the author’s own intellectual creation” is hard to reconcile with the CJEU’s case law that implies that the notion of a “work” is fully harmonised and therefore does not allow national laws to accord copyright protection under more lenient conditions.[445] According to some scholars, however, the British regime on computer-generated works is, “in substance”, not a copyright regime at all, but a species of related rights or even unfair competition protection.[446] Seen from that perspective the British and Irish rules would probably not conflict with EU law, since the EU *acquis* allows Member States to provide for specific related rights in addition to the rights expressly harmonised.[447]

The British approach towards computer-generated works is somewhat reminiscent of a provision that was once included in the draft proposal of the Computer Programmes Directive, but dropped in later versions. Under art. 2(5) of the draft proposal, the natural or legal person who caused the generation of subsequent programs would be entitled to exercise all rights in respect of the programmes.[448]

3.6. Protection of AI-assisted outputs by related rights

Having examined the protection AI-assisted outputs under EU copyright standards, the following section looks into the possibilities of protecting such outputs under the law of related (neighbouring) rights. Since the focus of this Report is on copyright law, this is necessarily a briefer analysis. Related rights are exclusive rights “relating” to (“neighbouring” upon) copyright that provide for copyright-like protection of persons or entities operating in the creative industries that do not qualify as creators of copyright-protected works. The major difference between related rights and copyright in this context is that related rights do not require originality or authorship. Related rights come in many shapes and forms, and differ from one Member State to the other. **EU law** has presently harmonised **six categories of related rights**.

The Rental and Lending Rental Rights Directive harmonises related rights of **performing artists**, **phonogram producers**, **broadcasting organisations** and **film producers**. In addition, the CDSM Directive obliges Member

(440) Art. 2(1) Copyright and Related Rights Act 2000 of Ireland.
(441) Art. 21 Copyright and Related Rights Act 2000 of Ireland.
(442) UK Copyright, Designs and Patents Act 1988 (UK, as updated), s178; Lionel Bently et al., *Intellectual Property Law*, Fifth Edition (OUP, 2018), 117–18.
(443) Nova Productions Ltd v. Mazooma games. *See also* Ramalho, “Originality Redux,” 13–14; Enrico Bonadio and Luke McDonagh, “Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity,” *Intellectual Property Quarterly* 2 (2020): 112–37.
(444) *See* Begoña Gonzalez Otero and Joao Pedro Quintais, “Before the Singularity: Copyright and the Challenges of Artificial Intelligence,” *Kluwer Copyright Blog* (blog), 10 2018. (reporting on the presentation of Professor Lionel Bently); and Bently et al., *Intellectual Property Law*, 118. (“Because the European standard now applies to all works, it must be doubted whether copyright protection (in an European sense) should be regarded as available at all to ‘computer-generated works’… It seems to follow that no computer-generated work can be protected by copyright in accordance with European Law”). On the latter point, *see also* Ginsburg, “People Not Machines.”
(445) See, in particular, Case C-604/10 (Football Dataco).
(446) *See* Bently et al., *Intellectual Property Law*, 118. On related rights protection for AI outputs, *see also later in this report*.
(447) See the discussion of related rights in section 3.6 below.
(448) Ramalho, “Originality Redux.”

States to introduce a related right for **press publishers** by 7 June 2021, the Directive's implementation deadline.[449] Apart from the performers' related right, these related rights have in common that they reward economic or entrepreneurial expenditure rather than human creativity. By the same token, most related rights may vest directly in legal persons to whom the entrepreneurial activities are to be attributed. At first sight, therefore, these related rights more easily accommodate AI-assisted outputs than copyright law with its focus on human creativity.

Here, we will briefly scrutinise each of these entrepreneurial related rights regimes (i.e. excluding perfomers' rights) for their potential to offer legal protection to AI-assisted outputs. In addition, we will look into the **related right in non-original photographs** that exists in Germany and a few other Member States, but which is presently **not harmonised**. Following this discussion of related rights regimes, we will examine the ***sui generis* "database right"** introduced by the Database Directive.[450] Finally, this section probes whether an expansion of current related rights to (better) protect "authorless" AI productions might be justified.

3.6.1. Related rights

Rights of phonogram producers

The Rental and Lending Rights Directive[451] and the InfoSoc Directives harmonise the related right of phonogram producers. Producers of "phonograms" (i.e. sound recordings) enjoy rights of reproduction, distribution and communication to the public.[452] The Directives do not define the notions of "phonogram" and "phonogram producer". From their legislative history, it transpires that these definitions can be derived from the 1960 Rome Convention and the 1996 WPPT, the two main international treaties on the protection of related rights.

The Rome Convention defines a "phonogram" as an "exclusively aural fixation of sounds of a performance or of other sounds", and a "producer of phonograms" as "person who, or the legal entity which, first fixes the sounds of a performance or other sounds".[453] The WPPT provides for similar, albeit slightly different definitions.[454] With digital technology in mind, the WPPT adds a definition of "fixation", which is absent from the Rome Convention. Fixation is defined as "the embodiment of sounds, or of the representations thereof, from which they can be perceived, reproduced or communicated through a device". In the case of computer-generated or computer-assisted music production, the notion clearly extends to sounds directly recorded on a computer hard drive or other digital equipment.[455]

The phonographic right is triggered by the act of "fixation" of sounds in a recording medium. No act of human authorship is needed, nor does the phonogram right require originality or any other threshold prerequisite.[456] As the diplomatic records of the Rome Convention reveal, even a simple recording (by means of a tape recorder) of "bird songs" suffices.[457]

(449) *See* arts 15 and 26 CDSM Directive.
(450) *See* arts. 7–11 Database Directive.
(451) Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental right and lending right and on certain rights related to copyright in the field of intellectual property (codified version) (Rental and Lending Rights Directive).
(452) *See* arts. 9(1)(b) Rental and Lending Rights Directive and 2(c) and 3(2)(b) InfoSoc Directive.
(453) Art. 3(a) and (c) Rome Convention.
(454) Art. 2 WPPT: (b) "phonogram" means the fixation of the sounds of a performance or of other sounds, or of a representation of sounds, other than in the form of a fixation incorporated in a cinematographic or other audio-visual work; (c) "fixation" means the embodiment of sounds, or of the representations thereof, from which they can be perceived, reproduced or communicated through a device; (d) "producer of a phonogram" means the person, or the legal entity, who or which takes the initiative and has the responsibility for the first fixation of the sounds of a performance or other sounds, or the representations of sounds.
(455) Jörg Reinbothe and Silke Von Lewinski, *The WIPO Treaties on Copyright: A Commentary on the WCT, the WPPT, and the BTAP* (Oxford, New York: Oxford University Press, 2015), para. 8.2.50. *See also* Mihály Ficsor, *The Law of Copyright and the Internet: The 1996 WIPO Treaties, Their Interpretation and Implementation* (Oxford, New York: Oxford University Press, 2002), para. 2.06. referring to the Records of the Diplomatic Conference ("sounds may be fixed in the form of data before they have been even aurally perceptible").
(456) The absence of a minimum threshold in neighbouring rights law has been criticised. *See* P. Bernt Hugenholtz, "Neighbouring Rights Are Obsolete," *IIC - International Review of Intellectual Property and Competition Law* 50, no. 8 (October 1, 2019): 1006–11, https://doi.org/10.1007/s40319-019-00864-3.
(457) WIPO Guide to the Rome Convention, WIPO 1981, p. 22.

As Advocate General Szpunar explains in his opinion in the *Pelham* case,

> *"[a] phonogram is not an intellectual creation consisting of a composition of elements such as words, sounds, colours etc. A phonogram is a fixation of sounds which is protected, not by virtue of the arrangement of those sounds, but rather on account of the fixation itself. [...]. Moreover, in the case of a phonogram, there is no requirement for originality, because a phonogram, unlike a work, is protected, not by virtue of its creativeness, but rather on account of the financial and organisational investment."*[458]

The phonographic rights vests in the "producer" of the phonogram. The international conventions define a "phonogram producer" in similar, but not identical ways. According to the Rome Convention this is "the person who, or the legal entity which, first fixes the sounds of a performance or other sounds". A phonogram producer may therefore be either a natural person or a legal entity. The WPPT's definition is more elaborate. A "phonogram producer" is "the person, or the legal entity, who or which takes the initiative and has the responsibility for the first fixation of the sounds of a performance or other sounds, or the representations of sounds". The WPPT's definition makes clear that the definition focuses on the entrepreneurial activity of the person or legal entity "taking the initiative and having the responsibility" for the recording, rather than on the physical person that actually makes the first recording.

In light of the phonogram right's absence of a requirement of human authorship or originality, and its rationale of rewarding economic or entrepreneurial activity, this right will fairly easily accommodate AI-assisted output. Whether any kind of human intervention is at all required is a matter of discussion in legal doctrine.[459] However, given that current AI systems will not spontaneously engage in producing "phonograms", this potential obstacle to protection need not be discussed here.

In practical terms, all that is required, is that the AI-assisted output at issue qualifies as a "phonogram", in other words: a recording of sounds. **This opens the door to a wide variety of AI produced audio output**, ranging from the generation through AI of electronic dance music to the generation through AI of aural translations. A remarkable example is OpenAI's MuseNet, based on the abovementioned GPT-2 transformer model, also used for automatic generation of text and images, and operating according to similar principles:

> *"We've created MuseNet, a deep neural network that can generate 4-minute musical compositions with 10 different instruments, and can combine styles from country to Mozart to the Beatles. MuseNet was not explicitly programmed with our understanding of music, but instead discovered patterns of harmony, rhythm, and style by learning to predict the next token in hundreds of thousands of MIDI files. MuseNet uses the same general-purpose unsupervised technology as GPT-2, a large-scale transformer model trained to predict the next token in a sequence, whether audio or text.*[460]

Other examples in the field of music are OpenAI's Jukebox and AIVA. OpenAI's Jukebox is a "neural net that generates music, including rudimentary singing, as raw audio in a variety of genres and artist style". [461] Jukebox uses a dataset, obtained from web crawling, "of 1.2 million songs (600,000 of which are in English), paired with the corresponding lyrics and metadata from LyricWiki".[462] AIVA markets itself as the AI "composing emotional soundtrack music" and a "creative assistant for creative people". It allows users to compose soundtracks with

(458) AG Szpunar, Opinion in Pelham GmbH v Hutter (C-476/17) EU:C:2018:1002; [2019], para 30.

(459) *See later* discussion with regard to the protection of non-original photographs.

(460) OpenAI, Musenet, https://openai.com/blog/musenet/. See also Jon Porter, "OpenAI's MuseNet Generates AI Music at the Push of a Button," *The Verge* (blog), April 26, 2019, https://www.theverge.com/2019/4/26/18517803/openai-musenet-artificial-intelligence-ai-music-generation-lady-gaga-harry-potter-mozart.

(461) OpenAI, Jukebox, https://openai.com/blog/jukebox/ ("a neural net that generates music, including rudimentary singing, as raw audio in a variety of genres and artist style"). See also Prafulla Dhariwal et al., "Jukebox: A Generative Model for Music," *ArXiv:2005.00341 [Cs, Eess, Stat]*, April 30, 2020, http://arxiv.org/abs/2005.00341.

(462) Id. ("The metadata includes artist, album genre, and year of the songs, along with common moods or playlist keywords associated with each song.")

"preset styles" and "influences".[463] For commercial uses, as noted, AIVA's "Pro Plan" allows users to "own the full copyright of any composition created with AIVA, forever".[464]

In respect of such audio recordings generated by AI, the phonographic right will be allocated to "the person, or the legal entity, who or which takes the initiative and has the responsibility for the first fixation of the sounds of a performance or other sounds, or the representations of sounds". This will in most cases be the user of the AI software, not the developer, since it is the user that triggers the act of fixation of the sounds by activitating the AI system.[465]

Rights of broadcasters

In line with the Rome Convention, the Rental and Lending Rights Directive and the InfoSoc Directive also accord related rights to "broadcasting organisations" in respect of their broadcasts.[466] Again, the Directives leave the definitions to the Rome Convention. The latter defines "broadcasting" as "the transmission by wireless means for public reception of sounds or of images and sounds",[467] but does not provide a definition of "broadcasting organisation." Apparently, this notion was considered self-explanatory.

Like the phonogram right, the related right of broadcasting organisations does not require any creative activity to trigger protection, nor does it provide for any other threshold requirement. All that is needed is an act of "broadcasting" of audio or audiovisual content by a "broadcasting organisation." Whereas the scope of the Rome Convention is limited to traditional "wireless" broadcasting, the Rental and Lending Rights Directive also protects broadcasts by cable or satellite.[468]

Likewise, the broadcasters' related right can easily give rise to protection of AI-assisted output if the preconditions of the right are met, in other words if a broadcast is automatically produced and transmitted by an AI system. Note that computer-assisted radio broadcasting has been a common phenomenon for some time,[469] Moroever, "artificial DJ's" are reportedly now making inroads into music radio.[470] Some examples of AI-assisted video and audio content susceptible of broadcasting can be found in Annex I.[471]

Rights of film producers

The Rental and Lending Rights Directive and the InfoSoc Directive provide for yet another category of related right: the right of "producers of the first fixations of films", generally known as the film producer's right.[472] In contrast to the preceding categories, the film producer's right has no history in the Rome Convention. Its origin can be traced to German law that traditionally protects non-original moving pictures (so-called *Laufbilder*) under related rights.[473] Note that a motion picture will normally (also) qualify as an original film work or audiovisual work protected by copyright, and that the copyright will usually be assigned to the film producer. The practical importance of having a separate related right is that it accords a minimum of legal protection to producers of *non-original* films.

Like the two preceding related rights, the **film producer's right** does not require originality or provide for any other threshold requirement. As a consequence, it **allows protection of all sorts of video content generated by AI systems**, varying from surveillance videos, to drone footage, to satellite imagery, to video content automatically generated for media channels.[474]

(463) AIVA, https://www.aiva.ai/.
(464) See AIVA, FAQs, "I don't understand the terms of License", https://aiva.crisp.help/en/article/i-dont-understand-the-terms-of-license-1wqvh5v/.
(465) In practice, phonographic rights may be transferred subject to contractual arrangement.
(466) Art. 13 Rome Convention. *See also* Art. 14(3) TRIPS.
(467) Art. 3(f) Rome Convention.
(468) See art. 7(2) Rome Convention; art. 8(3) Rental and Lending Rights Directive
(469) P.B. Hugenholtz, 'The WIPO Broadcasting Treaty. A Conceptual Conundrum', European Intellectual Property Review (EIPR), Vol. 41, #4, 2019, 199-202.
(470) See: R.J. Stine, 'Radio Streamlines Workflow With Artificial Intelligence', available at https://www.tvtechnology.com/news/radio-streamlines-workflow-with-artificial-intelligence.
(471) See Annex IA, under the heading "annex technologies".
(472) Art. 9(1)(c) Rental and Lending Rights Directive; art. 2(d) and 3(2)(c) InfoSoc Directive.
(473) Art. 95 German Copyright Act.
(474) See Annex IA, under the heading "generative technologies", for examples of AI-assisted video content.

Although the Directives are silent on the issue of ownership of the film producer's right, it may be assumed that the allocation rule of the phonographic right applies here mutatis mutandis. In other words, the right will belong to the person or legal entity, who or which takes the initiative and has the responsibility for the first fixation of the film. In case of AI generated audio-visual output this will in most cases be the user of the AI software – not the developer.

Rights of publishers of press publications

Art. 15 of the CDSM Directive obliges the Member States to grant certain exclusive rights to "publishers of press publications". Like the film producer's right, this new related right has its origin in Germany.[475] The right is meant to protect European newspaper publishers against online news aggregators that allegedly undermine the newspaper's business model. The CDSM Directive provides a detailed definition of "press publication". This is "a collection composed mainly of literary works of a journalistic nature, but which can also include other works or other subject matter, and which: (a) constitutes an individual item within a periodical or regularly updated publication under a single title, such as a newspaper or a general or special interest magazine; (b) has the purpose of providing the general public with information related to news or other topics; and (c) is published in any media under the initiative, editorial responsibility and control of a service provider."[476] Similar to the related rights previously discussed the press publisher's right does not require originality, and thus leaves room for content generated by AI that otherwise fits this definition. For example, **a blog generated by AI and publishing sports news would probably qualify for protection**, as long as it is published under the imprint of a European press publisher.[477] Whether from an information policy perspective it is desirable to protect and reward "robojournalism" by way of a legal instrument designed to promote a free and diverse press in Europe, is another matter.[478]

Other related rights in national law

Outside the related rights harmonised at EU level, national laws in Europe offer a "potpourri" of related rights.[479] For the purpose of this Report, the most relevant of these is the related right for non-original photographs that exists in Germany, Spain and a small number of other EU Member States.[480] The aim of this related right is to protect photographs that do not meet copyright's originality test, such as simple holiday snapshots, purely technical photographs or photographs produced with the aid of satellites or surveillance camera's.[481] The right therefore admits protection of photographs that would remain unprotected in other countries for lack of originality.[482]

Like the film producer's right discussed above, this related right *prima facie* seems to apply to all sorts of AI-assisted photographic content. However, the German *Lichtbildschutz* has given rise to discussion in legal doctrine and case law as to whether photographs that are produced by machines without any human intervention qualify for protection.[483] Case law also diverges on this point. But what is clear from the case law is that the photographic right does not encompass non-photographic computer-generated imagery. The right assumes photographic registration of a pre-existing physical object.[484]

3.6.2. *Sui generis* database right

The Database Directive offers special (*sui generis*) protection to databases that are the product of "substantial investment". The right applies to "databases", a notion defined as "a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other

(475) Art. 87f German Copyright Act.
(476) Art. 2(4) CDSM Directive.
(477) Note that the press publisher's right is not available for publishers that are not established in an EU Member State. *See* art. 15(1) CDSM Directive.
(478) See the critical observations in the report of the German AIPPI study group on AI & IP, p. 10. *See also later in this report*
(479) See Goldstein & Hugenholtz, 4th ed., p. 334 ff.
(480) Art. 72 German Copyright Act; art. 118 Spanish Copyright Act.
(481) *See e.g.* BGH 20 December 2018, I ZR 104/17 (photographs of public domain paintings protected by related right); see also OLG Dusseldorf, Feb. 13, 1996, 1997 GRUR 49 and OLF Hamburg, June 29, 1995, §997, ZUM-RD 217. OLG Hamm, Aug. 24, 2004, 2004, ZUM 927.
(482) TGI Paris, 3e Ch., 6 Oct. 2009, R.I.D.A. 2010, no. 226, 506 (aerial photographs not copyright protected for lack of originality).
(483) S*ee* the report of the [German AIPPI study group], p. 9-10. LG Hamburg 4 April 2003, case O 515/02, ZUM 2004, 675; LG Berlin 30 May 1989, case 16 O 33/89, GRUR 1990, 270 (Satellite photo).
(484) LG Berlin, Urteil vom 20.06.2017 - 16 O 59/16, ZUM 2017, 955. KG Berlin, Urt. v. 16.1.2020 – 2 U 12/16.Kart, GRUR 2020, 280.

means".[485] The database right is different from the related rights discussed above in that it does provide for a threshold criterion. Investment in the database must be "substantial", either in a "qualitative" and/or a "quantitative" sense.[486] Qualitative investment could, for instance, result from employing the expertise of a professional, e.g. an AI programmer or data scientist. In practice, most databases will result from quantitative investment, involving "the deployment of financial resources and/or the expanding of time, effort and energy".[487] **This might include expenditure in AI systems or technology used to produce a database**. For example, if a producer of a legal database would use AI technology to automatically create cross-references between documents and cases, the database producer's costs of procuring and implementing this technology would undoubtedly count towards (quantitative) investment.

The Directive defines the owner of the database right as the "maker of a database".[488] According to Recital 41, this "is the person who takes the initiative and the risk of investing", in other words: the database producer. However, **only nationals or residents of EU Member States, or companies and firms formed in accordance with the law of a Member State and having their registered office**, central administration or principal place of business within the EU, may benefit from the *sui generis* right.[489]

The substantial investment is to be made "in either the obtaining, verification or presentation of the contents" of the database.[490] The "obtaining" obviously refers to the collection (gathering) of pre-existing data, works or other materials comprising the database. "Verification" relates to the checking, correcting and updating of data already existing in the database. "Presentation" includes activities such as the digitalisation (scanning) of analogue files, the creation of a thesaurus or the design of a user interface.

In four cases concerning schedules of sports events ("fixtures"), the CJEU clarified that investment not in obtaining, verifying or presenting the contents of the database, but in generating its contents, does not count towards substantial investment.[491] According to the Court, "the expression 'investment in … the obtaining … of the contents' of a database must, […] be understood to refer to the resources used to seek out existing independent materials and collect them in the database, and not to the resources used for the creation as such of independent materials."[492] In other words, investment in creating the contents of a database may not be taken into account. Compilations of such "created" data will not qualify for database right unless some additional "substantial investment" can be demonstrated, for instance in presenting or verifying the database.

Following these judgements, scholars and courts have speculated on what distinguishes data "creation" from a mere "obtaining" of data.[493] Philosophically speaking, one could argue that any reporting of a fact amounts to data "creation". In the case of Football Dataco v. Stan James and Sportsradar – a case concerning a database of "live" statistics of football matches – the English Court of Appeals, however, squarely rejects this view. Facts observed, such as the scoring of a goal, are not "created"; subjective comments and interpretation, on the other hand, are.[494]

Since the EU database right does not require human authorship, it allows for protection of all sorts of AI-assisted outputs that qualify as "databases", including the database of pharmaceutical properties of molecules, weather reports and sports data generated by AI and discussed elsewhere as case studies in this report. The substantial investment requirement allows the costs of developing and implementing AI technology to be factored in, and therefore does not seem to present an unsurmountable obstacle to protection.

---

(485) Art. 1(2) Database Directive.
(486) Art. 7 Database Directive.
(487) Recital 40 Database Directive.
(488) Art. 7(1) Database Directive.
(489) Art. 11 Database Directive.
(490) Art. 7(1) Database Directive.
(491) Case C-46/02 Fixtures Marketing Ltd v Oy Veikkaus Ab (2004) ECLI:EU:C:2004:694 (*Fixtures Marketing Ltd v Oy Veikkaus Ab*); Case C-203/02 The British Horseracing Board Ltd and Others v William Hill Organisation Ltd. (2004) ECLI:EU:C:2004:695 (*British Horseracing Board and others*); Case C-338/02 Fixtures Marketing Ltd v Svenska AB (2004) ECLI:EU:C:2004:696 (*Fixtures Marketing Ltd v Svenska AB*); Case C-444/02 Fixtures Marketing Ltd v Organismos prognostikon agonon podosfairou AE (OPAP), (2004) ECLI:EU:C:2004:697 (*Fixtures Marketing Ltd v OPAP*).
(492) CJEU Fixtures Marketing Ltd v Oy Veikkaus Ab; CJEU British Horseracing Board and others; CJEU Fixtures Marketing Ltd v Svenska AB; and CJEU Fixtures Marketing Ltd v OPAP.
(493) *See e.g.* Mark J. Davidson and P. Bernt Hugenholtz, "Football Fixtures, Horseraces and Spinoffs: The ECJ Domesticates the Database Right," *EIPR*, no. 3 (2005): 113–18.
(494) Football Dataco Ltd & Ors v. Stan James plc & Ors and Sportradar GmbH and Anor [2013] EWCA Civ 27.

However, due to the vagueness of the creation/obtaining distinction discussed above, **it is currently uncertain whether investment in the machine-generation of data** – for example, the generation of weather data with the aid of AI – **may be factored in**. On the one hand, one could argue that such data are the direct result of observation of natural phenomena, and therefore "obtained". On the other hand, the use of (increasingly untransparent) AI systems in calculating these data, might suggest they are "created".[495]

Another potential obstacle to *sui generis* protection of data generated by AI is the prerequisite that the data be arranged in a "database". According to art. 1(2) of the Database Directive, data or other materials collected in a database must be "independent", that is to say, "materials which are separable from one another without their informative, literary, artistic, musical or other value being affected".[496] Therefore an audiovisual, cinematographic, literary or musical work or a sound recording does not qualify as a database, even if it can be perceived as a representation of data.[497] Moreover, the individual elements of the database must be "arranged in a systematic or methodical way". **This rules out protection of raw machine-generated data**. Consequently, for example, meteorogical data produced with the aid of AI will qualify for *sui generis* database protection only if the data are duly arranged or structured into a 'database'.

Whether and to what extent an AI system as such would qualify as a "database" protected under EU database right, is another matter altogether. This may or may not be the case depending on the circumstances, but since questions concerning the legal protection of AI systems as such are beyond the scope of this Report, we will not seek to answer this question here.

3.6.3. Special related right for AI-assisted outputs?

As is apparent, related rights regimes in the EU potentially extend to "authorless" AI productions in a variety of sectors: audio recording, broadcasting, film/video recording, news and databases. This has inspired some scholars to consider the introduction of a related or *sui generis* right that would apply to AI productions across the board.[498] Recognising, however, that AI systems (being machines) do not need economic incentives to produce output, Ramalho instead proposes a "dissemination right" somewhat akin to the (related) right granted to publishers of previously unpublished works, in order to promote and reward the making available of AI outputs to the public.[499] Similarly, Buijtelaer & Senftleben contemplate the introduction of a special neighbouring right that would serve as an incentive for "robot users" to train AI systems, and produce and disseminate AI outputs.[500] Both proposals, therefore, identify the *users* of AI systems as the potential beneficiaries of a new related or *sui generis* right.

An IP Strategy paper published in 2016 by the Japanese government suggests there might be a need for a special IP regime that protects "those AI-created works which are offered to the market and which generate a given amount of value (brand value)." The paper, however, does not offer any detail on the contours of such a regime.[501] According to a news report, the regime contemplated by the Japanese government would be akin to a rule of unfair competition. Protection would be granted to "[t]he person or company responsible for a technological system that produces creative work [...] Rights holders would be allowed to seek injunctions against or damages for unauthorised use, letting them more easily recover investment costs."[502] As far as can be ascertained, no concrete legislative proposals have however materialised in Japan.

(495) See Commission Staff Working Document, Evaluation of Directive 96/9/EC on the legal protection of databases, Brussels, 25.4.2018, SWD(2018) 146 final, p. 35. M. Leistner, 'Big Data and the EU Database Directive 96/9/EC: Current Law and Potential for Reform' in: Lohsse a.o. (eds), Trading Data in the Digital Economy: Legal Concepts and Tools (Nomos 2017), p. 25-58. See also P. Bernt Hugenholtz, 'Data Property in the System of Intellectual Property Law: Welcome Guest or Misfit?', in: Lohsse a.o. (eds.), Trading Data in the Digital Economy: Legal Concepts and Tools (Nomos 2017), p. 73 – 100.

(496) Fixtures Marketing Ltd v. Organismos prognostikon agonon podosfairou AE (OPAP)(note 31).

(497) Recital 17, Database Directive. See Hugenholtz, in Dreier/Hugenholtz, *Concise European Copyright* (2nd ed. 2016), Database Directive art. 1, note 3(b).

(498) "Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems by Ana Ramalho :: SSRN," accessed April 16, 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2987757.

(499) Ramalho, "Will Robots Rule the (Artistic) World?".

(500) Martin Senftleben; Laurens Buijtelaar, "Robot Creativity: An Incentive-Based Neighboring Rights Approach."

(501) 'Intellectual Property Strategic Program 2016', available at http://www.kantei.go. jp/jp/singi/titeki2/kettei/chizaikeikaku20160509_e.pdf.See Jani Ihalainen: Computer Creativity: Artificial Intelligence and Copyright, Journal of Intellectual Property Law and Practice, Issue 9/2018, p. 724-728, p. 727.

(502) N. Segawa, Nikei Asian review, 15 April 2016, 'Japan Eyes Rights Protection for AI Artwork', available at : http://asia.nikkei.com/Politics-Economy/Economy/Japan-eyes-rights-protection-for-AI-art

work.

Needless to say, introducing new rights or expanding existing related rights in order to protect AI-assisted outputs that do not qualify for protection under existing EU copyright or related rights regimes, would be justified only if solid empirical economic analysis were to reveal that the absence of legal exclusivity negatively affects overall economic welfare. This might be the case if the risks of copying and free riding would deter investment in AI productions to such an extent that the positive welfare gains of broad public access to public domain AI outputs are overshadowed. To be sure, establishing such "market failure" would require extensive economic study,[503] which is outside the terms of reference of the present report. In any case, no evidence of an industry need for IP protection beyond existing copyright and related rights regimes has been presented to the authors of this Report. [504]

Another important factor to consider before embarking on any legislative initiative in this field is the availability of other legal regimes that might offer producers or publishers of AI-assisted outputs legal protection. Such regimes might include, for example, trade secret protection, the law of unfair competition, and the law of contract – none of which are within the terms of reference of the present study.

3.7. Case Studies: Media and Science

3.7.1. Media: automated journalism

FUSSBALL.DE CASE

Fussball.de, operated by the German Football Association (DFB) using AI natural language generation (NLG) technology developed by Retresco, produces match reports of amateur football matches. The reports are generated based on the official match data provided by the DFB, which includes factual data on the team line-ups and the main match events, such as goals, penalties, warnings, substitutions, with corresponding times. The match reports are automatically generated applying standard text templates and storylines developed by Retresco. The generated match reports are directly posted to the Fussball.de site, without prior error checking or editing by human actors (see example in the state of the art review, Section 2 of this Report).

As the legal analysis concludes, copyright protection according to EU standards may arise if a "four-step" test has been met: (1) production in literary, scientific or artistic domain; (2) human intellectual effort; (3) originality/ creativity (creative choice); (4) expression. Creative choices may occur at three stages of the creative process: (1) conception, (2) execution and (3) redaction.

In this case, the AI-assisted output (match reports) are clearly within the "literary, scientific or artistic domain" (step 1). Also, they are the result of human intellectual effort (step 2), and "expressed" (step 4). As to step 3 (creative choices): the role of human beings is largely limited to the conceptual stage. The development stage involved making standard text templates and storylines. While these required substantial intellectual effort on the part of the human developers, these design choices were inspired by standard journalistic practices, and therefore probably lacked originality or creativity. For example, it will be hard to argue that the standard template sentence [example] required any measure of creativity.

In the operational stage, the execution of the texts does not require any further human intervention. While the AI system was still being developed, the system did receive extensive feedback from the human developers that trained the system to correctly interpret match data and generate proper journalistic language. However, these human interventions probably did not involve creating choices that are reflected in the generated output. Finally, the match reports are posted as is on Fussball.de, without any final human redaction.

(503) See Gonzalez Otero and Quintais, "Before the Singularity: Copyright and the Challenges of Artificial Intelligence." (reporting on the presentation of Professor Reto Hilty). See also Gervais, "The Machine As Author." (rejecting "arguments in favor of protection of machine productions by copyright for several reasons, not the least of which is that machines need no legal or financial incentives to run their code"). For general discussion of deontological and economic arguments that might justify IP protection of AI outputs, see Reto Hilty, Jörg Hoffmann, and Stefan Scheuerer, "Intellectual Property Justification for Artificial Intelligence," Oxford University Press, no. 5 (February 11, 2020): 1–29.

(504) See 2019 AIPPI World Congress – London Adopted Resolution on Copyright in artificially generated works, September 18, 2019: "As AI is still developing, it is too early to take a position on the question, whether AI generated works not covered by such existing protection should be eligible for exclusive rights protection as a Related Right or as exclusive rights under copyright (not in the meaning of the RBC)". See also submissions to WIPO Conversation on IP and AI, available at https://www.wipo.int/about-ip/en/artificial_intelligence/policy.html#submissions.

In summary, the match reports are probably not the result of creative choices by human beings that justify copyright protection under EU copyright law standards. This conclusion would probably be no different if the match reports were entirely produced by human beings. As the CJEU has stated in Funke Medien, reports that "constitute purely informative documents, the content of which is essentially determined by the information which they contain, so that such information and the expression of those reports become indissociable and that those reports are thus entirely characterised by their technical function, precluding all originality" cannot be copyright protected.

Moreover, the football match reports generated by Fussball.de are not eligible for related rights protection, except perhaps under the press publisher's right introduced by art. 15 CDSM Directive, given the broad definition of "press publication" that does not seem to rule out sports blogs such as Fussball.de. Additionally, the entire Fussball.de site – but not the individual reports – could benefit from database right protection, assuming the site fulfils the (broad) definition of "database". In light of the DFB's expenditure in AI technology, the substantial investment test will probably be met.

3.7.2. Science: weather forecasting

UBIMET CASE

UBIMET produces weather-related sales forecasts individually tailored to the specific needs of grocery retail stores. UBIMET also produces general weather forecasts for a variety of news media.

Both products are based on weather data that UBIMET derives from its own high-resolution weather model RACE, which are further refined by sophisticated data analysis. For its sales forecasts, UBIMET uses supervised Machine Learning technology developed in-house in order to produce weather-related sales forecasts geared to the needs of each individual client. The forecasting system is developed in close cooperation with the client. Archived sales data provided by the client are analysed and correlated with past weather data in order to develop and calibrate the forecasting system. During the development stage the model is trained and refined. Upon operationalisation, the forecasting operates largely automatically with little or no further human intervention.

For its media clients, UBIMET produces high-precision weather forecasts in the specific format desired by the media. It produces weather video content in its own studio, and for daily and weekly newspapers. The weather reports are prepared in UBIMET's own weather editorial system using individually designed graphical illustrations.

As the legal analysis concludes, copyright protection according to EU standards may arise if creative choices by human beings are reflected in the AI output. Creative choices may occur at three stages of the creative process: (1) conception, (2) execution and (3) redaction. Additionally, AI outputs may qualify for database right protection if the output qualifies as a "database" that is the result of "substantial investment".

As for the retail sales forecasts, the role of human beings is largely limited to the conceptual stage. The development stage mainly involved developing and calibrating a model of the forecasting system. While this clearly required substantial know-how and intellectual effort on the part of the human developers, these are not creative choices that give rise to copyright protection under EU standards. In the operational stage, the production of the forecasts does not require any further human intervention nor final human redaction. The forecasts are therefore not the result of creative choices by human beings that justify copyright protection under EU copyright law standards. As the CJEU has clarified in Football Dataco and Funke Medien, the mere intellectual skill and labour involved in reporting information cannot support a finding of originality that would justify copyright protection. However, the sales forecast reports generated by UBIMET might benefit from the database right, assuming the reports fulfils the (broad) definition of "database" and are the result of "substantial investment".

As for the weather reports produced for the news media, besides specifying the parameters and general format of the reports at the conceptual stage, human editors also play a role at the redaction stage, by designing illustrations, supplying texts, creating video content and doing final editing. This might be sufficient to justify copyright protection of the final output, depending on the facts of the case. The sample weather report reprinted in the state of the art review would probably qualify as a copyright protected "work".

Note that these conclusions would be no different if the forecasts were entirely produced by human beings.

4. Artificial intelligence and patent law

4.1. Outline, research question for patent law

This Section explains, first, the general patent law framework at both the multilateral and European levels (4.2). This includes in particular key international norms such as those contained in the Paris Convention (last updated at Stockholm in 1967); the TRIPS Agreement (administered by the WTO), other agreements administered by the WIPO, notably the Patent Law Treaty and the Patent Cooperation Treaty, and, for the European context, the EPC.

Second, it examines substantive areas of patent law that may be affected by AI and their specific application to AI-assisted outputs. This begins with a look at inventorship (4.3) and specifically the obligation contained in both European and international law to name the inventor, and the related question of inventors' "moral rights".[505] The matter of ownership of AI-assisted outputs that may involve patentable inventions is considered (4.4) and a variety of factual scenarios is provided. This includes a discussion of applicable law. The next part (4.5) examines the novelty assessment process. Novelty is recognised as a patentability criterion internationally. AI systems' ability to process and parse much higher volumes of information may have both quantitative and qualitative impacts on novelty determinations. Aspects of the inventiveness assessment process are discussed (4.6). As AI's use increases in certain industries as a tool to accelerate or amplify innovation processes, the application of the notion of obviousness may be challenged. Consideration is given (4.7) to whether AI-assisted innovation modifies aspects the sufficiency of disclosure. Finally (4.8), the Report applies the substantive findings to the two case studies (pharmaceutical research and weather).

4.2. The general patent law framework

4.2.1. The multilateral framework

There are four principal multilateral instruments that contain provisions relevant to the analysis of patent law. In chronological order, they are as follows:

The 1883 Paris Convention, last revised at Stockholm in 1967, provides a number of basic rules that are generally followed by its 177 Member States. This includes *national treatment*, that is, the obligation of States to treat foreign applicants and patent owners no less favorably than their own nationals. It also includes the *right of priority*, which means that a patent applicant may, within twelve months of a first application for a patent, file a subsequent application in one or more Paris Convention Member States. Using the priority date system, subsequent applications will be regarded as if they had been filed on the same day as the first application. Beyond those two basic rules, there are two substantive patent provisions that should be mentioned in the context of this Report. The first is that patents issued in each jurisdiction are independent of each other. This means that the "granting of a patent in one Contracting State does not oblige other Contracting States to grant a patent; [*and that*] a patent cannot be refused, annulled or terminated in any Contracting State on the ground that it has been refused or annulled or has terminated in any other Contracting State."[506] The second is that an inventor has the right to be named as such in the patent.[507]

The 1994 TRIPS Agreement, like the Paris Convention, provides for national treatment, but adds most-favoured nation treatment (MFN). MFN means "any advantage, favour, privilege or immunity granted by a Member to the nationals of any other country shall be accorded immediately and unconditionally to the nationals of all other Members."[508] Importantly, the TRIPS also provides that "patents shall be available for any inventions, whether

(505) NB The focus of this section is on AI aspects of inventorship, not the use of AI technologies in patent claim drafting. There is a difference between (a) AI that assists in developing a new invention, and (b) AI used to draft applications, and possibly broaden the scope if claims, just as there is between the work of an inventor and that of a patent agent optimizing the scope of claims in light of the prior art.

(506) WIPO, Summary of the Paris Convention. Online: https://www.wipo.int/treaties/en/ip/paris/summary_paris.html (accessed 15.2.2020)

(507) Art. 4*ter* Paris Convention.

(508) Art. 4 TRIPS.

products or processes, in all fields of technology, provided that they are new, involve an inventive step and are capable of industrial application."[509] A footnote to this article states that the terms "inventive step" and "capable of industrial application" may be deemed by a Member to be synonymous with the terms "non-obvious" and "useful" respectively.[510] The notions of utility and non-obviousness are used for instance in the patent law of Canada and the United States.[511]

The 2000 Patent Law Treaty (PLT), to which (as of June 2020) only 42 States had adhered—including 17 EU Member States, is also relevant in the context of this Report. As WIPO explains, its purpose is to "harmonise and streamline formal procedures in respect of national and regional patent applications and patents and, thus, to make such procedures more user friendly."[512] The PLT contains few provisions liable to impact directly the conclusions of this report. Rule 16(9) of the PLT Regulations may be relevant, however, as it provides inter alia that "[w]hat constitutes inventorship shall be determined under the applicable law."[513]

Finally, the 1970 Patent Cooperation Treaty (PCT) allows applicants to use a centralised application system administered by WIPO to file patent applications in up to 153 countries (as of June 2020). The PCT allows, but does not require, that States ask for data about the inventor.[514]

4.2.2. The European Patent System

The EPO was set up in 1973 by the EPC. As of June 2020, the EPO had 38 Member States, including all EU Member States.[515] Though its original text dates back to 5 October 1973, a substantial revision took place on 29 November 2000.[516] Also important in this discussion are the EPC Implementing Regulations (Rules), which were last amended on 27 March 2020. [517]

Patent applicants can directly file applications at the EPO, or combine this with national applications through the PCT system.[518] The patent laws of EU Member States vary considerably. It would be fair to say that there are two main groups of countries in the EU, however: those that provide a "registration" system for patents (with or without search), and those that provide a system of examination (which can be automatic or on request depending on the jurisdiction). [519]

Applications filed at the EPO are examined by the Office, which decides on their validity, and grants or denies the European Patent application. At the *post-grant stage*, however, "competence is transferred to the contracting states designated in the European patent."[520] In other words, national courts of the EPC Contracting States will ultimately determine the validity and scope of the patent in their jurisdictions. Particularly relevant rules in this context are contained in art. 81 of the EPC and Rule 19(1) of the EPC Implementing Regulations. The former provides in part that the application "shall designate the inventor." The EPC does not define the term "inventor" and has left its interpretation to the national laws of Contracting States, subject to the following.

(509) Art. 27.1. A TRIPS.
(510) Art. 27(1) TRIPS, fn. 5.
(511) For a discussion, see [Eli Lilly and Company v. Government of Canada, Expert Report of Daniel J. Gervais (Jan. 2015). Online: https://www.italaw.com/cases/documents/2910.]
(512) WIPO, Summary of the Patent Law Treaty (2000). Online: https://www.wipo.int/treaties/en/ip/plt/summary_plt.html.
(513) Rule 16 [(9)[Exclusion with Respect to Inventorship], *in fine* PLT (Regulations).
(514) Arts. 4(1)(v), 22(1) and 27(3) PCT.
(515) See European Patent Office, Member states of the European Patent Organisation (accessed 18 June 2020). Online: https://www.epo.org/about-us/foundation/member-states.html.
(516) Convention on the Grant of European Patents (European Patent Convention) of 5 October 1973 as revised by the Act revising Article 63 EPC of 17 December 1991 and the Act revising the EPC of 29 November 2000. Online: https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ma1.html.
(517) Implementing Regulations to the Convention on the Grant of European Patents of 5 October 1973 as adopted by decision of the Administrative Council of the European Patent Organisation of 7 December 2006 and as last amended by decision of the Administrative Council of the European Patent Organisation of 27 March 2020 Online: https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ma2.html. (**EPC Implementation Regulations**).
(518) Applicants may also apply directly at one or more national patent offices, in application of the national treatment principle.
(519) EPO, National Applications. Online: https://www.epo.org/applying/national.html.
(520) EPO, European patents and the grant procedure (2016), at p. 16. Online: https://bit.ly/2HpA3pI

The Rules provide in the relevant part that the "designation [*of the inventor*] shall state the family name, given names and full address of the inventor."[521] However, the EPO "shall not verify the accuracy of the designation of the inventor".[522] This rule is arguably reinforced by art. 62 of the EPC, which contains the inventor's right, "vis-à-vis the applicant for or proprietor of a European patent, to be mentioned as such before the European Patent Office." Article 60 of the EPC is also worth noting. It provides in part as follows: "The right to a European patent *shall belong to the inventor* or his successor in title. If the inventor is an employee, the right to a European patent shall be determined in accordance with the law of the State in which the employee is mainly employed" (our emphasis). An incorrect designation can be rectified.[523]

This **limited role of the EPO in determining both inventorship and ownership** can be explained by a look at the negotiating history. When the first version (1973) of the EPC was being drafted, two options were considered. First, including substantive ownership provisions, for example to deal with employee-made inventions; second, providing the EPO with a power to determine entitlement to a patent.[524] The negotiators decided against both, arguing that it would be "impossible to standardise the laws on ownership of inventions for all European States" and because it would be "equally impossible" for the EPO to determine applicable law.[525] The matter is thus left to national law, and national courts. This is reflected *inter alia* in the EPO Examination Guidelines, which provide specifically that it "may be adjudged by decision of a court or competent authority (hereinafter "court") that a person … other than the registered applicant, is entitled to the grant of a European patent".[526] In other words, "if a person is of the opinion that he is entitled to the grant of a European patent instead of the person whom the EPO has registered as applicant, he can obtain his rights only through a national court."[527]

The EPC makes a distinction between a *procedural* right to the patent, which, under art. 60 of the EPC, is **deemed to belong to the applicant, and the *substantive* right to the patent**.[528] The Enlarged Board of Appeal noted that only the inventor or his successor in title is entitled to apply for the grant of a European patent.[529] However, "the relation between the appli**cant and the person having the substantive right is governed by national law."[530] The EPO has no power to determine disputes regarding substantive entitlement**.[531]

Laws concerning the rights of employees vary considerably among Contracting States. The very question whether an employer is entitled to the European patent resulting from an employee's inventions received different answers.[532] The rules concerning the distinction between private and public sector employees and especially academic inventors are "particularly diverse".[533] For example, Germany has rules dating back to 1957 that allow an arbitration board to adjudicate disputes between employer and employee.[534] The law distinguishes inventions made by the employee in the course of the employee's duties ("service invention"), and those made outside of such duties ("free inventions"). Complex rules mandate compensation to the employee based on the value of the invention. Different rules apply in the UK, which contain as "notable features" detailed rules on ownership disputes between employers and employees, and a right of fair reimbursement or compensation for employees where it is determined that ownership of the invention belongs to the employer.[535] Still, like in Germany, such rules are mandatory and cannot be derogated from by contract.[536] As those rules typically protect "inventors", a determination of inventorship – under *national law* – may be required.

---

(521) Rule 19(1) EPC Implementation Regulations.
(522) Rule 19(2) EPC Implementation Regulations.
(523) See Rule 21 EPC Implementation Regulations. To remove a designated inventor from an application requires that person's consent.
(524) *See* Justine Pila and Paul Torremans, *European Intellectual Property Law*, Second Edition (Oxford, New York: Oxford University Press, 2019), 125.
(525) [Document BR/144/171, cited in *ibid.*] *See also* Bently et al., *Intellectual Property Law*, 620 (& fn.5). (noting that the reason for this choice is "because disputes about entitlement potentially raise questions about legal personality, contract, equity, and labour law, rather than patent law.").
(526) Guidelines for Examination in the European Patent Office (November 2019), para. 2.1. Online: http://documents.epo.org/projects/babylon/eponet.nsf/0/8654640290C2DBE7C12584A4004D2D9A/$File/epo_guidelines_for_examination_2019_hyperlinked_en.pdf.
(527) Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 138.
(528) Ibid. See also EPO Legal Board of Appeal, The Trustees of Dartmouth College, 4 February 2004, Case J2/01, r.2.6.
(529) Case G 0003/92 (13 June 1994). Online: https://www.epo.org/law-practice/case-law-appeals/recent/g920003ex1.html.
(530) Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 138.
(531) *Ibid.* See also Case G 0003/92 above.
(532) Pila and Torremans, *European Intellectual Property Law*, 125.
(533) Pila and Torremans, 125.
(534) *See* Philip W. Grubb et al., *Patents for Chemicals, Pharmaceuticals, and Biotechnology*, Sixth Edition (Oxford, New York: Oxford University Press, 2016), 412–13.
(535) Bently et al., *Intellectual Property Law*, 629–635.
(536) Grubb et al., *Patents for Chemicals, Pharmaceuticals, and Biotechnology*, 413; Bently et al., *Intellectual Property Law*, 631–32.

4.3. Inventorship of AI-assisted or generated outputs

4.3.1. Naming the inventor

Although, the **main international treaties seem to imply that an inventor is a natural person**, this is hardly surprising as all of them predate the emergence of AI technologies susceptible of assisting in or generating inventive outputs.

In the analysis of this issue, the report is based on the descriptive application of the relevant law to this new reality and guided by the results of consultations with experts and interviews.[537] As noted in Section 2, at this point in time, outputs that are fully generated by AI – i.e. autonomously created or produced by an AI system without human intervention or contribution – are essentially non-existent.[538] Even though terms such as "automated drug discovery" are fairly common in the legal literature, they may overstate the actual autonomy of the AI.[539] Hence, the focus of this section is mostly on what could be qualified as AI-*assisted* outputs, that is, with one or more humans involved in the invention process.[540]

Against this background, **the key question is *whether human inventorship is a substantive patentability requirement or rather merely a formal requirement***. The **EPO decided in 2020 that an AI system could not be the *named* inventor on a patent application**.[541] The case stemmed from an application filed with the UKIPO and forwarded to the EPO. The EPO asked the applicant to fill the field "inventor", which had been left blank. The applicant entered the name of an AI system known as DABUS. Two weeks later, the applicant filed another form indicating that he was the successor in title as the employer of DABUS. The EPO refused to accept that "entities other than natural persons" could be accepted as inventors under the EPC. It found that the legal requirement contained in the EPC was that "an inventor designated in the application has to be a human being, and not a machine."[542] At the time of writing, these decisions were under appeal.

The EPO decisions can be read in two ways. A first, expansive reading is to infer that AI-assisted inventions with little if any discernible human contribution are not patentable. This assumes that, to claim inventor status, a person must have made a contribution to the invention. As noted elsewhere, the link between the substantive elements required, if any, of contributions to an invention, on the one hand, and inventorship, on the other hand, are best viewed as a matter for the national laws of member States. The second reading is simply that, **because AI systems do not have the status of a person under the law (i.e. legal personality), they cannot be named inventors**.[543] The authors of this Report read the EPO decisions as suggesting the latter, not the former.[544] **Naming the inventor is**, in other words, **a formal requirement that a *human person* be named as inventor**, nothing more.

The UK IPO had earlier found along similar lines that, although the Office accepts indications of inventors at face value, for the purposes of sections 7 and 13 of the Patents Act, the inventor must be a natural person.[545] The

(537) *See* Annex IVA - Expert Workshop Report.

(538) ibid

(539) *See* Janet Strath & Reuben Jacob, 'Actavis v ICOS: obvious-to-try a dose of clarity', Eur. Int. Prop. Rev. 2019, 41(7), 461-464, 464.

(540) *See* previously in this report. NB a number of other experts and commentators beyond experts interviewed or consulted for the preparation of this Report agree with this position. *See* Annex - Expert Workshop Report.

(541) EPO decision of 27 January 2020 on EP 18 275 163. For a discussion, see e.g. Kaelyn Knutson, "Anything You Can Do, AI Can't Do Better: An Analysis of Conception as a Requirement For Patent Inventorship And A Rationale For Excluding AI Inventors," Cybaris An Intell. Prop. L. Rev 11, no. 2 (January 1, 2020), https://open.mitchellhamline.edu/cybaris/vol11/iss2/2. Others disagree. Professor Abbott, for example, suggested that "[M]achines have been autonomously generating patentable results for at least twenty years and … the pace of such invention is likely increasing"). See Abbott, "I Think, Therefore I Invent," 1083.. NB Professor Abbot is part of the "AI Inventor" team behind the patent applications designating DABUS as the inventor.

(542) EPO, 'EPO publishes grounds for its decision to refuse two patent applications naming a machine as inventor' (28 January 2020). Online: https://www.epo.org/news-issues/news/2020/20200128.html. In the same vein, the recently updated UK Patent Office 'Formality Guidelines' provide that "Where the stated inventor is an 'AI Inventor', the Formalities Examiner request a replacement. An 'AI Inventor' is not acceptable as this does not identify 'a person' which is required by law. The consequence of failing to supply this is that the application is taken to be withdrawn under s.13(2)." UKIPO, Formality Guidelines (October 2019), ch. 3.

(543) *See* arts. 62, 81 and 90 EPC, and Rule 19 EPC Implementing Regulations.

(544) This was also the reading of the experts consulted. *See* Annex – Expert Workshop Report.

(545) Case BL O/741/19 (4 December 2019), online: https://bit.ly/2zCPdYm.

High Court of England and Wales essentially agreed.[546] This decision is reflected in the April 2020 update of the UKIPO Manual of Patent Practice.[547]

It may be worth noting also that the United States Patent and Trademark Office (USPTO) decided - in April 2020 in a case involving the same DABUS system - that US law similarly requires that an inventor be a "natural person".[548]

4.3.2. Inventors' moral right

A possible counter-argument to the "mere formal requirement" point made in the previous paragraph is that the EPC provides a "right" to be named as an inventor, which is sometimes referred to as the "moral right" of inventors, thus drawing (at least terminologically) a parallel with authors' rights.[549]

Are inventor's (moral) rights indicative of a substantive or formal requirement of demonstrable human inventorship? As noted in the discussion of how AI technology can help generate or accelerate innovation in the pharmaceutical area, AI systems can use existing patent-related and other data to generate new leads.[550] They can also broaden the claim scope or add claims in patent applications. In both cases, AI systems seems to be predicting the next incremental steps in a given field of activity by analysing innovation trajectories. As Yanisky Ravid and Liu noted:

> *"AI systems create a wide range of innovative, new, and non-obvious products and services, such as medical devices, drug synthesisers, weapons, kitchen appliances, and machines, and will soon produce many others that, had they been generated by humans, might be patentable inventions under current patent law."*[551]

There is little doubt that AI will help innovate, and that, as part of this function, it will assist in producing inventions. AI systems are more likely to generate incremental innovations than true pioneer inventions or "game-changers". The example of Tu Youyou is telling in that regard. Her Nobel Prize (for a new malaria treatment) was the result of combining ideas from modern science and knowledge dating back to the fourth century BC in a way that is probably impossible for an AI system.[552] While AI systems can process huge amounts of data in a given field much faster and possibly better than humans, pioneer inventions that often result from a rapprochement between ideas from different fields are (still) the exclusive domain of humans. To quote Arthur Koestler, the "essence of discovery is that unlikely marriage of cabbages and kings — of previously unrelated frames of reference or universes of discourse whose union will solve the previously insoluble problem."[553] This is not to say that AI systems will not "innovate". The point is rather that the nature of AI innovation might belong to a more limited universe of "possibles", though the pace of AI innovation in that more limited space may, as explained above for pharmaceutical inventions, be much greater. For example, one company active in the field markets itself as creating "commercially relevant inventions at high speed and with great diversity" and notes "[h]undreds of patents based on *our inventions* have been filed by some of the best-known technology companies worldwide".[554]

(546) Thaler v The Comptroller-General of Patents, Designs And Trade Marks [2020] EWHC 2412 (Pat).

(547) UKIPO, Manual of Patent Practice (April 2020 update), Section 13. Online: https://www.gov.uk/guidance/manual-of-patent-practice-mopp/section-13-mention-of-inventor.

(548) In Re Application No 16/524,350 ("FlashPoint IP Ltd."), Decision on Petition, 22 April 2020. Online: https://bit.ly/2YMINyo.

(549) *See* arts 58 to 62 EPC. See also Annex – Expert Workshop Report. The experts cautioned, however, that this terminology can be misleading and that the moral right protection is significantly stronger in the authors' rights area.

(550) *See* the state of the art review, Section 2 of this report

(551) *See* Yanisky-Ravid and Liu, "When Artificial Intelligence Systems Produce Inventions," 2219–20. *See also* World Economic Forum, "Artificial Intelligence Collides with Patent Law," White Paper (Geneva: World Economic Forum - Center for the Fourth Industrial Revolution, April 2018), 6.

(552) See Jia-Chen Fu, "Project 523 Behind Tu's Nobel Prize Win for Malaria," The Conversation, October 6, 2015, https://www.usnews.com/news/articles/2015/10/06/project-523-behind-tus-nobel-prize-win-for-malaria. This development prompted new interest in pharmaceutical research based on traditional medicine. See [Michael Blakeney & Candy Cheung, 'Protecting traditional medicinal knowledge to promote the economic interests of the state: the case of the People's Republic of China', E.I.P.R. 2018, 40(12), 775-783.] It is possible of course that AI systems will be developed to help in this area.

(553) A. Koestler, *The Act of Creation* (New York: Arkana, 1964), p. 201.

(554) Iprova Invent First, http://www.iprova.com/about-us/.

If one were to look at the future, doctrinally one should thus distinguish three possible categories:

1. Inventions that are produced by human inventors, the development of which was supported by AI technology;
2. Inventions where the "inventive activity" was co-produced by humans and machine[555]; and
3. AI-generated "inventions", i.e. "inventions generated autonomously by an artificial intelligence (AI) under circumstances in which we believe that no natural person, as traditionally defined, qualifies as an inventor."[556]

The first and second category refer to AI-assisted inventions, which are those likely to be the only ones in existence for the predictable future. They do not pose insurmountable challenges. The (human) inventor or a successor in title (in the case of an employee for example) is the owner of the entitlement to the right to a patent under art. 60 of the EPC, independently of whether there is a partial "contribution" to the invention by the AI system. (The scope and extent of the contribution, if any, that need be made by a natural person to claim inventorship is a matter for national law.) A partial AI contribution might raise ownership issues,[557] which we address below.[558] Indeed, even if, as some experts are predicting, human-AI collaboration in innovation (a model in which humans work with AI systems) becomes the norm in a number of industries, a "partly human" invention should be patentable as any other invention would be, that is, subject to the usual patentability criteria.[559]

A full answer to the third category – fully AI-generated inventions – is beyond the scope of this Report as it is not an issue harmonised under the EPC and thus depends on a comparative analysis of national laws. A prevailing view seems to be that such inventions do not exist at this time, though they may emerge in the future.[560] Professor Shemtov's useful study on this question for the EPO concluded the following:

> *US, Japan, P.R. China, R. Korea, Germany, France, UK and Switzerland all appear to require creative or intelligent conception of the invention, or contribution thereto; it is a feature that either explicitly or implicitly runs throughout the definition of inventorship in all above jurisdictions... When assessing the nature of contribution to the conception stage it must be borne in mind that the real inventive 'spark', the decisive element that makes an invention work and differentiates it from that which has gone before does not necessarily have to originate from the inventive effort of the inventor.*[561]

Extending protection to purely AI-generated inventions (assuming such a category exists or comes into existence) may lead to a more comprehensive discussion, including an economic analysis of whether granting patents to such inventions would increase or hinder innovation. A standard textbook on the economics of IP notes that the appropriate degree of patent protection depends on several factors: "the patentee's fixed costs, the inherent

(555) But this should be distinguished from any claim drafting or "broadening" algorithm, which are excluded from the scope of our analysis. *See previously in this report*

(556) Ryan Abbott, "The Artificial Inventor Project." For an exploration of the notion of "AI-generated" invention and its relationship to "AI-aided" invention *see* Kim, "'AI-Generated Inventions.'"

(557) Some commentators note that allocation of patents on AI inventions is a key policy question to solve. *See, e.g.,* W. Michael Schuster, "Artificial Intelligence and Patent Ownership," *Washington and Lee Law Review* 75, no. 4 (February 19, 2019): 1945.

(558) *See* previously in this report

(559) See Liza Vertinsky, "Boundary-Spanning Collaboration and the Limits of Joint Inventorship Doctrine," Houston Law Review 55, no. 2 (November 20, 2017): 432. See also Kim, "'AI-Generated Inventions,'" 455. arguing that "as long as a computer is bound by an algorithm, there is no reason to assign to it 'cognitive' autonomy, irrespective of the complexity of the computational process". The author concludes that because "AI techniques represent computational methods of problem solving enabling partial automation of inventive activity... the application of such techniques cannot be prejudicial for the allocation of the inventor entitlement to a natural person... Instead, the use of such techniques should be a matter of the assessment of inventive step."

(560) For example, in a letter submitted to WIPO, IBM stated that "autonomously generated AI inventions are, in our view, far into the future and currently not capable of patent protection". IBM submission to WIPO, 30 June 2020. Some scholars have argued that those inventions do exist and indeed have for many years. See Ryan Abbott, "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law", (2016) 57 *Boston College Law Review* 1079, at 1083: "machines have been autonomously generating patentable results for at least twenty years". Others disagree. See e.g. Mark Summerfield, "The Impact of Machine Learning on Patent Law, Part 2: 'Machine-Assisted Inventing':, *Patentology*, 21 January 2018, 2:22 PM, online: https://blog.patentology.com.au/2018/01/the-impact-of-machine-learning-on_21.html. This Report need not take a definitive view on this question. It is sufficient to note that, in the case of the inventive processes analysed for the preparation of this report, and in particular the case studies, we have always found significant human contributions to the inventive process and have not otherwise seen any conclusive proof of the existence of entirely autonomously generated inventions.

(561) Shemtov, "A Study on Inventorship in Inventions Involving AI Activity."

difficulty of inventing around the patent […], and the extra profits that the patentee can expect to receive from greater protection."[562] Here, the high fixed costs are in designing and programming the system to "invent", not in developing the invention.

Then one could consider, as a number of authors have suggested, more **normative or deontological arguments**, some of which suggest that "pure" AI-generated inventions should not be patented.[563] Dr. Shemtov has suggested that, in the case of AI-generated inventions, there could be a break in the causation chain (what he refers to as the "chain of creation[564]) between human programmers/designers and the invention. The autonomy of the AI system, in other words, would lead to "non-human inventiveness". This does not lead to any particular conclusion about patentability (or absence thereof), however.

The PLT Regulations also provide that what constitutes inventorship shall be determined under the applicable law.[565] According to Dr Shemtov, current laws of a number of Member States requires human inventorship. European law only requires that one or more human inventors be named, but under EPC rules this does not amount a substantive requirement of human inventorship. It may very well be that the strongest policy case is for nonhuman inventions to be patentable. This is not likely a matter requiring a short-term policy decision, however.

For our immediate purposes, the legal answer to the question is likely to come not from the EPO but rather from national courts asked to decide in infringement proceedings whether the lack of (human) inventorship is a ground for invalidation. As technology stands today, on the specific issue of inventorship, no immediate action seems to be required at this stage at the level of the EPC.

This does not mean that change is not happening, however. If "entirely new paradigms" emerge, then the issue could be revisited.[566] In the following sections, observed and foreseeable changes in the application of key patent and other legal doctrines are discussed.

4.4. Ownership of AI-assisted or generated outputs

4.4.1. Ownership scenarios

The underpinning of the patent system is to provide an incentive to "increase investment in research and innovation, especially in areas that involve a high cost of research and a low cost to replicate the result, such as many pharmaceutical compounds."[567] One can argue that the design of the AI system used to assist in innovation processes requires a significant investment and that granting patents in the output of that system to the (human) programmer of the AI system is warranted.[568] The question of ownership of an AI-assisted (or even produced) invention might then arise if the programmer of the AI system is only one of parties involved in an innovation process that produces an output meeting the criteria for patentability.

(562) William M. Landes and Richard A. Posner, *The Economic Structure of Intellectual Property Law*, First Edition (Cambridge, Mass: Belknap Press: An Imprint of Harvard University Press, 2003), 300.

(563) *See* McLaughlin, "Computer-Generated Inventions," 23; Blok, "The Inventor's New Tool," 73; Vertinsky and Rice, "Thinking about Thinking Machines," 586.

(564) Shemtov, "A Study on Inventorship in Inventions Involving AI Activity," 28.

(565) Rules 16(2)(c) and 16(9) PLT Regulations.

(566) Erik Vermeulen and Wulf Kaal, "How to Regulate Disruptive Innovation: From Facts to Data," *Jurimetrics: The Journal of Law, Science, and Technology* 57, no. 2 (2017): 174–175, https://doi.org/10.2139/ssrn.2808044.

(567) Susy Frankel and Daniel J. Gervais, *Advanced Introduction to International Intellectual Property* (Cheltenham, UK ; Northampton, MA, USA: Edward Elgar Pub, 2016), 90.

(568) The law already has a system of protection for computer code, of course, based on copyright law, with some possible role by patent law for computer-implemented inventions that meet the patentability criteria, bearing in mind that under art. 52(1)(c) EPC, "programs for computers… shall not be regarded as inventions. Patentable subject matter current is not, however, part of this report. However, it should be noted that "prevailing patent law is the basis for all other legal assessments with regard to computer programs, no other granting requirements can be devised. In this respect, same granting requirements exist for computer programs as for other inventions, leaving only regular questions of novelty and inventive step to be discussed." *See* Oliver Baldus, "A Practical Guide on How to Patent Artificial Intelligence (AI) Inventions and Computer Programs within the German and European Patent System: Much Ado about Little," *EIPR* 41, no. 12 (2019): 751.

One may envisage various **ownership scenarios**, including at least three possible claimants to ownership of the invention:

- The programmer or developer of the AI system;
- The owner of the AI system; and
- The (authorised) user of the AI system who provided it with training or other data.

A **2020 UK court decision** left open the question of whether the owner or "controller" of an AI machine can claim ownership if and when such a machine "'invents' something".[569] In **reality**, it is **likely to be more complex** than this binary choice, as the AI "machine" (system) may well have various components, such as a standard AI platform on which a particular application was developed. Naturally, more than one of the roles can be performed by the same person or entity, though it is similarly true that each role could also be split among different persons or entities (for example, co-ownership of the AI system). As the **IKTOS case study** discussed in 4.8 illustrates, **humans can play several different roles in generating new AI-assisted inventions**. In the IKTOS case, those include: preparation/definition of the initial dataset (for ML); setting up of specific objectives for the ML process; preselection of compounds; interventions by engineers during the iterative trial and error process; and final selection of the compounds.

4.4.2. Applicable rules

Neither international law, nor the EPC provide clear rules concerning ownership of patents. In the case of the TRIPS, a WTO dispute-settlement panel specifically ruled, in a case dealing with trademarks but implicating the same two sets of multilateral rules that apply to patents (namely the TRIPS and the Paris Convention provisions incorporated into TRIPS) that "neither the TRIPS Agreement, nor the provisions of the Paris Convention (1967) that are incorporated into the TRIPS Agreement, determines who owns or does not own a trademark. Accordingly, the regulation of such ownership is, in principle, reserved to the legislative discretion of [WTO] Members pursuant to art. 6(1) of the Paris Convention (1967)."[570] As already noted, the **PLT Regulations** provide that what constitutes inventorship shall be determined under the applicable law. They also provide that a change of ownership of the patent may be recorded as ordered by a national court decision.[571]

Under the **PCT**, national law "determines whether a designated Office may request the applicant to furnish evidence of the identity of the inventor (R.4.17(i); R.51bis.1(a)(i); R.51bis.2(i) or a declaration or oath of inventorship (R.4.17(iv); R.51bis.1(a)(iv); R.51bis.2(iv))."[572]

As noted above, ownership is not a matter that can be solved simply by applying the EPC, and it was specifically excluded during the negotiation of the Convention. As the legal systems stand in the EU and many other jurisdictions, however, AI systems do not have legal personality and, therefore, cannot be "employees, or have a 'successor in title…in accordance with the law of the State in which the employee is mainly employed', to use the EPC's terminology.[573]

4.4.3. Ownership analysis roadmap

If one were to suggest a roadmap, **it may be advisable not to focus on exactly how much of the inventive activity, if any, is attributable to an AI system, for two reasons. First**, the AI system's "inventiveness" may well be derivative of the person(s) who programmed and/or supervised or directed the ML process. **Second**, as already noted, if a human person made an inventive contribution to an invention that meets patentability criteria, the fact that part of the inventive activity can be attributed in whole or in part to an AI system should not prevent

(569) Thaler v The Comptroller-General of Patents, Designs And Trade Marks [2020] EWHC 2412 (Pat), at para 52.

(570) United States - Section 211 Omnibus Appropriations Act of 1998 - Report of the Appellate Body, WT/DS176/AB/R (02/01/2002), at para. 222.

(571) Rules 16(2)(c) and 16(9) PLT Regulations.

(572) Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 844.

(573) A radical option, and one which would have major consequences on too many levels to mention here, would be to grant legal personality to AI systems. The fact that such systems often "collaborate" and work more "collectively" than humans might in itself pose definitional challenges, beyond the obvious normative questions. But such an option is currently not in the cards at EU policy level.

the invention from being patented. **There should be no issue in establishing the existence of a sufficient nexus between the invention and the applicant**.[574]

As noted above, under the EPC, the applicant has procedural rights and will generally be the owner of the right over a European patent on an AI-assisted output. This implies that problems between applicant and other persons who claim inventorship and ownership rights should be addressed under applicable national law**. At this time, there is an insufficient number of cases on entitlement to a patent (whether based on AI or not) to indicate that this is a significant enough problem to merit further harmonisation.** Importantly, this is not an AI specific problem though it might become more prominent of an issue in the context of AI-assisted outputs.

As technology now stands, it is reasonable to foresee that the programmer(s) of the AI system may in some cases have a claim to the AI system's output, but so might the person(s) who trained the machine (those in charge of dataset selection or creation, supervision of the learning process or reinforcement learning), as would the owner of the system who may have risked capital to produce the output, including the cost of developing the AI system, e.g. obtaining the input training data, developing the operational logic model(s) and the learning process, and turning outputs into applications.

This will not necessarily lead to a major flow of ownership conflicts before national courts. Hopefully, many of the bilateral relationships between those parties may be defined in contractual arrangements that may provide for right transfers in any valuable outcomes but also the inverse, namely liability limitations for negative, infringing or other harmful outcomes. A significant emergence of litigation in this area would likely constitute an additional incentive to consider contractual arrangements. In some cases, those arrangements will be under employment agreements, where applicable laws often include (mandatory) rules on ownership of inventions developed by employees.

As Professor Daryl Lim explained, "AI-generated works might infringe pre-existing rights of others in several ways. First, the AI could perform steps recited in the claims of a patented invention. A lens designed by or with the assistance of an AI system could infringe patents on optical lens and electrical circuit technology. Similarly, AI-enabled systems that map the fastest route between two points could ensnare both the AI developer and the agent who implements the system under a direct infringement theory where they cumulatively perform all the patented steps." [575] This is also a matter for national courts to sort out.

4.4.4. Applicable law

If, as noted above, the law of the relevant member State(s) applies, **possible conflicts may arise if those laws contain different ownership rules**. The **EPC Protocol on Recognition** provides a number of specific rules.[576]

Normally, proceedings against an applicant for a European patent will be brought in the courts of the applicant's residence or principal place of business within one of the Contracting States. If the applicant's residence or principal place of business is outside the Contracting States and the party claiming the right to the grant of the European patent has residence or principal place of business within one of the Contracting States, then courts of the latter party's State have exclusive jurisdiction. Parties can, however, make an agreement to grant exclusive jurisdiction to the court of a particular Contracting State. In the case of employee-made inventions, special rules apply.[577]

The Protocol can lead to a situation where a court of a Contracting State can have jurisdiction to hear a dispute about the law of another. In one case, a dispute between two English scientists who had collaborated on an invention (one had filed an application without involving the other) was found to be a matter for the exclusive jurisdiction of Swedish courts because the application had been assigned to a Swedish firm.[578]

(574) *See* McLaughlin, "Computer-Generated Inventions." p.243.
(575) Daryl Lim, "AI & IP: Innovation & Creativity in an Age of Accelerated Change," *Akron Law Review* 52 (2018): 865.
(576) Protocol on Jurisdiction and the Recognition of Decisions in respect of the Right to the Grant of a European Patent (Protocol on Recognition) of 5 October 1973. Online: https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ma4.html.
(577) Ibid., arts 2-5.
(578) See Grubb et al., Patents for Chemicals, Pharmaceuticals, and Biotechnology, 417.

In other cases, other private international law rules, including the Rome II (applicable law) and Brussels (jurisdiction) Regulations.[579] Normally, this will be the law of the country where a patent has been applied for and/or granted. As noted above, the Paris Convention provides that patents issued (and possibly then invalidated) in one jurisdiction are independent of patents issued in other jurisdictions.[580]

If a patent is issued by a patent office for an invention without an identifiable human inventor (as opposed to a named human proxy inventor), then of course inventorship could be challenged before national courts. As noted, at the current state of technology the scenario of an entirely AI-generated invention seems unlikely. Similarly, a refusal by a patent office to grant such a patent could lead to a court challenge. **This reinforces the general conclusion that this is a matter for national law**.

Overall, therefore, this Report can conclude, first, that under the EPC, the applicant has procedural rights and will generally be the owner of the right over a European patent on an AI-assisted output. This implies that problems between applicant and other persons who claim inventorship and ownership rights should normally be addressed under applicable national law. On that second point, the Report offers a roadmap but cannot draw a single legal conclusion applicable to all member States as to the ownership of every AI–assisted invention, as the possible impact of the AI assistance of the AI on inventorship, and how that, in turn, might influence ownership, is a matter for national law**.**

4.5. Novelty Assessment of AI-assisted outputs

This section explores issues regarding the application of the novelty requirement under the EPC to AI-assisted (or generated) outputs as inventions.[581] This part of the analysis can be brief since **there appear to be few if any significant novelty-related challenges associated with such outputs for our purposes**. Still, there are few issues to explore.

The basic rules for novelty under the EPC is that an invention is new only if it does not form part of the "state of the art" available to the public "by means of a written or oral description, by use, or in any other way, before the date of filing of the European patent application."[582] To defeat novelty, a single item of the state of the art must contain the elements of a claim in the application and enable the person skilled in the art (POSITA) to "practice the technical teaching which is the subject of the document, taking into account also the general knowledge at that time in the field to be expected of him".[583] In contrast, for inventiveness or inventive step (discussed below), elements of prior art may be combined.[584]

It should also be made clear at the outset that determining novelty can always be a difficult process, whether AI is involved or not.[585] "The requirement of novelty … is becoming increasingly difficult to meet and determine with any great level of certainty."[586] **The role of AI in this field means that several parallel changes are happening.** More data can be parsed by AI systems used by patent applicants or by patent offices. AI can assist inventors in myriad ways, including as explained in the case studies, in selecting the most relevant data for humans to work with. AI systems can also be used by patent offices to analyse more potentially relevant prior art faster. Applicants can perhaps use this same feature to modify or even "broaden" claims in an application. As noted above, this Report does not consider any of those changes anissue in need of remedial or other regulatory response. Thus far**, those are essentially *quantitative* changes to the amount of data that an AI system might be able to process during a patent application or examination process as compared to a human**

(579) See: Regulation (EC) No 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations (Rome II Regulation); and Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast) (Brussels Regulation).

(580) Frankel and Gervais, *Advanced Introduction to International Intellectual Property*, 117.

(581) *See* Arts. 54 and 55 EPC.

(582) Art. 54(1) and (2), EPC.

(583) EPO, Guidelines for Examination: Enabling disclosure of a prior-art document, online: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_vi_4.htm.

(584) See Ibid: State of the art pursuant to Art. 54(2), online: https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_vi_1.htm.

(585) Joff Wild, "Artificial Intelligence and the Future of the Patent System | IAM," *IAM* (blog), July 11, 2018, https://www.iam-media.com/law-policy/artificial-intelligence-and-future-patent-system.

(586) Ibid.

examiner. Admittedly, an experienced human examiner might, however, perform a much more targeted search and thus need to process much less data to arrive at their conclusion.

One possibly ***qualitative* change** that this quantitative change (and the related analytical approach) might induce is worth noting, however. It is that the **novelty assessment will be increasingly performed by AI systems**. As Professor Ebrahim explains:

> *Unlike past technological advancements in tools for the invention process, artificial-intelligence technology ushers in a form of omniscience in the patent-prosecution process and disintermediates the patent-prosecution process. The move toward automation and predictive analytics in patent prosecution will undoubtedly decrease reliance on patent legal judgment. In economic terms, artificial-intelligence technology reduces the transaction costs of acquiring patents.... Therefore, the economic impact of artificial-intelligence technologies will reshape patent law from a policy perspective. The danger in artificial-intelligence technology, particularly predictive analytics that can make predictions from large data sets, is the complex and opaque effects on interactions.*[587]

One policy consideration in that respect might be to maintain a level of technical capability at patent offices that matches the technology available to (most) patent applicants, a process apparently well underway at the Japanese Patent Office.[588] Otherwise, it is suggested that this is an issue to keep on the policy radar but no more.

4.6. Inventiveness Assessment of AI-assisted outputs

4.6.1. Basic rule

To be patentable, an invention must involve an inventive step. According to Art. 56 of the EPC, an invention is considered to involve an inventive step if, having regard to the state of the art, it is not obvious to the POSITA. The EPO uses the so-called “problem-solution” approach to determine whether an inventive step is present in a claimed invention.[589] In essence, this approach “is intended to give objectivity to the assessment of inventive step and to avoid *ex post facto* analysis of the prior art, i.e. hindsight.”[590]

The invention must provide a technical solution to a technical problem. Some computer-implemented inventions present both technical and non-technical features. Non-technical features may legitimately be claimed “when taken in isolation, [*if they*] do, in the context of the invention, contribute to producing a technical effect serving a technical purpose, thereby contributing to the technical character of the invention.”[591] This implies that the applicant must identify the technical effect claimed. For our purposes, “**[i]n order to show the technical effect of an AI-based invention, the application could also include objectifiable parameters and a corresponding meaningful test**.”[592]

(587) Tabrez Ebrahim, “Automation & Predictive Analytics in Patent Prosecution: USPTO Implication & Policy,” *Georgia State University Law Review* 35, no. 4 (June 1, 2019): 1188, https://readingroom.law.gsu.edu/gsulr/vol35/iss4/5.

(588) *See* Ed Lauder, “The Japan Patent Office Deploying AI to Screen Patent Application - AI Business,” *Informa - AI Business*, April 24, 2017, https://aibusiness.com/document.asp?doc_id=760272&site=aibusiness.

(589) This approach results from Rule 42(1)(c) EPC Implementing Regulations: “(1) The description shall:... disclose the invention, as claimed, in such terms that the *technical problem*, even if not expressly stated as such, *and its solution* can be understood, and state any advantageous effects of the invention with reference to the background art.” (our emphasis).

(590) Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 124.

(591) European Patent Office, *Guidelines for Examination* G-VII (5.4).

(592) Stefan Luginbuehl, “Patent Protection of Inventions Involving Artificial Intelligence,” in *Transition and Coherence Ii Intellectual Property Law*, 2020. (*forthcoming*). For a comparative study of the inventive step requirement vis-à-vis AI-generated invention in Europe, Japan and the US, s*ee* Ramalho, “Patentability of AI-Generated Inventions.”

4.6.2. The "three steps" of inventiveness

There are three main steps in the inventiveness analysis:

1. Determining the "closest prior art";
2. Establishing the "objective technical problem" to be solved; and
3. Considering whether or not the claimed invention, starting from the closest prior art and the objective technical problem, would have been obvious to the skilled person".[593]

Inventiveness, or the assessment of whether an inventive step is present in a claimed invention when considered against the relevant prior art, poses challenges that are often harder to parse than those related to novelty. It implies that the specification reveals a technical problem that the invention solves.[594] Stating a new problem is not, in itself, sufficient.[595] Identifying a new compound is also insufficient unless the applicant can identify a hitherto unidentified technical effect.[596]

Regarding the **first step**, the closest prior art is the "one single reference [that] discloses the combination of features which constitutes the most promising starting point for a development leading to the invention".[597]

On the **second step**, the "objective technical problem" is determined objectively by studying "the application (or the patent), the closest prior art and the difference (also called 'the distinguishing feature(s)' of the claimed invention) in terms of features (either structural or functional) between the claimed invention and the closest prior art, identifies the technical effect resulting from the distinguishing features, and then formulates the technical problem."[598]

Finally, the **third and last step** is identifying "any teaching in the prior art as a whole that *would* (not simply could, but would) have prompted the skilled person, faced with the objective technical problem, to modify or adapt the closest prior art while taking account of that teaching, thereby arriving at something falling within the terms of the claims".[599]

Beyond these three steps, the EPO *may* also take into account so-called **secondary indicia, "indicators" or "considerations"** when assessing arguments for and against inventive step.[600] The EPC Guidelines identify secondary indicators for "predictable disadvantage"[601], unexpected technical effect[602], and long-felt need or commercial success.[603] Importantly, these indicia are truly secondary, meaning that their presence alone is insufficient to establish inventive step; they merely "indicate the possible presence of an inventive step".[604]

(593) European Patent Office, *Guidelines for Examination*, G-VII (5).
(594) *See* EPO Boards of Appeal, case T 641/00 (Two Identities/Comvik), 26 September 2002. Online: https://www.epo.org/law-practice/case-law-appeals/recent/t000641ep1.html. See also Torremans and Pila, note x, at 173.
(595) *See* EPO Boards of Appeal, case T 1329/04 (Factor-9/John Hopkins), 28 June 2005. Online: https://www.epo.org/law-practice/case-law-appeals/recent/t041329eu1.html.
(596) *See* EPO Technical Board of Appeal, case T 939/92 (AgrEvo), 12 September 1995. Online: https://www.epo.org/law-practice/case-law-appeals/pdf/t920939ex1.pdf.
(597) European Patent Office, *Guidelines for Examination* G-VII (5).
(598) Ibid.
(599) Ibid.
(600) See, e.g., Ramalho, "Patentability of AI-Generated Inventions," 15–16.; and Daniele Fabris, "From the PHOSITA to the MOSITA: Will 'Secondary Considerations' Save Pharmaceutical Patents from Artificial Intelligence?," *IIC - International Review of Intellectual Property and Competition Law* 51, no. 6 (July 1, 2020): 685–708, https://doi.org/10.1007/s40319-020-00953-8.
(601) EPC Guidelines Part G, Chapter 7, 10.1 Predictable disadvantage; non-functional modification; arbitrary choice, https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_vii_10_1.htm ("If an invention is the result of a foreseeable disadvantageous modification of the closest prior art, which the skilled person could clearly predict and correctly assess, and if this predictable disadvantage is not accompanied by an unexpected technical advantage, then the claimed invention does not involve an inventive step (see T 119/82 and T 155/85)").
(602) EPC Guidelines Part G, Chapter 7, 10.2 Unexpected technical effect; bonus effect, https://www.epo.org/law-practice/legal-texts/html/guidelines/e/g_vii_10_2.htm
(603) EPC Guidelines Part G, Chapter 7, 10.3 Long-felt need; commercial success
(604) Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 128. For and overview of all indicia, *see* ibid. pp. 128-129.

4.6.3. How AI may change the three steps

Step 1 is typically based on a fairly straightforward data analysis, namely identifying features and finding the closest match in the dataset. In contrast, steps 2 and 3 - and the latter more than the former - may require more "judgement" on the part of the examiner.

The **use of AI technology has already changed the step 1 process**. AI is now regularly used to identify prior art, as noted in the previous section.[605] Indeed, AI systems may be used to identify the *relevant* prior art – the EPO is doing this already – for both novelty and inventive step, though the actual assessment of the inventive step is still made by a human. **Beyond a certain level of complexity of the claimed AI-assisted invention, EPO examiners would have difficulty in establishing the causal link required to a finding of obviousness**. It should, therefore, be rare for applications concerning AI-assisted outputs to be rejected on the grounds that they fail the inventive step requirement.[606] Most likely, **opposition proceedings will become the backstop** for the system to reject certain patent applications on grounds of obviousness.[607]

For the examination of applications concerning AI-assisted inventions, the current problem-solution approach followed by the EPO requires examiners to identify and assess the state of the art and the claimed invention and then ask whether the solution presented to the problem in the patent application is obvious, or not, to the POSITA.[608] Thus, **determining who is a POSITA is crucial** to the determination of the patentability of a claimed invention. Like many other legal tests (e.g. "reasonable person" "average consumer", "informed user") the POSITA is a legal fiction, a "composite entity", but it is at the very core of the patent system.

The **EPO *Examination Guidelines*** are directly relevant in this analysis, in several ways. According to the Guidelines the **POSITA** is:

(a) presumed to be a skilled practitioner in the relevant field of technology;

(b) who is possessed of average knowledge and ability;

(c) is aware of what was common general knowledge in the art at the relevant date;

(d) is presumed to have had access to everything in the "state of the art", in particular the documents cited in the search report,

(e) to have had at his disposal the means and capacity for routine work and experimentation which are normal for the field of technology in question.

(f) if the problem prompts the person skilled in the art to seek its solution in another technical field, is the person qualified to solve the problem.

(g) is in constant development in his technical field; and

(h) may be expected to look for suggestions in neighbouring and general technical fields or even in remote technical fields, if prompted to do so.[609]

In light of these *Guidelines*, **the use of AI in innovation raises a few potential questions**. A first question is whether AI systems can process more data much faster than a human applicant or examiner, and find correlations that a human may not find. Does that change the analysis of who a POSITA is? In particular, how does it relate to the "relevant field of technology" and his "average knowledge and ability"?[610] The answer is that **the use of**

(605) *See* previously in this Report 4.5.
(606) See Annex - Expert Workshop Reports.
(607) Idem.
(608) Art. 56 EPC requires that the invention be not obvious to a POSITA. *See also* European Patent Office, European Patent Guide: How to Get a European Patent. at para. 3.4.02 (updated 2019)
(609) EPO, Guidelines for Examination, G-VII(3).
(610) On this, see points *a* and *b* of the *Guidelines* cited above in the text.

AI systems could indeed change the analysis of the POSITA. This change could impact all inventiveness analyses, not just those for inventions made using AI tools.

In the examination processes, AI tools provide access to more data, and such data will be processed more "intelligently" as AI technology progresses. This will assist (human) examiners in determining novelty and obviousness. To be sure, an applicant will still have to identify the technical effect. Using an AI system may raise some differences if that AI system cannot "explain" the technical contribution it helped make – this relates to aforementioned AI "black-box" or explainability problem. This issue can be assessed during examination.

A second question is whether, to assess "common general knowledge in the art at the relevant date", access to the state of the art, and the constant development in the technical field, AI systems have access to a dataset that is limited in time?[611] In our view, this is a matter for Patent Offices to provide access to the necessary and updated dataset(s). To determine the state of the art on a given date, **AI systems should have access to all the data worldwide on that date that is relevant for a determination of inventiveness (and novelty)**.

A third question, related to the previous one, is whether AI systems have "common general knowledge"?[612] Common general knowledge is a class of prior art that refers to "knowledge which the skilled person has at his immediate disposal ..., either because he knows it or is aware of the place where to find it"; it is "the knowledge the skilled person has in his specific technical field".[613] **AI systems probably do not have "common general knowledge"**, as such notion reflects what is expected of, say, a (human) technical education in the field and gained from experience. This is why a **determination of obviousness is still a human process**.[614]

A fourth question is how to define access to "general technical" legal and "remote technical fields" in the context of AI systems, which rely on access to a limited dataset.[615] Human examiners at patent offices might limit their search by area, for example using the International Patent Classification (IPC), established by the Strasbourg Agreement 1971[616] or the Cooperative Patent Classification developed by the EPO and the USPTO.[617] A (human) POSITA is assumed to know their field of endeavor. An AI system would have no such limitation unless it were specifically programmed to look at a clearly delineated set of data. This might affect the definition of remoteness of technical fields and lead to a reconsideration of the EPO case law on the notion of "neighbouring fields", which draws a distinction between technical and non-technical fields.[618]

The answer to this question might be a combination of the answers to the two previous questions. AI systems will provide examiners with access to more and better processed data, and may better help identify the closest prior art. **Whether access to remote fields is permitted or possible for AI systems depends on programming and ML. Whether access to such remote fields was "prompted" by an application is a human determination to be made in the examination process**.

A fifth question stems from the fact that humans and AI systems have different cognitive, heuristic and epistemological processes and thus innovate differently. In that light, how does one apply points **(e)** and ***(f)*** of the Guidelines above and, in particular, determine what is "routine work and experimentation which are *normal* for the field"? As AI technology progresses, normalcy, which is necessarily a dynamic notion, will certainly evolve. However, the **cognitive functions involved in the assessment of steps 2 and especially 3** in the determination of inventiveness/obviousness a**re still human and will remain** so for the predictable future.

Also relevant in examining inventive step for AI-assisted outputs is whether to adjust the interpretation of secondary indicia. Some authors suggest "adding a 'made by AI' factor as an indication of obviousness" with reference

(611) On this, *see* points *c, d,* and *g* of the *Guidelines* cited above in the text.

(612) On this, *see* points *c* of the *Guidelines* cited above in the text.

(613) Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 115. (citing T206/83 r.5 and T939/92 r.2.3, and adding "It is defined as the knowledge an experienced man in his field is expected to have, or at least to be aware of, to the extent that he knows he can look it up if he needs it (T766/91 r.8.2). The skilled person does not necessarily have knowledge of the whole technology (T206/83 r.5)." NB this concept is relevant is assessing novelty, inventive step and sufficiency of disclosure. Id, pp. 115-116.

(614) *See also* in this respect Annex - Expert Workshop Reports.

(615) On this, *see* point h) of the *Guidelines* cited above in the text.

(616) WIPO, International Patent Classification, online : https://www.wipo.int/classifications/ipc/en/

(617) EPO, Cooperative Patent Classification (CPC), online: https://www.epo.org/searching-for-patents/helpful-resources/first-time-here/classification/cpc.html.

(618) *See* EPO, Case Law of the Boards of Appeal: Neighbouring Fields, online: https://www.epo.org/law-practice/legal-texts/html/caselaw/2019/e/clr_i_d_8_2.htm.

to "the general trend in the field of the pertinent art".[619] Others, fearing that a "wide range of new [AI-generated] pharmaceutical inventions" fall into the public domain, argue that "objective indicia of non-obviousness can serve as an important… instrument to rebut the technical findings of obviousness of new pharmaceutical inventions".[620] Even if the latter conclusion seems unjustified at the current state of technology, the **adjustment of interpretation guidelines (e.g. at EPO level) on secondary considerations to assess AI-assisted outputs might provide some help in improving the determination of inventive step for such inventions.**

4.6.4. A look ahead

The assessment of inventiveness by Patent Offices is likely to increasingly "allow for machine-augmented human capabilities [*which*] would create a heightened standard for patentability with widespread implications for all kinds of areas, not just the inventions arising from machine-human collaborations."[621] To paraphrase Professor Abbott, perhaps "inventive is the 'new skilled' and hence, the 'widespread use of [AI] should raise the bar for obviousness'".[622] In the longer term, we note Michael McLaughlin's suggestion that a "deeper fundamental understanding of 'inventorship' will be evaluated as society delves toward the inevitable depths of this 'artificial invention age,' to determine whether inventions made with the assistance of AI (computer-assisted or computer-generated inventions) *should* result in patents."[623]

Looking at future developments on the policy level, if the patent system is seen as creating incentives for inventions *that would not otherwise occur* without such incentives, innovations easily produced by readily available and widely used AI systems may not or no longer be considered inventive enough to require a patent as incentive. As Peter Blok noted,

> ***[T]he main question is whether the use of a particular artificial intelligence application is or has become a "normal" tool for routine work in the relevant field of technology…*** *On the one hand, this means that, as long as the artificial intelligence application is not "normal" in the relevant field of technology, patents can be granted if the invention is not obvious to the average skilled person without use of artificial intelligence, even if an artificial intelligence application has done most or all of the inventive work. On the other hand, when artificial intelligence becomes a standard tool for routine work, the abilities of the average skilled person are assumed to have improved accordingly and, then (sic), include the "inventive" capabilities of the artificial intelligence application. Under those circumstances, patents will not be granted if the invention is obvious to the average skilled person equipped with the artificial intelligence application, even if the inventor did not use artificial intelligence.*[624]

At this juncture, however, this is a mere dot on the distant horizon policy radar. For now, Patent Offices may be expected to adjust to the increasing capability of AI systems and be equipped with tools that reflect the current state of the art in AI.[625] Beyond this promising development to keep an eye on, no need to effect policy changes emerges on the immediate horizon.

4.7. Sufficiency of Disclosure

Under the EPC, a patent application must disclose the invention in a manner sufficiently clear and complete for it to be carried out by a POSITA[626], who may avail themselves of "contemporaneously available common general knowledge and the literature referred to in the description".[627] This requirement is based on concepts of repro-

(619) Ramalho, "Patentability of AI-Generated Inventions," 26.
(620) Fabris, "From the PHOSITA to the MOSITA."
(621) Vertinsky, "Boundary-Spanning Collaboration and the Limits of Joint Inventorship Doctrine," 432–433.
(622) Ryan Abbott, "Everything Is Obvious by Ryan Abbott," *UCLA Law Review* 66, no. 2 (2019): 33–34.
(623) Fabris, "From the PHOSITA to the MOSITA," 704.
(624) Blok, "The Inventor's New Tool," 71. (our emphasis)
(625) McLaughlin, "Computer-Generated Inventions," 5. (Internal footnotes omitted.)
(626) *See* art. 83 EPC.
(627) Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 184. adding that the "[t]he last two elements widen the scope of the disclosure".

ducibility and plausibility of the claimed invention.[628] **Does the disclosure requirement change due to the use AI of systems in the inventive activity?[629] The short answer is no, but a few interesting issues emerge none the less.**

AI-assisted claims might complicate the analysis due to the “black box” nature of certain AI systems that may make it challenging to provide a sufficiently clear and complete disclosure for the invention to be carried out by a POSITA. This may be more so in the case of process claims, but attention should be drawn to any claimed invention, whether product or process, where an AI system, or one of its outputs, is a part of the claimed invention. An AI system may help identify an innovation (or even just a lead) but it may not be able to explain why it works, or how it is making its contribution.[630] This lack of “explainability” is not necessarily fatal – applicants need not explain *why* their invention works if it is reproducible and otherwise plausible. Disclosing whether an invention was conceived with the aid of AI is only required if necessary to meet patentability criteria.

One potentially problematic case identified was that of diagnostic processes, where it may be challenging for the applicant to explain the results obtained when these come from the operation of an AI system.[631] One could draw a parallel with arts. 21 and 22 (and Recital 71) of the General Data Protection Regulation (**GDPR**)[632], which may be characterised as providing a **right of “explanation”** of decisions made by AI systems. This “right” is arguably often quite theoretical because “[n]o one really knows how the most advanced algorithms do what they do”.[633] However, **a duty to explain an inventive contribution made by the AI system could emerge. Then, instead of *ex post* explainability, standards surrounding ongoing interpretability of AI systems and processes might be useful**.

The EPO Examination Guidelines state that an invention is described not only in terms of its structure but also in terms of its function, unless the functions of the various parts are immediately apparent. [634] This is underlined by Luginbuehl explaining (in his personal capacity) that: :

> ***AI often performs its processes in a “black box”*** *which humans cannot understand to the fullest extent due to the different approach that machines take towards solving a problem. This makes it very* ***difficult for an applicant to fulfil the disclosure requirement if he is not able to explain how the invention works so that the technical solution can be re-worked on the basis of the specifications in the patent. Thus he will often be forced to define the function from a different perspective, such as from a network structure or a functional perspective***.[635]

(628) *See* in this context Rules 31 (Requirements relating to biological material), Rule 33 (Availability of Biological Material) and Rule 34 (New deposit of biological material) EPC Implementing Regulations.

(629) See also on this question Rule 42 EPC on content of the description, especially paras (c) and (e), as well as the Guidelines on Sufficiency of Disclosure F.III.1, fourth paragraph.

(630) For a discussion of the issues of black-box AI models and unpredictability of solutions, *see* Kim, "'AI-Generated Inventions,'" 453–455.

(631) *See* Annex - Expert Workshop Reports.

(632) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). See also Articles 13 to 15 GDPR as a legal basis for a right to explanation under the GDPR. On the topic, *see, e.g.,* Andrew D. Selbst and Julia Powles, “Meaningful Information and the Right to Explanation,” *International Data Privacy Law* 7, no. 4 (November 1, 2017): 233–42, https://doi.org/10.1093/idpl/ipx022; Sandra Wachter, Brent Mittelstadt, and Luciano Floridi, “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation,” *International Data Privacy Law* 7, no. 2 (May 1, 2017): 76–99, https://doi.org/10.1093/idpl/ipx005; Margot Kaminski, “The Right to Explanation, Explained,” *Berkeley Technology Law Journal* 34 (January 1, 2019), https://doi.org/10.15779/Z38TD9N83H.

(633) Will Knight, “The Dark Secret at the Heart of AI,” *MIT Technology Review*, April 11, 2017, https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/. *See also* Sandra Wachter et. al., ‘Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR’ (2018) 31 Harvard J L & Tech 841, at 851.

(634) Guidelines, F-III (1).

(635) Luginbuehl, “Patent Protection of Inventions Involving Artificial Intelligence.” (our emphasis)

Then, as Charles Clark once put it in a different technological context, for some AI inventions the "answer to the machine may be in the machine".[636] That is to say, as Professor Vertinksy explains:

> *A person working with a machine may be able to make and use an invention based on a very limited patent disclosure, meaning that even limited patent disclosures could enable a broader range of inventions. Machine-based enablement might thus allow for an expanded patent scope for any particular patent disclosure.*[637]

Like inventiveness, this is a **matter that should likely be solved primarily by patent offices**. If an invention is "sufficiently disclosed" so that readily available AI systems in a given field can determine reproducibility and plausibility, then the application of the standard may indeed change somewhat, though the standard itself remains unaffected.

Although there could be a problem in relation to sufficiency of disclosure AI outputs (for example if an AI system is part of a claimed process), **it is not clear that a deposit system for AI algorithms (similar to the existing system of deposit of biological material) would be of use** in the context of disclosure. The deposit of biological material is useful because a written description is often insufficient.[638] Whether, and more importantly perhaps, how an algorithm could be sufficiently described (again, assuming that such description is required to fulfill patentability criteria) in a patent application involving an AI-assisted invention would require a specific analysis well beyond the scope of this report.

The current regime for deposit of biological material (microorganisms) is set out in Rule 31 of the EPC Implementing Regulations. According to this, if an invention involving the use of or concerning biological material which is not available to the public and which cannot be described in the patent application in such a manner as to enable the invention to be carried out by a POSITA, then to fulfill the disclosure requirement under art. 83 of the EPC, a sample of the biological material must be "deposited with a recognised depositary institution on the same terms as those laid down in the Budapest Treaty."[639] The application must disclose relevant information on the characteristics of the biological material such as classification of the material, morphological or physiological characteristics relevant for recognition, and propagation of the material if it cannot replicate itself.[640]

As **applied to AI-assisted outputs, this regime might distinguish between standard AI systems and proprietary systems**. It might include providing information on data used in one or other aspects of the ML function. The principle remains unchanged. The patent application must disclose enough for replicability by a POSITA but not more than is the case for non-AI assisted inventions. The idea of requiring disclosure/deposit of algorithms, and possibly also of a description of the way in which the AI that assited in the inventive process was trained, including a reference to the training data and its main characteritics, training technique and method used, may be required in appropriate cases.[641] It has been suggested in this context that "every learned coefficient or weight of the model" could be required to be disclosed to meet the disclosure requirement.[642] The idea has intuitive appeal .[643] Whether this requirement would actually be useful, and to what extent would require careful consideration of several factors, including the protection of trade secrets or other confidential information.

Finally, **AI technologies might also facilitate contact between potential licensors and licensees**.[644] This kind of matchmaking between licensors and licensees interested in technology could be done by bots linking

(636) Charles Clark, The Answer to the Machine is in the Machine & Other Collected Writings (edited by Jon Bing & Thomas Dreier, 2005)

(637) Vertinsky, "Boundary-Spanning Collaboration and the Limits of Joint Inventorship Doctrine." 432-433.

(638) As WIPO explains: "Take, for example, an organism isolated in soil that has been "improved" by mutation and further selection. It would be practically impossible to describe the strain and its selection in a way that would guarantee that another skilled microbiologist would obtain the same strain." *WIPO Magazine*, April 2015.

(639) Rule 31 EPC Implementing Regulations. *See also* Visser et al., *Visser's Annotated European Patent Convention (EPC)*, 465–67.

(640) Visser et al., 466.

(641) Some sets of training data may include personal data, which could be protected under the GDPR.

(642) See Harm van der Heidjen, AI inventions and sufficiency of disclosure – when enough is enough, iam, 3 October 2019, available at https://www.iam-media.com/ai-inventions-and-sufficiency-disclosure-when-enough-enough.

(643) See Andrew D. Selbst and Solon Barocas, "The Intuitive Appeal of Explainable Machines," *Fordham Law Review* 87 (2018), https://doi.org/10.2139/ssrn.3126971; Hannah Bloch-Wehba, "Access to Algorithms," *Fordham Law Review* 88 (2020), https://papers.ssrn.com/abstract=3355776.

(644) The authors are grateful to Florent Thouvenin (University of Zurich) for this insight.

technologies with people already exploiting similar technologies, for example in different parts of the world. Whether using blockchain or any other technology, AI-assisted systems could find matches - for example, between claims in newly issued patents and existing ones belonging to third parties - and suggest matches. This is not, however, a matter requiring immediate policy attention.

4.8. Case Study: Drug discovery

In certain industries, and for certain problems, inventive machines will become the norm. For example, in the pharmaceutical industry, IBM's Watson is now identifying novel drug targets and new indications for existing drugs. Soon, it may be the case that inventive machines are the primary means by which new uses for existing drugs are researched, as the following case study illustrates.

IKTOS CASE

IKTOS is a relatively small French entity working in collaboration with major international pharmaceuticals companies and a United States-based research centre. It has developed proprietary deep-learning (AI) ligand-based de novo drug design technology to accelerate discovery and development of novel anti-viral therapies, including new molecules to fight COVID-19 . The technology uses deep generative models previously used in fields such as image and natural language processing. The IKTOS system accelerates research and reduces its costs by replacing costly in vitro tests with "in silico", namely computer-based predictions of therapeutic effectiveness. This process is largely automated. It first generates millions of potential new molecules, that are then scored based on a series of parameters. The weights used to evaluate the molecules are adjusted iteratively. IKTOS is also using AI to develop new pathways to generate specific compounds and makes this application freely available to chemical engineers, who can use it to perform retrosynthetic analysis with machine-augmented capacity in minutes.

As the legal analysis concludes, patent protection may arise if a patentable subject-matter is produced that is novel (world-wide), involves and inventive step and is capable of industrial application. Inventiveness implies that the invention offers a technical solution to a technical problem. The inventor(s) must also be named on the application and indeed the inventor has a (moral) right to be named. Finally, the invention must be described in the patent application not only in terms of its structure but also in terms of its function, unless the functions of the various parts are immediately apparent.

In this case, the application of patent law seems fairly straightforward. Whether a molecule produced by the IKTOS system is actually novel world-wide as of the date of application is a determination made by the patent office. The fact that the molecule was created by an AI system does not modify the legal standard for this analysis. Inventiveness is also determined according to the current standard. As discussed in the legal analysis, it is possible that more new inventions will be considered obvious as AI technology becomes more commonly used. However, the interplay between human programmers, chemical engineers and others in various phases of the IKTOS invention process provides strong indicia that what the IKTOS tool produces may be considered inventive This frequent human intervention at various stages also demonstrates that naming at least one human inventor will not be a barrier to patentability. Determining who the inventors are as a matter of law — if multiple persons claim that status and/or ownership of the invention — is a matter of national law. This may in theory lead to a fragmentation of answers in different member States. Some national laws in EU Member States (e.g. Germany) provide rights to employees in this context. The ownership of patent rights in the outcome of any collaboration between two or more entities (e.g. IKTOS and the US-based centre) would presumably be handled via contractual arrangements.

BENEVOLENTAI CASE

BenevolentAI is a small biotechnology company based in London. It uses biomedical data and machine learning to improve drug development at various stages of research: basic science, target identification, molecular design, and clinical outcomes. Specifically, its drug discovery platform uses AI to analyse biomedical information to hypothesise and validate new ways to treat disease and personalise drugs for patients in three ways: (1) target identification, (2) molecular design, and (3) precision medicine. Its platform has been used to research ALS, Parkinson's, ulcerative colitis and sarcopenia.

The platform illustrates strengths of deep learning and other AI techniques, namely their ability to process enormous amounts of both structured and unstructured datasets including scientific literature and then derive inferences and contextual relationships, which can help, for example, to identify new drug targets. The platform can use this ability to develop new molecules, score them according to the factors that innovators are seeking to optimise, and, in doing so, it can help to rank and select compounds for synthesis or further research. In the precision medicine space, the same general technique can be used to identify patients more likely to respond to a particular treatment, an approach that can be used to accelerate clinical trials.

As the legal analysis concludes, patent protection may arise if a patentable subject-matter is produced that is novel (world-wide), involves and inventive step and is capable of industrial application. Inventiveness implies that the invention offers a technical solution to a technical problem. The inventor(s) must also be named on the application and indeed the inventor has a (moral) right to be named. Finally, the invention must be described in the patent application not only in terms of its structure but also in terms of its function, unless the functions of the various parts are immediately apparent.

In this case, the application of patent law poses few difficulties. The AI platform is used to process more data than any human could, and it can help human researchers focus their efforts by identifying and ranking new molecules to test. The process of innovation is guided at various steps from beginning to end by human researchers. Hence, identifying one or more human inventors on the application does not seem problematic. The novelty of a molecule suggested by the AI system would be determined by the relevant patent office as it does now for any application. The intervention of an AI system in the process does not change this analysis in any substantial way.

As explained in the legal analysis, issues of inventiveness/obviousness may eventually arise if systems such as BenevolentAI's platform become widely used and de facto increase the threshold for obviousness. However, as the Report also explains, as technology progresses this is a matter which can be dealt with appropriately by patent offices and, ultimately national courts.

This case study also illustrates why, as discussed in the legal analysis, a deposit requirement for the underlying AI code would not necessarily be useful. What matters is that the patent office can determine novelty and that appropriate indicia of industrial applicability and inventiveness are provided in the application. Whether the new molecule was "invented" by an AI system, or with its assistance, matters very little in that determination. Indeed, as the case study report in the state of the art review, AI is seen in this context only as a technological input into the process of invention, "there is no difference whether data for validation is generated by human researchers or a machine learning model".

CONCLUSIONS AND RECOMMENDATIONS

5. Conclusions and recommendations

This Report examines the state of the art of copyright and patent protection in Europe both for AI-assisted outputs in general and in three priority domains: pharmaceutical research, science (meteorology) and media (journalism). "AI-assisted outputs" are meant as including productions or applications generated by, or with the assistance of, AI systems, tools or techniques. As the state of the art review demonstrates, the use of AI systems in the realms of creation, innovation and science has grown spectacularly in recent years and is expected to continue to do so in the years to come. Whereas AI is making deep inroads into media and journalism, AI systems have become almost ubiquitous in pharmaceutical research and meteorology. Moreover, outside these distinct domains, AI systems are being used to generate all manner of literary and artistic content, including translations, poems, scripts, novels, photos, paintings, etc. Likewise, a wide variety of innovative and inventive activity relies on AI systems for its development and deployment, from facial recognition to autonomous driving.

While AI systems have become – and will become – increasingly intelligent and autonomous, this Report nonetheless assumes that fully autonomous creation or invention by AI does not exist, and will not exist for the foreseeable future. We therefore view AI systems primarily as sophisticated *tools* in the hands of human operators. For this reason, we do not enquire whether AI systems should one day be recognised as a legal personality and accorded authorship or inventorship status under future IP Law – a question we leave for now to science fiction.

An important trend that emerges from the state of the art review is that more and more AI capability is being offered "as a service" rather than as "bespoke" (tailored) AI systems. Consequently, the emphasis of our analysis is on the users (operators) of AI systems, rather than on their developers. Because we assume that human intervention remains essential in AI-assisted creation and invention, our report concentrates on legally qualifying the role of the human operators or users of AI systems in respect of the AI-assisted outputs. Does this role justify protection of these output as "works" protected according to current EU copyright standards, or "inventions" under the European Patent Convention? Based on legal analysis, we answer these questions in a descriptive way, primarily by interpreting existing legal doctrines and case law.

5.1. Copyright law: summary and recommendations

As our inquiry into EU copyright law reveals, four interrelated criteria are to be met for an AI-assisted output to qualify as a protected "work": the output is (1) a "production in the literary, scientific or artistic domain"; (2) the product of human intellectual effort; and (3) the result of creative choices that are (4) "expressed" in the output. Whether the first step is established EU law is however uncertain. Since most AI artefacts belong to the "literary, scientific or artistic domain" anyway, and are the result of at least some "human intellectual effort" (however remote), in practice the focus of the copyright analysis is on steps 3 and 4.

Based on a thorough analysis of the CJEU's case law, and in light of the findings of the two experts workshops,[645] we conclude that the core issue is whether the AI-assisted output is the result of human creative choices that are "expressed" in the output. In line with the CJEU's reasoning in the *Painer* case, we distinguish three distinct phases of the creative process in AI-assisted production: "conception" (design and specifications), "execution" (producing draft versions) and "redaction" (editing, refinement, finalisation).

While AI systems play a dominant role at the execution phase, the role of human authors at the conception stage often remains essential. Moreover, in many instances human beings will also be in charge of the redaction stage.

(645) See Annexes IVA and IVB.

Depending on the facts of the case, this will allow human beings sufficient creative choice. Assuming these choices are expressed in the final AI-assisted output, the output will then qualify as a copyright-protected work. By contrast, if an AI system is programmed to automatically execute content without the output being conceived or redacted by a person exercising creative choices, there will be no “work”.

Due to the “black box” nature of some AI systems, persons in charge of the conception phase will sometimes not be able to precisely predict or explain the outcome of the execution phase. This, however, need not present an obstacle to the “work” status of the final output, assuming that such output stays within the ambit of the person’s general authorial intent.

Authorship status will be accorded to the person or persons that have creatively contributed to the output. In most cases, this will be the user of the AI system, not the AI system developer, unless the developer and user collaborate on a specific AI production, in which case there will be co-authorship. If “off-the-shelf” AI systems are used to create content, co-authorship claims by AI developers will also be unlikely for commercial reasons, since AI developers will normally not want to burden customers with downstream copyright claims. We therefore expect these arrangments to be clarified in the contractual terms of service of providers of such systems.

An authorship problem that might arise is potential false claims of authorship in respect of AI-assisted outputs that do not qualify as “works” for lack of human creativity. Producers or publishers might falsely attribute authorship in such productions in order to benefit from the authorship presumptions granted under EU law, which allow the person whose name is mentioned as an author to initiate infringement procedures.

British and Irish copyright law accord authorship status to persons undertaking the arrangements necessary for creating computer-generated works in cases where no (human) author can be identified. These provisions have been criticised as being incompatible with EU copyright standards, since “authorless works” do not meet the EU standard of “the author’s own intellectual creation”. Although this criticism is valid, we note – in line with some scholarship – that these regimes are perhaps better understood or qualified as a species of related rights, in which case they would be compatible with the EU *acquis*.

As the case studies presented in this report illustrate, it is impossible to make general assessments of the copyright status of AI-assisted outputs in individual cases. In some cases, where the creative role of human beings is evident at various stages of the creative process, such as *The Next Rembrandt* project, the output will most likely be copyright protected. In other cases, where it is difficult or even impossible to identify creative choices – such as the automatically generated match reports of Fussball.de or the weather forecasts produced by Uibimet – copyright protection will be unlikely. Note however that this is no different for simple sports reports and weather forecasts produced without any machine assistance. Nevertheless, producers of “authorless” AI-assisted outputs might find recourse in related (neighbouring) rights.

In the final part of the section, we examine these regimes. As our analysis reveals, the related rights harmonised under the EU *acquis* offer various possibilities for protecting AI-assisted outputs that do not qualify for copyright protection. In light of the general absence in related rights’ law of a requirement of human authorship or originality, and its rationale of rewarding economic or entrepreneurial activity, related rights should accommodate AI-assisted output that falls within the scope of its subject-matter in cases of insufficient human creative input.

While AI-assisted outputs in the form of aural signals (audio data) may benefit from the phonographic right, audio-visual outputs will qualify for protection under the film producer’s right. Moreover, AI-assisted broadcasts may find protection under the related rights of broadcasters. None of these related rights provide for a threshold requirement, making these regimes available for AI-assisted outputs that are generated without any creative human involvement – even absent significant economic investment. In most cases, the user – not the developer – of the AI system will be deemed the owner of the related right, since it is the user that triggers the acts that give rise to these rights by activitating the AI system.

Additionally, AI-generated databases will qualify for *sui generis* protection under the EU Database Directive if the databases are the result of substantive investment. This includes investment in AI technology and know-how applied in producing the database. In light of the broad legal notion of “database”, the *sui generis* right potentially offers protection to a wide range of AI-assisted productions. However, it is currently uncertain whether investment in the machine-generation of data – for example, the generation of weather data with the aid of AI – may be factored in. In any case, the prerequisite of a "database" rules out protection of raw machine-generated data.

"Authorless" AI-assisted outputs will remain completely unprotected only in cases where no related right or *sui generis* right is available. Since such rights attach primarily to aural and audio-visual signals, as well as to databases, such cases are most likely to occur if the AI-assisted output is in alphanumerical form (letters and numbers). Whether this absence of IP protection might justify regulatory intervention, is a primarily economic question that cannot be addressed in the context of this report. Such intervention would be justified only if no alternative protection (e.g., under trade secret protection, unfair competition or contract law) were available, *and* solid empirical economic analysis were to reveal that the absence of protection harms overall economic welfare in the EU. We have seen no such evidence.

The preceding analysis with regard to copyright and AI leads to the following **conclusions and recommendations**:

- Current EU copyright rules are generally sufficiently flexible to deal with the challenges posed by AI-assisted outputs.
- The absence of (fully) harmonised rules of authorship and copyright ownership has led to divergent solutions in national law of distinct Member States in respect of AI-assisted works, which might justify a harmonisation initiative.
- Further research into the risks of false authorship attributions by publishers of "work-like" but "authorless" AI productions, seen in the light of the general authorship presumption in art. 5 of the Enforcement Directive, should be considered.
- Related rights regimes in the EU potentially extend to "authorless" AI productions in a variety of sectors: audio recording, broadcasting, audivisual recording, and news. In addition, the *sui generis* database right may offer protection to AI-assisted databases that are the result of substantial investment.
- The creation/obtaining distinction in the *sui generis* right is a cause of legal uncertainty regarding the status of machine-generated data that could justify revision or clarification of the EU Database Directive.
- Further study on the role of alternative IP regimes to protect AI-assisted outputs, such as trade secret protection, unfair competition and contract law, should be encouraged.

5.2. Patent law: summary and recommendations

As our analysis of European patent law – and in particular the European Patent Convention (EPC) – demonstrates, the requirement that an inventor be named on a patent application means that one or more human inventors must be identified. Under the EPC regime, this is essentially a formal requirement. The EPO does not resolve disputes regarding substantive entitlement; this is an issue that is governed by national law.

Following this approach, the EPO decided two cases in 2020 (currently under appeal) where it considered that, because AI systems do not have legal personality, they cannot be named inventors on a patent application. Similar decisions were reached by the UKIPO (confirmed on appeal by the England and Wales High Court) and the USPTO (currently under appeal).

A human inventor typically has the right to be named on the application. Beyond this, inventorship and co-ownership are mostly a matter for national law. We note, however, that as AI technology stands today, the possibility that an AI system would invent in a way that is not causally related to one or more human inventors (e.g. the programmer, the trainer, the user, or a combination thereof) seems remote. We therefore focus our analysis on AI-assisted inventions rather than on fully AI-*generated* inventions. The challenge of determining whether a machine can be assimilated in some way to an inventor, employee, or right holder is a matter that will emerge, if at all, before national courts rather than the EPO. As technology stands today, no immediate action appears to be required on the issue of inventorship at EPC level.

As regards ownership, there are at least three possible (sets of) claimants to an AI-assisted invention: the programmer or developer of the AI system; the owner of the system; and the authorised user of the system who provided it with training data or otherwise supervised the training. Neither international law nor the EPC provide clear rules concerning how ownership of patents may be affected by this new type of AI-assisted inventive activity. Ownership was explicity excluded during the negotiation of the EPC. It is therefore a matter for national laws. A survey of such laws shows that because AI systems lack legal personality they cannot qualify as "employees" or "successors in title". However, that might not be an issue that requires harmonisation. This is because, on the one hand, there does not seem to be a problem in establishing a sufficient connection between an AI-assisted invention and a patent applicant, and, on the other hand, there is insufficient case law on entitlement to a patent to suggest a need for legislative intervention. Therefore, at this stage, issues of ownership concerning AI-assisted inventions are best dealt with by national laws.

A patent can only be granted if the claimed invention was new as of the date of filing (or priority date if applicable). Novelty means that the claimed invention did not form part of the "state of the art" available to the public "by means of a written or oral description, by use, or in any other way". Our analysis suggests that European patent law does not require legislative change to apply to AI-assisted inventions. However, its practical application may be changed by AI systems. Quantitative changes relate to the amount of data an AI system can process during application or examination proceedings, as compared to a human examiner. Quantitative changes involve the possibility that the novelty assessment is increasingly carried out by (or with the aid of) AI systems. These quantitative and qualitative changes suggest the need for IPOs and the EPO to invest in order to mirror the technical capability of sophisticated patent applicants. This is liable to affect the examination of all patent applications, not just those concerning AI-assisted inventions.

A claimed invention must also involve an inventive step, that is, it must not be obvious (again, as of the date of filing) to a person of ordinary skill in the art ("POSITA"). In other words, an invention does not involve an inventive step if it is obvious to a POSITA. The three main factors to make this analysis under the EPC involve: 1) determining the closest pior art; 2) establishing the objective technical problem to be solved; and 3) considering whether the claimed invention, starting from the closest prior art and the objective technical problem, would be obvious to the POSITA. In addition, the EPO may also take into account secondary indicia.

Although it may not require legislative change, the use of AI technology affects the assessment of these factors in practice. For instance, some claimed AI-assisted inventions may be too complex for EPO examiners to establish the causal link required in step 3 above. That is to say, the link between the closest prior art and objective technical problem, on the one hand, and the obviousness of the invention to the POSITA, on the other. The result is that EPO examiners will rarely reject a claim for such AI-assisted inventions on the basis of lack of inventive step. In those cases, possible opposition proceedings could operate as a backstop to further assess the inventive step for AI-assisted inventions. In addition, the use of AI technology may have an impact on the POSITA, e.g. regarding the type of data that can be accessed to determine the state of the art. Still, many aspects of the analysis point towards the conclusion that the determination of obviousness is still to a great extent a human process.

In this light, it is our view that current law is capable of handling changes caused by the increasing use of AI in innovative processes. That being said, as with novelty, the practical application of inventiveness may change along the lines mentioned above, meaning that certain claimed inventions may be "obvious" to AI systems. This change, if it happens, will likely emerge as and when EPO Boards of Appeal and national courts are asked to enforce or invalidate patents. This may justify adjustments of EPO guidelines to improve the determination of the inventive step in AI-assisted inventions. This could be particularly helpful regarding the definition of the POSITA and the adjustment of secondary indicia.

Finally, a patent application must sufficiently disclose the invention. One question that emerges in this respect is how to apply the disclosure requirement for AI-assisted inventions. AI-assisted claims – and this may apply even more directly to certain *process* claims – might complicate the application of this requirement due to the "black box" nature of certain AI systems that may make it challenging to provide a sufficiently clear and complete disclosure for the invention to be carried out by a POSITA. This lack of "explainability" is problematic but not fatal, since applicants are not typically required to explain why the invention works as long as it is reproducible and otherwise plausible. One potentially problematic case is that of diagnostic processes, where it may be difficult to explain the results obtained if these originate from an AI system. Inspired by the GDPR's right of explanation, we consider the possible emergence of a duty to explain an inventive contribution made by an AI system in appropriate cases, which would focus on ongoing interpretability rather than *ex post* explainability.

Proposals have been floated to provide a mechanism to deposit AI algorithms and information on training data and methods, somewhat similar to the deposit system for biologicals i.e. microorganisms (the Budapest Treaty). Although it is as yet unclear that a deposit system for AI algorithms would be useful in this context, it seems advisable to at least consider the possibility of requiring applicants to provide this type of information (whether as a "deposit" or via some other legal mechanism), while maintaining sufficient safeguards to protect the applicant's confidential information to the extent it is required under EU or international rules. That is to say, as applied to AI-assisted inventions or outputs, such a regime could distinguish between standard AI systems and proprietary systems. Depending on the type of invention and AI system used, such a regime could require the disclosure or deposit of algorithms and/or a description of some of all of the following: the training of the relevant model(s), training data and its main charactecristics, training technique and methods used, the coefficients and weights of the model.

Finally, it may be useful to point out that inventions that might otherwise be patentable can be protected under undisclosed information laws and doctrines, including, at the international level, art. 10*bis* of the Paris Convention and art. 39 of the TRIPS and, at the EU level, the 2016 Trade Secrets Directive.[646] This is relevant because it may modify the role that patents play. Patents confer a right without having to incur the expense of keeping the invention secret but the protection is often more fragile as it may cease with disclosure. Hence, the choice of a patent is not advantageous for all inventors, which is one reason why secrecy abounds even in domains of inventive activity in which patent protection is obtainable.[647]

The preceding analysis with regard to patent law and AI leads to the following **conclusions and recommendations**:

- The EPC is currently suitable to address the challenges posed by AI technologies in the context of AI-assisted inventions or outputs.
- While the increasing use of AI systems for inventive purposes does not require material changes to the core concepts of patent law, the emergence of AI may have practical consequences for national Intellectual Property Offices (IPOs) and the EPO. Also, certain rules may in specific cases be difficult to apply to AI-assisted outputs and, where that is the case, it may be justified to make minor adjustments.
- In the context of assessing *novelty*, IPOs and the EPO should consider investing in maintaining a level of technical capability that matches the technology available to sophisticated patent applicants.
- In the context of assessing the *inventive step*, it may be advisable to update the EPO examination guidelines to adjust the definition of the POSITA and secondary indicia so as to track developments in AI-assisted inventions or outputs.
- In the context of assessing *sufficiency of disclosure*, it would be useful to study the feasibility and usefulness of a deposit system (or similar legal mechanism) for AI algorithms and/or training data and models that would require applicants in appropriate cases to provide information that is relevant to meet this legal requirement, while including safeguards to protect applicants' confidential information to the extent it is required under EU or international rules.
- For the remaining potential challenges identified in this report arising out of AI-assisted inventions or outputs, it may be good policy to wait for cases to emerge to identify actual issues that require a regulatory response, if any.

(646) Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ L 157, 15.6.2016, p. 1–18 (Trade Secrets Directive).

(647) Landes and Posner, *The Economic Structure of Intellectual Property Law*, 295.

GLOSSARY OF ABBREVIATIONS

AI	Artificial Intelligence
AIaaA	Artificial Intelligence as a Service
AIPPI	Association for the Protection of Intellectual Property
AI-HLEG	EU AI High-Level Expert Group
CJEU (or Court)	Court of Justice of the European Union
CLSPA	Conseil supérieur de la propriété littéraire et artistique
EC	European Commission
EPO	European Patent Office
EU	European Union
EP	European Parliament
GUI	Graphic user interface
IPO	Intellectual Property Office
JURI	EP Committee on Legal Affairs
ML	Machine Learning
MFN	Most-Favoured Nation
IP	Intellectual Property
IP5	Forum of the five largest intellectual property offices in the world – EPO, Japan Patent Office (JPO), Korean Intellectual Property Office (KIP), State Intellectual Property Office of the People's Republic of China (SIPO) and the USPTO)
POSITA	Person skilled (or having ordinary skill) in the art
SOTA	State of the art report on AI (in this Report)
Study	This research was carried out in response to Call for Tender SMART 2018/0052
UKIPO	United Kingdom Intellectual Property Office
USPTO	United States Patent and Trademark Office
WIPO	World Intellectual Property Organisation
WTO	World Trade Organisation

ANNEXES

ANNEX IA - Scoping: Detailed analyses of application areas

Pharmaceutical research

AI applied to finding molecular targets

A “drug target” is a (macro)molecular structure with which an active pharmaceutical ingredient interacts and thus triggers its effects. Most drug targets are proteins: receptors, enzymes, transporters and ion channels. Nucleic acids and structural proteins can also be targets. Pharmaceutical chemists often search libraries containing thousands of small organic molecules to find some that have a desired effect on cells. This process to determine exactly how these molecules work, and what targets they hit, can be very time-consuming and labour-intensive. AI can help to significantly speed up this process.

Weill Cornell Medicine scientists have recently developed a machine-learning algorithm capable of predicting biological targets of prospective drug molecules[648]: BANDIT (Bayesian analysis for the determination of drug interaction targets) is designed to use virtually all available data on any potential drug to predict which enzyme or receptor or other target in the cells it interacts with to achieve its therapeutic effect. BANDIT can greatly streamline the time- and labour-intensive process of target identification by limiting the possibilities to perhaps only one or two targets, which can then be studied in detail.

AI applied to finding a ‘hit’ or ‘lead’

In pharmaceutical research, automated high-throughput screening has established itself as a method for finding so-called ‘hits’ and thus candidates for lead structures (‘leads’). A lead is a chemical substance that is investigated in pharmaceutical research in the course of drug design as a starting point for the development of a drug candidate. A lead already shows a desired biological effect in vitro, but does not yet have the qualities required for a drug substance in terms of potency, selectivity and pharmacokinetics. A lead differs from any hit from high-throughput screening in that its structure allows the synthesis of analogue molecules. A lead can also be a natural substance whose pharmacological properties must be improved by chemical modifications.

The research robot “Eve”, which was presented in 2015 in Journal of the Royal Society Interface, uses statistical models and machine learning to produce and test assumptions, check observations, conduct experiments, interpret results, change hypotheses and repeat this again and again. In this way, the robot can predict promising substances and thus make the process of finding lead structures more efficient[649]. With the help of this robot, researchers found out in 2018 that Triclosan, which is also used in toothpaste, could combat malaria infections in two critical stages, namely the infestation of the liver and blood. Following this discovery achieved by using AI, the possibility of developing a new drug emerged[650].

AI in drug repurposing

Many drugs are small, flexible molecules that not only interact with the originally determined target structure (their target) - usually a protein - but can also interact with other biological structures. Such “side effects” have already led to the identification of new applications in numerous cases. If such substances are reused for new indications, essential data on safety, side effects and interactions in the human organism are already available and the mechanism of action is known, so that the duration and costs of development can be significantly reduced. One example is the sedative thalidomide (Contergan), which was sold millions of times in the 1960s and caused

(648) https://news.weill.cornell.edu/news/2020/01/al-algorithm-predicts-drug-targets-leads-to-promising-results-for-experimental-cancer
(649) https://phys.org/news/2015-02-artificially-intelligent-robot-scientist-eve.html
(650) https://www.independent.co.uk/news/science/triclosan-malaria-toothpaste-ai-robot-scientist-university-cambridge-drug-resistant-a8170991.html

severe birth defects in newborns, but is now successfully being used against leprosy and certain cancers (including the blood cancer myeloma).

AI can help in identifying such candidates by screening existing databases such as the Drug Repurposing Hub of the Broad Institute.[651] A large team[652] from MIT, Harvard University and the MIT/Harvard Broad Institute has demonstrated this by applying ML models to identify completely new antibiotically-active substances. The ML models can analyse molecular structures and correlate them with certain properties - such as the ability to kill bacteria - and convert them into algorithms. The key aspect is that the algorithm learns without needing to know how substances work[653].

In this work, the algorithm - a self-learning neural network – was fed with data on the structure of 2,500 molecules, 1,700 of them from a database of the US Food and Drug Administration (FDA), the other 800 from a database for natural substances. After the algorithm was trained to recognise patterns, it was released on the much larger Broad Institute's Drug Repurposing Hub, a database of about 6,000 molecules, many of had been developed or discovered without ever being marketed as drugs. The algorithm scanned all the molecules in the database to find those that possessed two previously learned properties - and found them in a molecule, now called halicine, which was originally developed as a diabetes drug, but was never approved due to its lack of efficacy. Another algorithm was used to examine halicine to see how well it would be tolerated by human cells. The researchers tested the new drug named Halicin *in vitro* on a number of E.coli strains carrying resistance genes against antibiotics such as polymyxins, chloramphenicol, beta-lactams, amine glycosides and fluoroquinolines and then on dozens of bacterial strains and isolates from infected patients. With the exception of Pseudomonas aeruginosa, which affects the lungs, Halicin was able to kill many resistant germs[654].

AI in polypharmacology

Drug development in recent decades has been guided by the paradigm "one drug - one target structure - one disease". This principle is based on the approach that a pharmacological agent is used to selectively address a target structure in the body in order to treat a specific disease pattern. A large number of serious illnesses, such as cancer, mental disorders or cardiovascular diseases, often cannot be explained by a single biological defect, but only the interplay of complex and usually networked malfunctions that together produce the clinical picture. A large number of active substances already on the market interact with several targets in pharmacologically-relevant concentrations, whereas selective substances are the exception rather than the rule. These findings can contribute to the understanding of possible side effects and at the same time open up new paths in drug development. The understanding of multitarget drugs and the resulting efforts to rationally design new multitarget ligands is being investigated through the growing research area of polypharmacology[655].

DeepDDI is a novel AI platform that can accurately predict drug–drug interactions (DDI) types for a given drug pair. Looking at the DrugBank DDI dataset covering 192,284 DDIs contributed by 191,878 drug pairs, the deep learning method, developed by researchers at the Korean Advanced Institute of Science and Technology (KAIST), predicts 86 DDI types with a mean accuracy of 92.4%. DDI types are generated in the form of human-readable sentences that describe changes in pharmacological effects and/or the risk of adverse drug events (ADEs) from the interaction between two drugs in a pair. For example, DeepDDI output statements describing potential interactions between oxycodone (opioid pain medication) and atazanavir (antiretroviral medication) were generated as follows: "The metabolism of Oxycodone can be decreased when combined with Atazanavir" and "The risk or severity of adverse effects can be increased when Oxycodone is combined with Atazanavir." Through this process, DeepDDI can provide more specific information on drug interactions far beyond that reported to date for the risk of occurrence of DDIs or ADEs. DeepDDI can also be used to suggest which drug or food to avoid in order to minimise the chance of adverse drug events or to optimise drug efficacy[656].

(651) https://www.broadinstitute.org/drug-repurposing-hub
(652) led by Regina Barzilay and James Collins of MIT, see http://news.mit.edu/2020/artificial-intelligence-identifies-new-antibiotic-0220
(653) https://www.pharmazeutische-zeitung.de/neues-superantibiotikum-ex-silico-115843/
(654) https://www.theguardian.com/society/2020/feb/20/antibiotic-that-kills-drug-resistant-bacteria-discovered-through-ai
(655) Jalencas, X.; Mestres, J. On the Origins of Drug Polypharmacology. Medchemcomm2013, 4, 80.
(656) https://www.americanlaboratory.com/349165-DeepDDI-Improves-Prediction-of-Drug-Drug-Drug-Food-Interactions/

AI to find vaccines

Currently, more than 100 companies worldwide are researching a vaccine against the covid-19 virus[657]. Normally, the development of a vaccine takes many years, however development of a covid-19 vaccine is being driven at a faster pace due to the urgency for the need for a vaccine and the great competition between pharmaceutical companies, an urgency that will shorten the test phases. Artificial intelligence is potentially a strong tool in the development of vaccines.

On the one hand, artificial intelligence helps with gene sequencing. This means decoding the genetic material of the virus. On the other hand artificial intelligence can also help with simulation of vaccines. In a simulation possible vaccines can be released on the virus. The data on the nature of the virus and the vaccine are sufficient to simulate how they react to each other. This makes it possible to exclude less promising vaccines even before the actual tests begin.

The Alphabet subsidiary DeepMind is currently using Deep Learning to investigate the protein structures of the virus. These are crucial in the search for a vaccine and drugs because the antibodies of a vaccine must act precisely on the proteins of the virus in order to neutralise it. In drug research, this is a lengthy process, but it can be accelerated by algorithmic calculations[658].

Data to support AI

There are many databases, such as ZINC[659], PubChem[660], Ligand Expo[661], KEGG[662], ChEMBL[663], DrugBank[664], STITCH[665], BindingDB[666], Supertarget[667], PDB[668] that are available to integrate diverse information of molecular pathways, crystal structures, binding affinities, drug targets, disease relevance, chemical properties and biological activities[669]. AI could be used to harvest information from these databases to design new compounds both for single targets but also in polypharmacology[670].

In addition to research driven platforms also many pharmaceutical companies are now involved in initiatives such as IMI[671], Allotrope[672] and Pistoia[673], which help to ensure that research data, for example, is prepared in such a way that it can be used across companies. For example, that they are machine-readable and retrievable. That sounds simple, but it involves an immense amount of effort. According to experts, such initiatives are very important if the strengths of artificial intelligence in the industry are to be exploited one day. For example, by giving companies access to the necessary amount of clinical data to train the algorithms so that reliable analyses can be carried out later. It is too expensive for one company alone to build up a sufficient data pool. Incidentally, the companies do not give up their corporate knowledge, their intellectual property, in such cooperations. One can imagine that they give their algorithm to other companies for training - and get it back better trained.

Data for AI is also provided by private companies. Innoplexus[674] from Eschborn in Germany, which was founded in 2011 and now has around 250 employees, sees itself as a specialist that can make both the necessary software

(657) https://www.netzpiloten.de/kuenstliche-intelligenz-gegen-corona/
(658) https://t3n.de/news/corona-diese-kuenstlichen-suchen-1272580/
(659) https://zinc.docking.org/
(660) https://pubchem.ncbi.nlm.nih.gov/
(661) http://ligand-expo.rcsb.org/
(662) https://www.genome.jp/kegg/
(663) https://www.ebi.ac.uk/chembl/
(664) https://www.drugbank.ca/
(665) http://stitch.embl.de/
(666) https://www.bindingdb.org
(667) http://insilico.charite.de/supertarget/
(668) http://www.wwpdb.org/
(669) Mak K.K., Pichika, P.P. (2019), Artificial intelligence in drug development: present status and future prospects, Drug Discovery Today,Volume 24, Number 3 March 2019, 778.
(670) Ibid.
(671) https://www.imi.europa.eu/
(672) https://www.allotrope.org/allotrope-framework
(673) https://www.pistoiaalliance.org/
(674) https://www.innoplexus.com/

and publicly available information available to financial service providers and pharmaceutical companies, among others, with the help of structured data pools and artificial intelligence. According to the company's own statements, the company's software can be used, amongst other things, to analyse more than 27m specialist publications, 1m doctoral theses and 375,000 clinical studies. Innoplexus believes that preclinical research in particular can be considerably accelerated by such systems[675].

AI in clinical trials

AI could also support clinical trials by detecting the disease in patients, identifying the gene targets and predicting the effect of the designed molecule and the on- and off-target effects. AiCure is a novel AI platform that has already been used as a mobile application to measure drug adherence in a Phase II study in patients suffering from schizophrenia. AiCure[676] has increased adherence by 25% compared to traditional "modified, direct-observed therapy"[677].

Patient selection is a critical process in clinical trials. In clinical studies, the relationship between human biomarkers and in vitro phenotypes is essential. It allows a more predictable, quantifiable evaluation of the uncertainty of therapeutic responses in a given patient. The development of AI approaches for the identification and prediction of human relevant disease biomarkers enables the recruitment of a specific patient population in Phase II and Phase III clinical trials. The success rate in clinical trials can thus be increased by predictive modelling using AI in the selection of a patient population[678].

Science/Meteorology

AI in Numerical Weather Prediction (NWP) Models

For numerical weather forecasts, equations from flow and thermodynamics are used to make statements about the future state of the Earth's fluid envelope, using Numerical Weather Prediction (NWP) Models. The starting point of the forecast is a snapshot of the (current) initial state of the envelope. Atmospheric, hydrospheric and biospheric effects, sea basin and ice masses, the influence of the sun and many other factors are considered and statistically processed. On the basis of this assimilation of several million weather measurements from all over the world, the future state of the geosystem is calculated. In addition to precipitation, temperatures, cloud formation or soil moisture, even the height of ocean waves can be predicted. As is well known, forecasts are never absolutely reliable. The weather forecast is no exception. One of the reasons for this uncertainty is that there is an inadequate picture of the initial state of the geosystem. It is not possible to measure everything everywhere. Models are always only approximations to reality.

The ability of ML to extract knowledge from complex and extensive databases makes it particularly suitable for numerical weather forecasting. In addition, such methods can play a role in preparing observations, optimising workflows and modelling complex physical processes, thus speeding up calculations. These methods are already used to estimate soil moisture curves in data assimilation. The methods and their practical application are still in their infancy, but already show great potential. At the same time, the spectacular rise of AI raises questions about the role of the meteorologist, which may require adjustments, since many tasks that are currently performed manually by a human being will likely be automated in the future.

While ML has been mainly used in the past to improve numerical models, increasing discussions on attempts to tackle weather forecasting itself can be noted in the recent literature. The ideal solution is to use machine learning, and especially deep learning, to completely replace the numerical weather prediction models. Whether, and when, this will happen, however, is still the subject of intensive debate amongst experts. It is also seen as a very

(675) https://www.handelsblatt.com/unternehmen/industrie/novartis-roche-pfizer-merck-big-pharma-setzt-auf-big-data-gegen-die-forschungsflaute/22931524.html?ticket=ST-11784262-pcXSf0H6exIAS55WGPe1-ap3

(676) https://aicure.com/

(677) Bain, E.E. et al. (2017). Use of a novel artificial intelligence platform on mobile devices to assess dosing compliance in a Phase 2 clinical trial in subjects with schizophrenia. JMIR Mhealth Uhealth 5, e18

(678) Perez-Gracia, J.L. et al. (2017). Strategies to design clinical studies to identify predictive biomarkers in cancer research. Cancer Treat. Rev. 53, 79–97; Deliberato, R.O. et al. (2017) Clinical note creation, binning, and artificial intelligence. JMIR Med. Inf. 5, e24

attractive idea in climate research to use ML or DL to reproduce very costly general circulation models (GCMs). Both ideas have already been tested for simplified realities with some success.[679]

AI in Post-Processing of Weather Data

Weather data are often subjected to additional processing steps to improve quality. The main reasons for post-processing are quality control (detection and removal of errors, filling of gaps), accuracy and transformation (turning a signal (e.g. reflection) into a meteorological (radiation) or other value)[680].

Post-processing methods are therefore applied to different data sources: measurements (quality control), observations (transformation, interpolation) and the weather model itself (i.e. downscaling, multimodel, etc.) or a combination of those (e.g. Model Output Statistics[681]). Most post-processing methods correct systematic errors in the raw forecast and, from a machine learning perspective, post-processing can be viewed as a supervised learning task[682].

A recent study from the **Maximilian Ludwigs University** (MLU) in Munich shows how ML can eliminate systematic errors in numerical weather forecasts[683]. All numerical weather forecasts have systematic errors. Before the forecasts can be passed on to users, they must be calibrated by statistical means. Traditionally, this has been done with simple linear methods. In the study by MLU, modern AI methods were used. Neural networks offer a flexible method to model nonlinear relationships between different predicted variables and station-specific properties. The neural network provides better predictions than previous methods and is faster. The MLU study shows the potential of neural networks in the field of post-processing of numerical weather forecasts.

MeteoBlu AG[684] from Switzerland uses ML for post processing of output from numerical weather forecast models using actual weather measurements. On its website and in a leaflet, it presents its neural network-based Meteoblue learning multimodel (mLM) which it describes as follows: "The mLM reads actual weather measurement data and selects the best simulation model to create a forecast. Currently the mLM is validated and implemented for air temperature, dew point temperature, precipitation and wind speed. The development of the mLM for further weather variables is planned".

The model accuracy of the mLM was tested for one year before its introduction in August 2018. The accuracy of the operational mLM was verified over a period of two months (September and October 2018) at more than 30,000 meteorological stations worldwide. The analysis showed a model accuracy of 1.2 K[685] for 24-hour forecasts (24h) and an even better model accuracy of 2.0 K for six-day forecasts. By way of comparison, the established standards are:

- 0.8 K better than with "stand-alone" numerical weather forecast models (24-hour forecast).
- 0.3 K better than model simulations with MOS.
- 0.3 K better than the reanalysis model ERA5 (which uses measurements for model correction).

The model accuracy of the 24-hour forecast of mLM is therefore significantly higher than the established standards. It could also be seen that the six-day air temperature forecast of the mLM is as good as the 24-hour forecast of "stand-alone" (raw, unprocessed) forecasts of numerical weather forecast models. This improvement corresponds to the average improvement achieved by weather forecasting every ten years over recent decades" [686].

(679) Scher, S., Messori, D. (2019), Weather and climate forecasting with neural networks: using general circulation models (GCMs) with differentcomplexity as a study ground, Geosci. Model Dev., 12, 2797–2809
(680) Göbel M. U. (2018). Statistische Nachbearbeitung eines numerischen Wettervorhersagemodells mit neuronalen Netzwerken. Innsbruck
(681) Model Output Statistics (MOS) is a statistical procedure in modern weather forecasting that was developed in the USA in the 1960s/1970s. It often involves multi-linear regression equations that are applied to numerical weather models. Nowadays, MOS methods are used worldwide and serve as an aid especially for local weather forecasting.
(682) Rap et al. (2018). Neural networks for post-processing ensemble weather forecasts
(683) Neural Networks for Postprocessing Ensemble Weather Forecasts. Rasp, S. and S. Lerch, 2018. Monthly Weather Review. https://doi.org/10.1175/MWR-D-18-0187.1
(684) https://www.meteoblue.com/de/
(685) K represents the Mean Absolute Error and is a measure of the quality of the forecasts
(686) MeteoBlu (2019), meteoblue learning multimodel (mLM) – a quick introduction

Weather forecasting in predictive analytics

Predictive analytics uses data, statistical algorithms and machine learning to predict the probability of future outcomes based on historical data. The goal is to use an analysis of past events to best predict what will happen in the future. Several companies have started in recent years to offer such solutions based on weather data.

UBIMET[687] a Vienna-based (Austria) company uses many thousands of weather influencing factors and many different data sets for forecasting. These include retailer data with day- and branch-specific sales, information about special promotions and general shopping frequency at the location. In addition, UBIMET uses machine learning to compare long-term high-resolution weather data with retail data. Special algorithms help to significantly improve purchasing and warehousing. The formulas reduce the error rates in customer and store environment forecasts by 10 to 12 percent, through knowledge of meteorology and customer behaviour.

WindAI[688] is a wind analysis and prediction tool for operators of offshore wind farms. It uses predictive analytics to reduce maintenance costs for wind turbines without losses in energy production. Offshore wind farms are in an environment of high-velocity winds. Abrupt changes in wind speed and direction lead to increased wear on the turbines, which in the long term leads to higher maintenance costs. WindAI generates predictions of future wind patterns based on earth observation (EO) data and combines them with turbine-related data to achieve a more even power flow in the wind farm. Machine learning is used to optimise both the positioning and the switch-on and switch-off times for each individual turbine in a wind farm on the basis of the wind forecasts.

Since 2017, the start-up **Kaiserwetter**[689] from Hamburg (Germany) has been using data on the actual output and the expected output of renewable energy plants - both from a technical and a financial point of view. Its cloud-based Internet of Things (IoT) platform Aristoteles aggregates, structures, visualises and analyses historical and real-time data from wind and solar parks in eight countries on three different continents. General values such as weather data and, in future, data from various electricity markets will also be included in the analyses. With the help of predictive analytics and machine learning, precise forecasts can be made: What yield can be expected from a plant in the next month? Where are there technical problems? Which plants remain below their potential? This can also provide valuable information for operators - especially with regard to how reliably their plants are operating. If important components of a turbine threaten to fail, they can initiate repairs or bring forward maintenance processes at an early stage.

StormGeo[690] (formerly Storm Weather Centre) is a company specializing in the delivery of weather data, with headquarters in Bergen (Norway). With 25 offices in 15 countries, the company is able to continuously provide advanced solutions around the clock to a global customer base in the shipping, offshore, renewable energy, media, corporate and aviation industries. In eight global forecast centres, key figures related to climatic conditions such as temperature, precipitation, humidity, cloud cover, wind and/or waves are analysed around the clock. These data form the basis for advice on weather-sensitive processes: How does the drilling platform relate to the sea state? What is the best shipping route given the weather conditions? And which location is suitable for a wind farm? The company uses predictive analytics and weather data to improve operational efficiency in supply chain logistics. It helps freight and logistics companies reduce weather-related risks, delays and costs throughout the supply chain. StormGeo's fleet management solution uses deep leaning algorithms to provide operational solutions such as weather impedance calculators, estimation of weather route delays, and refined ETA (Estimated Time of Arrival) prediction. Additionally, the solutions are fully supported by StormGeo's meteorologists and data scientists through its global network of weather centres.

AI in Nowcasting

Nowcasting, means 'now-prediction', i.e. very short-term forecasting in the period from minutes to two hours in advance. For this purpose, horizontal and vertical movements in the atmosphere and also the intensity of already existing or just emerging weather systems such as showers, thunderstorms, snowfall areas, etc. are extrapolated.

(687) https://www.ubimet.com
(688) https://www.windai.net/
(689) https://www.kaiserwetter.energy/en/
(690) https://www.stormgeo.com/

For this area of forecasting, the forecaster is fully challenged despite all automation of data preparation and presentation. Nowcasting requires completely different methods from those of weather forecasting. For example, numerical weather forecasting cannot be used for such short periods of time[691].

Scientists at the **Swiss Federal Institute of Technology** in Lausanne (EPFL) have been recently using machine learning methods to predict lightning strikes within a radius of 30 kilometres with an accuracy of 10 to 30 minutes. In doing so, they used information that has been historically recorded for many years by ordinary weather stations. For their detection and prediction procedure, EPFL researchers evaluated data from twelve Swiss meteorological recording systems from a total of ten years. The data came from urban regions as well as from rural and mountainous areas. Using the four weather parameters of temperature, relative humidity, wind speed and atmospheric pressure, as well as geodata on meteorological events, the researchers created the algorithms that enable them to predict lightning strikes with almost 80 percent accuracy[692]. Such data is useful, for example, in the aviation industry to enable aircraft to avoid probable lightning strike zones.

Google has presented a AI-based nowcasting approach that requires far less data than previous systems and, according to the developers, can adjust its forecasts almost immediately. At the heart of this new methodology is a neural network that Google has trained with radar data collected by NOAA (the US National Oceanic and Atmospheric Administration) between 2017 and 2019. Basically, this training was the solution to a kind of advanced image recognition problem. The system is said to have learned how weather works by analysing the progression of the data in the context of the laws of nature. This seems to have worked, with convincing results being reported by Google. In tests, Google's new AI approach was able to prevail over all classic forecasting models. The forecast generated by AI is able to react to current changes in the weather situation within minutes and adjust the forecast accordingly.

Journalism

Assistive Technologies

AI supports journalists in researching and creating media content through automated analyses. This allows for more efficient work, as time-consuming, repetitive activities and research tasks are automated. In addition, AI technologies make it possible to create higher quality work.

Increases in effectiveness through AI result, for example, from the automation of audio transcriptions. Transcription services automatically convert the spoken word into written text, resulting in significant time savings for journalists. Following an interview, a complete transliteration is available, enabling journalists to focus to a greater extent on content analysis.

Speech recognition services that already use AI include Sonix[693], Trint[694], or the cloud services Speech-to-Text from Google[695] and Amazon Transcribe[696].

AI-based content analysis also helps journalists to discover new stories and manage their material. Some of the applications include tools for information extraction, clustering, checking the integrity of tweets[697], summarising, and machine translation capabilities for multi-lingual access to sources. For example, the DataMinr[698] tool searches millions of tweets on Twitter and attempts to filter out potentially news-worthy information by detecting unusual patterns.

(691) https://www.wmo.int/pages/prog/amp/pwsp/Nowcasting.htm
(692) https://actu.epfl.ch/news/using-ai-to-predict-where-and-when-lightning-will-/
(693) https://sonix.ai/
(694) https://trint.com/
(695) https://cloud.google.com/speech-to-text/
(696) https://aws.amazon.com/transcribe/
(697) Reuters uses the "News Tracer", an algorithm-based forecasting tool, to check the credibility of Twitter tweets. This involves evaluating "credibility" and "news value". https://www.reuterscommunity.com/topics/newsroom-of-the-future/reuters-news-tracer-filtering-through-the-noise-of-social-media/
(698) https://www.dataminr.com/

Data journalistic AI applications, for example, take over the extraction of relevant information from media content. In doing so, extensive data sets ("Big Data") are automatically evaluated and visualised. These automated analysis procedures also enable the realisation of journalistic projects of ever-increasing scope.

One example of a large-scale journalistic project in which the use of AI-based analysis procedures was indispensable is the Panama Papers research by the Süddeutsche Zeitung and the International Consortium for Investigative Journalists (ICIJ). Through a data leak from the Panamanian law firm Mossack Fonseca, a total of 2.6 TB of (unstructured) data came into the public domain. Only AI-based big-data analysis made a detailed evaluation of this data possible. AI tools for data processing and conversion were used, resulting in, among other things, a faster revelation of relationships. Visualisation software was used to graphically depict complex relationships between the actors involved[699].

AI is also used for photo archives, e.g. at the New York Times[700] where cloud and AI applications are used to digitally backup their extensive photo archive. With the help of various Big Data and machine learning methods, the search for archived photos is made easier by assigning each image to thousands of categories, capturing individual objects and faces, and recognising printed words in the images. For the purpose of recognition, both the back and front of the images are scanned, analysed and, if necessary, contextualised.

Generative Technologies

Photography

In photography, the most modern form of AI, so-called Generative Adversarial Networks[701] (further: GANs) are used. Here, two neural networks train each other: A "generator" algorithm generates a multitude of new data, e.g. photos from a data set of image files. A second so-called "discriminator" algorithm tries to identify the generated images in a series of real images. This process is repeated until the images created by the "generator" are indistinguishable from the real images[702].

GANs are extremely powerful and, unlike classic neural networks, do not require comprehensive training data sets. GANs are already being used to create photorealistic images[703], to model motion patterns in videos, to create 3D models from 2D images and for video games. Facebook also uses GANs[704].

Only recently, OpenAI has demonstrated that in the future there might be an alternative to GANs. It applied unsupervised learning to an image AI: Image GPT (iGPT) has learned to visually complete half submitted images. This way, a whole cat is created from half a cat - without iGPT ever seeing the word "cat" on a cat image.[705]

Potential journalistic uses of automated images include visualisation of events for which no real images or videos are (immediately) available. Such images could be suitable to visualise news reports in a more media effective way and represent an alternative solution to archive images. For example, there is already an AI solution that automatically creates the finished dishes on a plate as a "photo" from the information provided by cooking recipes alone, without the need to cook[706].

Videos

Videos can also be created largely automatically. The two most successful software solutions, Wochit[707] and Wibbitz[708], allow video contributions to be created based on texts or (news) articles in just a few steps. To create these

(699) Torinus, G. (2016). Die Entschlüsselung der Panama Papers, online: https://www.heise.de/tp/features/Die-Entschluesselung-der-Panama-Papers-3267916.html
(700) https://www.cnet.com/news/google-ai-helps-nyt-get-a-handle-on-5-million-photo-archive/
(701) https://pathmind.com/wiki/generative-adversarial-network-gan
(702) https://towardsdatascience.com/understanding-generative-adversarial-networks-gans-cd6e4651a29
(703) https://www.cdotrends.com/story/14300/rise-ai-supermodels
(704) https://research.fb.com/publications/eye-in-painting-with-exemplar-generative-adversarial-networks/
(705) https://openai.com/blog/image-gpt/
(706) https://news.developer.nvidia.com/ai-generates-images-of-a-finished-meal-using-only-a-written-recipe/
(707) https://www.wochit.com/
(708) https://www.wibbitz.com/

videos, the platforms rely on a combination of Natural Language Processing and neural networks to analyse texts first by keywords. The software then selects contextually suitable clips from keyworded video databases such as Getty Images, Reuters, current agency videos or its own databases, which are then used for the automated creation of the video. Some platforms also offer integration possibilities of content management systems as well as interfaces to social media channels for publishing the created videos in different formats[709].

The degree of automation in the creation process is variable: both small and extensive editing, the insertion of additional original material or even completely automated creation are possible. In this way, media organisations can prepare articles or other texts as video quickly and for various platforms.

The use of automatically generated videos is already relatively widespread in many media houses and news agencies. According to Wochit, customers include AP, PA Press Association and AFP, while Wibbitz names Reuters, Bloomberg, Forbes and The Weather Channel, among others[710]. In France, Le Figaro and Le Parisien are named as Wibbitz customers. In Germany, for example, Focus Online itself shows that videos are produced with the help of Wochit, and Bauer Verlagsgruppe and Pro7 and Conde Nast are also named as customers of Wochit.

Reuters uses a partnership with the Wibbitz platform to automate the creation of news videos that are then offered by Reuters to agency customers as material that is ready for publication. The videos, which are automatically created with the help of predictive analytics, are available within minutes of an event as video summaries in various languages. Reuters is currently also using the software to produce summaries of football matches[711], for example.

At the beginning of 2019, the Chinese news agency Xinhua also presented its first (of now three) AI-based digital news anchor, modelled on a living news presenter, who continuously presents the latest news in Chinese or English[712].

Audio content

The automated creation of audio content is also interesting for media companies. With the help of algorithms it is possible, similar to automated video production, to create audio clips, e.g. as a news story or for sports information.

Based on keywords in an article, algorithm-based summaries of news and other reports are created and offered as audio clips via Alexa or Google Assistant. Among other things, IBM offers a service through its Watson platform called "Text-to-Speech", which makes it possible to convert text written using AI into human-sounding audio files in various languages and voices[713]. Amazon also offers this service under the project name "Polly"[714]. With its product "VoCo", which has so far been presented as a prototype, Adobe is able to have written text read aloud and edited. The software can be trained with its own voice recordings and is then able to imitate the voice of the speaker from newly written texts[715].

Distributing technologies

Chatbots

Another potential field of application of AI with journalistic reference are so-called Conversational User Interfaces (CUI). These include chat bots and voice-controlled digital assistants such as Siri, Amazon Alexa and the Google Assistant. CUIs allow a natural, intuitive human-machine interaction, whereby the software is able to receive, understand and process questions and statements asked in natural language and to react with an appropriate answer[716]. A distinction must be made between chat and voicebots. While chat bots interact with the user via a text-based dialogue system, voicebots use human speech as a communication channel. According to experts,

(709) https://www.wochit.com/products/studio/
(710) Herrman, J. (2016). „As Online Video Surges, Publishers Turn to Automation", online: https://www.ny-times.com/2016/07/11/business/media/as-online-video-surges-publishers-turn-to-automation.html
(711) https://www.thomsonreuters.com/en/press-releases/2016/september/reuters-partners-with-wibbitz-to-create-videos-using-automation-technology.html
(712) https://www.youtube.com/watch?v=GAfiATTQufk
(713) https://www.pubnub.com/blog/ibm-watson-text-to-speech/
(714) https://aws.amazon.com/de/polly/what-is-text-to-speech/
(715) https://www.zdnet.com/article/adobes-voco-voice-project-now-you-really-can-put-words-in-someone-elses-mouth/
(716) Goldhammer et al. (2019). Künstliche Intelligenz, Medien und Öffentlichkeit, Berlin/Bern

voicebots in particular will continue to gain in importance in the future and could represent a paradigm shift in the way people receive information or messages - and thus also affect the field of algorithmic journalism[717].

The British Guardian newspaper has been pioneering this technology with the introduction of a chatbot that communicates with users through Facebook Messenger, by sending them a news briefing every morning with the top news stories[718]. The German Bild-Zeitung wants to use the bots even far away from Facebook Messenger. High priority is currently being given to the development of internal programmes that provide support in everyday editorial work. For example, 'slack bots' make the staff aware if apps receive a bad rating. Another sounds the alarm when a story is subject to widespread sharing in a particular local editorial office. Then the other offices can consider whether it is worth distributing the story further[719].

Programmatic Advertising

"Programmatic Advertising" refers to the automatically-controlled presentation of advertising messages to specific users. AI-based, it is possible to automatically display personalised advertising that is oriented to the special characteristics and preferences of certain persons[720].

This approach is being taken further. The applications for the automated creation of texts, photos and videos described above are also used for the so-called "Programmatic Creation" - the automated creation of advertising content. Programmatic Creation uses socio-demographic data, exogenous factors (such as weather and time) and behavioural data to automatically generate personalised and as relevant as possible ad content for a user. This form of automated advertising is already used by Google, YouTube and Facebook[721], among others. For publishers, Programmatic Creation represents a way of minimising wastage in advertising and thus offering higher quality advertising space in online offerings.

The market for so-called "branded content", i.e. media content produced by companies for the general public and/or their customers, is showing strong growth. Advertising companies are competing here with the classic media houses, and in some cases they are also using their expertise or buying their journalistic services. Especially in the online sector, a competition for attention and reach has developed, which is also based on quantity. Companies with ambitions in the area of branded content use the same instruments for automation and mass personalisation of content as classic media providers[722].

(717) Batish R. (2018). Voicebot and Chatbot Design: Flexible conversational interfaces with Amazon Alexa, Google Home, and Facebook Messenger

(718) Good, N., Wilk, C. (2016). In troducing the Guardian Chatbot. Inside the Guardian blog. https://www.theguardian.com/help/insideguardian/2016/nov/ 07/introducing the guardian chatbot

(719) https://t3n.de/magazin/chatbots-hype-smart-kleinen-programme-wirklich-chatbots-243315/2/

(720) https://www.sn.at/online-marketing/programmatic-advertising-80253256

(721) Prien, T. (2017). Automatisierte Werbekommunikation bringt Paradigmenwechsel im Marketing, online: https://t3n.de/news/automatisierte-werbekommunikation-888766/; Petereit, D. (2017). Conversational Interfaces: Das Interface der Zukunft erobert das Web, online: http://t3n.de/news/conversational-interfaces-812724

(722) https://www.bcg.com/de-at/publications/2015/media-entertainment-branded-content-growth-for-marketers-and-media-companies.aspx

ANNEX IB – Scoping: Literature and information sources

Literature

Title of the work	Year of publication	Full reference	Author/publisher
OECD Digital Economy Outlook 2017	2017	OECD (2017), OECD Digital Economy Outlook 2017, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264276284-en	OECD
Growing the AI Industry in the UK	2017	Hall, W. and J. Pesenti (2017), "Growing the AI Industry in the UK", independent report, Government of the United Kingdom, London, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf.	Hall, W. and J. Pesenti
Artificial Intelligence: A Modern Approach	2016	Russell, S., and P. Norvig (2016), "Artificial Intelligence: A Modern Approach. Global Edition", Pearson Education Limited, Harlow, England.	Russell, S., and P. Norvig, Pearson Education Limited, Harlow, England.
When Reporters Get Hands-on with Robo-Writing	2017	Thurman, Neil; Dörr, Konstantin; Kunert, Jessica (2017). "When Reporters Get Hands-on with Robo-Writing". Digital Journalism. 0 (10): 1240–1259. doi:10.1080/21670811.2017.1289819. ISSN 2167-0811	Thurman, Neil; Dörr, Konstantin; Kunert, Jessica
Fully Automatic Journalism: We Need to Talk About Nonfake News Generation	2019	Belz, Anya (2019), Fully Automatic Journalism: We Need to Talk About Nonfake News Generation, University of Brighton,	Belz, Anya
Automated Journalism 2.0: Event-driven narratives	2017	Caswell, David; Dörr, Konstantin (2017). "Automated Journalism 2.0: Event-driven narratives". Journalism Practice. 0 (4): 477–496. doi:10.1080/17512786.2017.1320773. ISSN 1751-2786.	Caswell, David; Dörr, Konstantin, Journalism Practice
Guide to Automated Journalism	2016	Graefe, Andreas (2016). Guide to Automated Journalism. New York City: Columbia Journalism Review.	Graefe, Andreas, Columbia Journalism Review
Mapping the field of Algorithmic Journalism	2016	Dörr, Konstantin Nicholas (2016). "Mapping the field of Algorithmic Journalism". Digital Journalism. 4 (6): 700–722. doi:10.1080/21670811.2015.1096748. ISSN 2167-0811	Dörr, Konstantin Nicholas, Digital Journalism
The Associated Press will use automated writing to cover the minor leagues	2016	Mullin, Benjamin (2016). "The Associated Press will use automated writing to cover the minor leagues". The Poynter Institute.	Mullin, Benjamin
Media Law and Copyright Implications of Automated Journalism	2014	Weeks, Lin (2014) Media Law and Copyright Implications of Automated Journalism, Jipel No. 4 - Vol. 1, http://jipel.law.nyu.edu/vol-4-no-1-3-weeks/	Weeks, Lin, Jipel
The Next Newsroom: Unlocking the Power of AI for Public Service Journalism	2019	European Broadcasting Union (2019), The Next Newsroom: Unlocking the Power of AI for Public Service Journalism	European Broadcasting Union

Title of the work	Year of publication	Full reference	Author/publisher
News Automation – The rewards, risks and realities of 'machine journalism'	2019	WAN-IFRA - World Association of News Publishers, (2019), News Automation – The rewards, risks and realities of 'machine journalism'	WAN-IFRA - World Association of News Publishers
Artificial Intelligence: Practice and Implications for Journalism	2017		Hansen, Mark; Roca-Sales, Meritxell; Keegan, Jonathan M.; King, George
New powers, new responsibilities A global survey of journalism and artificial intelligence	2019	Becket, Ch. (2019), New powers, new responsibilities A global survey of journalism and artificial intelligence, LSE	Charlie Beckett
Can an Algorithm Write a Better News Story Than a Human Reporter?	2012	Steven Levy, "Can an Algorithm Write a Better News Story Than a Human Reporter?" Wired, 24 April 2012, http://www.wired.com/2012/04/can-an-algorithm-write-a-better-news-story-than-a-human-reporter/.	Steven Levy, Wired
Putting Europe's Robots on the Map: Automated journalism in news agencies	2018	Fanta, A. (2018), Putting Europe's Robots on the Map: Automated journalism in news agencies	Fanta, A., Reuters Institute Fellowship Paper, University of Oxford
Estimation of clinical trial success rates and related parameters	2019	Estimation of clinical trial success rates and related parameters Chi Heem Wong, Kien Wei Siah, Andrew W Lo Biostatistics, Volume 20, Issue 2, April 2019, Pages 273–286, https://doi.org/10.1093/biostatistics/kxx069	Chi Heem Wong, Kien Wei Siah, Andrew W Lo
Tackling AI Impact on Drug Patenting	2018	King, R.D. (2018), "Tackling AI Impact on Drug Patenting", Nature, Vol. 560, correspondence, 16 August, p. 307, Springer Nature.	King, R.D., Nature
Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory	2018	Klucznik, T. et al. (2018), "Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory", Chem, Vol. 4/3, pp. 522-532, Elsevier, Amsterdam, https://doi.org/10.1016/j.chempr.2018.02.002.	Klucznik, T. et al., Chem,
Automating drug discovery	2017	Schneider, G, (2017), "Automating drug discovery", Nature Reviews Drug Discovery, Vol. 17, pp. 97-113, Springer Nature, http://doi.org/10.1038/nrd.2017.232	Schneider, G, Drug Discovery
Cheaper Faster Drug Development Validated by the Repositioning of Drugs Against Neglected Tropical Diseases	2015	Williams, K. et al. (2015), "Cheaper Faster Drug Development Validated by the Repositioning of Drugs Against Neglected Tropical Diseases", Journal of the Royal Society Interface, Vol. 12/104, The Royal Society Publishing, London, https://doi.org/10.1098/rsif.2014.1289.	Williams, K. et al., Journal of the Royal Society Interface
Machine learning at the energy and intensity frontiers of particle physics	2018	Alexander, R et al. (2018), "Machine learning at the energy and intensity frontiers of particle physics", Nature, Vol. 560, pp. 41-48, Springer Nature, https://www.nature.com/articles/s41586-018-0361-2.	Alexander, R et al., Nature
Machine learning for molecular and materials science	2018	Butler, K.T. et al. (2018), "Machine learning for molecular and materials science", Nature, Vol. 559, pp. 547-555, Springer Nature.	Butler, K.T. et al., Nature
Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data	2018	George, D. and E.A. Huerta (2018), "Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data", Physics Letters B, Vol. 778, pp. 64-70, https://doi.org/10.1016/j.physletb.2017.12.053.	George, D. and E.A. Huerta, Physics Letters B
Constrained Bayesian Optimisation for Automatic Chemical Design	2017	Griffiths, R-R. and J.M. Hernández-Lobato (2017), "Constrained Bayesian Optimisation for Automatic Chemical Design", arxiv.org, Cornell University, Ithaca, NY, arXiv:1709.05501v5.	Griffiths, R-R. and J.M. Hernández-Lobato
Planning chemical syntheses with deep neural networks and symbolic AI	2018	Segler, M.H.S., M. Preuss and M.P. Waller (2018), "Planning chemical syntheses with deep neural networks and symbolic AI", Nature, Springer Nature, http://doi.org/10.1038/nature25978.	Segler, M.H.S., M. Preuss and M.P. Waller, Nature

Title of the work	Year of publication	Full reference	Author/publisher
A Framework for Assessing the Performance of Pulsar Search Pipelines", Monthly Notices of the Royal Astronomical Society	2016	Van Heerden, E., A. Karastergiou and S. Roberts (2016), "A Framework for Assessing the Performance of Pulsar Search Pipelines", Monthly Notices of the Royal Astronomical Society, Vol. 467/2, pp. 1661-1677, Oxford University Press, Oxford, https://doi.org/10.1093/mnras/stw3068.	Van Heerden, E., A. Karastergiou and S. Roberts, Monthly Notices of the Royal Astronomical Society
Artificial Intelligence to Win the Nobel Prize and Beyond: Creating the Engine for Scientific Discovery	2016	Kitano, H. (2016), "Artificial Intelligence to Win the Nobel Prize and Beyond: Creating the Engine for Scientific Discovery", AI Magazine, Vol. 37/1, Spring 2016, pp. 39-49, Association for the Advancement of Artificial Intelligence, Palo Alto, CA, https://doi.org/10.1609/aimag.v37i1.2642.	Kitano, H., Ai Magazine
Artificial Intelligence and Machine Learning in Science	2018	Ross D. King and Stephen Roberts (2018), Artificial Intelligence and Machine Learning in Science, OECD	Ross D. King and Stephen Roberts, OECD

Other Information Sources

AI: General

https://www.artificial-intelligence.blog/

https://machinelearningmastery.com/blog/

https://aitopics.org/search

https://intelligence.org/blog/

https://www.sciencedaily.com/news/computers_math/artificial_intelligence/

https://www.aitrends.com/

AI and Pharma

https://www.digitalauthority.me/resources/artificial-intelligence-pharma/

https://verdict-ai.nridigital.com/issue-one/state-of-ai-artificial-intelligence-business-data

https://blogs.sciencemag.org/pipeline/archives/2020/01/31/another-ai-generated-drug

https://www.chemistryworld.com/news/patent-busting-ai-tool-navigates-around-protected-drug-pathways/3010015.article

https://www.managingip.com/article/b1kbptdv5sqw1y/life-sciences-firms-reveal-how-ai-could-disrupt-patent-strategies

https://www.nytimes.com/2019/02/05/technology/artificial-intelligence-drug-research-deepmind.html

https://hackaday.com/2019/01/30/ai-patent-trolls-now-on-the-job-for-drug-companies/

AI and Journalism

https://bbcnewslabs.co.uk/projects/juicer/

https://medium.com/fertiges-unfertiges/notes-on-ai-in-journalism-2fa8e9000964

https://emerj.com/ai-sector-overviews/automated-journalism-applications/

https://journalismai.com/

https://blogs.lse.ac.uk/polis/2019/10/25/journalism-and-artificial-intelligence-a-bibliography/

https://robotwritersai.com/2019/09/30/ai-journalism-a-second-chance-for-news-media/

https://patwhite.com/artificial-intelligence-and-its-impact-journalism

AI and Science

https://ai.googleblog.com/2020/01/using-machine-learning-to-nowcast.html

https://eos.org/opinions/artificial-intelligence-may-be-key-to-better-weather-forecasts

https://www.futurity.org/weather-forecasts-artificial-intelligence-2141972/

https://blogs.microsoft.com/ai/ai-subseasonal-weather-forecast/

https://amadeus.com/en/insights/blog/artificial-intelligence-weather-forecast-tool

https://www.ingenieur.de/technik/fachbereiche/ittk/kuenstliche-intelligenz-verbessert-wettervorhersagen/

https://futurezone.at/science/jeder-glaubt-wir-verhauen-staendig-die-wettervorhersage/400709511

https://scilogs.spektrum.de/via-data/w-wie-wetter-oder-kuenstliche-intelligenz-am-himmel/

ANNEX IIA – Scoping interviews: Interviewees

AI: General

- Prof. Dr.-Ing. Philipp Slusallek, Scientific Director at the German Research Center for Artificial Intelligence (DFKI)

AI– Science

- Dr. Lynn Kaack, co-chair of Climate Change AI, postdoctoral researcher at Energy Politics Group at the ETH Zürich
- Günther Tschabuschnig, CIO at the Zentralanstalt für Meteorologie und Geodynamik in Vienna

AI - Pharma

- Stefan Geissler, Entrepreneur/Co-Founder at Kairntech
- Ray Barlow, Former Chief Executive Officer at E-Therapeutics PLC

AI - Journalism

- Johannes Sommer, CEO at Restresco GmbH, Berlin
- Hendrik Alscher, CEO at text-on GmbH, Meckenheim

ANNEX IIB – Scoping interviews: Topics to be investigated

Thematic Area	Topic
AI, general	Strong AI
	Weak AI
	Deep learning
	Evolutionary algorithms
	Machine learning
	Natural language processing
	Neural Networks
Pharmaceuticals	The role of AI in developing new drugs
	AI and rare diseases
	AI and drug adherence
	AI to make sense of clinical data
	AI and finding the correct patients for clinical trials
Science	Numerical Weather Prediction
	Big Data
	data assimilation systems
	Accuracy and reliability of weather forecasting
	Parameterisation Problems
	Application markets (Energy, transport, Farming)
Journalism	Ad targeting
	Deepfake
	Fake News
	New journalistic products
	Natural language generation
	Ethical challenges of automated journalism

ANNEX IIC – Scoping interviews: Interview guidelines

- General background on the interviewee, including direct and indirect involvement of AI developments;
- General sketch of AI applications and algorithms in the area;
  - What is the current state of the art of AI applications in the field of Science/Pharmaceuticals/Media?
  - What are main barriers and drivers for the application of AI and algorithms?
- Future perspectives;
  - What of developments in AI that could be expected in 5 years from now in the field of Science/Pharmaceuticals/Media?
  - What main barriers and drivers for the future development of AI and algorithms?
- Perspectives on creation of new work, inventions or products by AI;
  - In which particular application fields do you see the biggest potential for inventions created with the help of AI and algorithms?
  - Which applications fields for AI in your domain are irrelevant for IPR created with the help of AI?
- Examples and cases;
  - Could you give us examples of AI applications that could be relevant to IPR issues?
  - Are there good practice-cases in your business environment you would like to tell us?
- References, articles.
  - What articles would you recommend to delve deeper into the issue?
  - What interview partners would you recommend to delve deeper into the issue?

ANNEX III – Case studies: Interviewees

- **Manfred Spatzierer**, CEO of UBIMET, Vienna
- **Quentin Perron**, CEO of IKTOS, Paris
- **Yann Gaston-Mathé,** CTO of IKTOS, Paris
- **Alexander Siebert**, CEO of Retresco, Berlin
- **Parjeet Tawana,** IP Counsel at BenevolentAI, London
- **Nikki Robas,** Director, Drug Discovery at BenevolentAI, London

ANNEX IVA – Final Experts Workshop Report

Trends and Developments in Artificial Intelligence - Challenges to the IPR Framework

FINAL WORKSHOP

10th September 2020, 14:30 – 17:30 CET

By Go To Webinar

Introduction and Welcome

There was a brief introduction (Dr Jacqueline Allan, JIIP) to JIIP (the Joint Institute for Innovation Policy), the Study Team (JIIP and IViR (the University of Amsterdam)) and the Study itself. This was followed by a welcome by the EC Project Officer (Ms Yordanka Ivanova) to the participants. There were over thirty participants in the workshop (including industry, patent and copyright experts (from institutes, companies, national governments and EU institutions), the European Parliament and European Commission, and the Study Team). The agenda for the workshop is in the Annex.

The moderators for the workshop were Dr Christian Hartmann (JIIP-Joanneum) and Dr João Quintais (IViR, University of Amsterdam).

Presentation of the state-of-the-art (Dr Christian Hartmann, JIIP-Joanneum)

AI has had many ups and downs in popularity since its inception in the 1950's, including the so-called 'AI winter'. The current state-of-the-art is artificial narrow intelligence (ANI) or "applied" AI, that is designed to accomplish a specific problem-solving or reasoning task. The most advanced AI systems available today, such as the IBM Watson or Google's AlphaGo, are still "narrow". Currently several successful applications are increasing the interest in AI, based on Machine Learning (ML) in various forms. The main drivers of this trend are dramatic increases in computer power, data availability, storage and transfer.

Artificial narrow intelligence is often contrasted to a (hypothetical) artificial general intelligence (AGI), in which autonomous machines would become capable of general intelligent action, like a human being, including generalising and abstracting learning across different cognitive functions. Projections from the few computer scientists active in AGI research on the time frame for the realisation of AGI range from a decade to a century or more.

AI as a service (AIaaS) is a new business trend that enables everyone to use artificial intelligence regardless of their knowledge and is evident in all three of the application areas considered in this study - pharmaceutical research, science (meteorology).

AI used in **pharmaceutical research** is mostly machine learning (supervised, unsupervised and reinforcement learning) which can produce disease diagnoses, predictions of drug efficacy and characteristics (e.g. toxicity, absorption, excretion). AI can help cutting costs in drug development by saving time and resources. AI applications in pharmaceutical research include:

- Identifying molecular drug targets (e.g. proteins, nucleic acids) by searching through libraries of candidates;
- Accelerating the high throughput screening needed to find a candidate substance for further investigation in drug development;
- Repurposing of drugs to meet new or different need;
- Polypharmacology (where a disease is due to multiple complex and networked malfunctions of the body); and
- Finding vaccines and speeding up their development (helping in both gene sequencing and simulations of vaccines).

In most processes, some human intervention is required, whether only at the start-up phase on inputting parameters to search engines or defining datasets, or throughout the process to act as a feedback loop in order to optimise the process at each stage.

Meteorology is an area in which machine learning technologies are already used as standard for tasks that are largely automated and, once coded and implemented, run without further human intervention. Applications of ML in meteorology include post-processing of weather data; bias correction of meteorological observations; parameterisation of models to correct for radiation, turbulence, cloud microphysics, etc.; and data assimilation.

AI in **journalism** is used for software-supported or automated aggregation, production and distribution of content of any kind (data, text, images, audio or video) in the so-called 'algorithmic journalism'. Currently the most developed area is automated content creation, where applications include: writing product descriptions, preparing patient summaries in hospitals, reporting on sports events and share prices and local information on property markets.

Companies working on AI for journalism continuously develop their systems to meet customer needs and also work with customer-specific models that reflect the characteristics of each customer's domain. Know-how is commonly protected through licensing models, rather than by applying for intellectual property protection, e.g. through patents.

So far, automation in journalism seems still quite limited. In a recent survey, journalists voiced concern about, among other things, the inability of current news robots to interrogate data. They noted that current automation is reliant on single, isolated data streams and is one-dimensional due to the quantitative data feeds on which it relies. The development of automated journalism will depend on advances in the wider field of AI. For example, self-learning software could eventually lead to breakthroughs in data mining and make it much easier to work with unstructured data. User feedback could help to improve Natural Language Generation and enhance the linguistic diversity of stories. Meanwhile, technologies such as speech-to-text and face recognition from images can vastly improve the workflow speed of journalists and add value to their existing archives by making them machine-searchable in new ways.

In the **question and answer session**, it was also noted that AI is being used in pharmaceutical clinical trials to help in identifying a relevant patient population from patient genetic profiles. This is a large task with, for example, 30,000 patients required for clinical trials related to Covid-19. In addition, AI can be used to review medical histories which can include information that will result in the exclusion of some patients from tests on certain molecules. AI algorithms can also individualise new compounds (e.g. for antibiotics) but the 'AI-constructed' molecules also need to be tested.

So far only one pharmaceutical has been put forward for Covid-19. In the two pharmaceutical-related company case studies, it was noted that IKTOS is involved in identifying compounds, while Benevolent AI (with over 400 employees) wants to be a stand-alone company that only collaborates with big pharms when clinical trials start.

One of the most significant concerns of big pharma is that monopolistic patent rights may be lost if there is a heavy reliance on big data and thereby on third parties. This may lead to consolidation between big pharma and big tech with them even taking large shareholdings in one another. It was also noted that already a few big pharma companies are joining together in a shared data space to train algorithms.

Presentation of the legal review – Copyright law (Professor Bernt Hugenholtz, IViR)

The Study Team identified four (interrelated) requirements for the copyright protection of "works". These must be met by the AI-assisted output in order for it to qualify as a work. The four requirements (and the ways in which AI qualifies for each) are shown below:

1. Production in the "literary, scientific or artistic domain": AI qualifies as much AI-assisted output stays within traditional copyright domain.

2. Human intellectual effort: AI qualifies because there is currently no such thing as a completely autonomous AI creator, some human effort always being required.

3. Originality/creativity: AI qualifies where a human creator(s) makes a creative choice or choices.

4. Expression: AI qualifies if creative choices are "expressed" in the AI-assisted output (as a text or sound, for example), and output remains within scope of general authorial intent.

Expanding further on the topic of creativity and the role of the human creator in AI production, there are three phases of creativity:

- Conception: plan/design/specifications (e.g. choice of genre, technique, materials, AI system, training data) – here there is a large role for a human creator;

- Execution: where the first draft version of the work is created (e.g. writing, painting, composing, recording, coding) – here there is a limited role for a human creator;

- Redaction: finalization (e.g. editing, formatting, framing, cropping, selection (e.g. where an AI machine produces many version of a work and the human selects just one from those), other "post-production") – here there is a variable role for a human creator.

Human intervention is only needed in one of these phases in order for creativity through human intervention to be shown.

Authorship of AI-assisted outputs is relevant only if the AI output is a "work". In this area, there is broad discretion for Member States, some national rules possibly being incompatible with EU principles. There are no harmonized EU rules on joint authorship, co-authorship or collective works except for some provisions in the EU Software & Database Directives. The author of AI output is the person(s) that engaged in creative choices, even if s/he did not execute the work her/himself, such as in the case of the user of the AI system. The AI developer will qualify as a co-author of the output only in the case of a concerted creative effort, not in the case of general-purpose AI. False claims of authorship may arise (and perhaps increasingly so in the future) when a name is attached to a work (e.g. music) without human involvement in its production (where copyright infringement may be claimed and there is no legal requirement to reveal creative process).

"Authorless" AI output is still protected in the EU by related rights if it is in audio/video/database form. Related rights for AI in the EU include those for phonogram producers (recorded audio), broadcasters (transmitted signals), film producers (recorded video), database producers (aggregated and structured data) and press publishers (press publications). In these, the rights are directly attributed to the entrepreneur (usually a legal person) with no human authorship/effort requirement. There is also no threshold requirement (e.g. of time or money), except for database rights (where "substantial investment" is required).

Specific examples of authorship were considered for journalism and weather forecasting. In these, there is a limited creative role for humans at the conception stage, AI is dominant during the execution phase, and the redaction stage may or may not require creative choices. The output is therefore unlikely to qualify as "work", but this will be case-dependent. For example, a simple news report is a work of fact and is not normally copyright protected, whether produced by humans or AI, but there may be protection available under database rights.

In the Arts, there is a creative role for humans at all three stages and the outputs are likely to qualify as "work" and to be suitable for copyright protection.

The presenter concluded by noting that the "work" status of AI outputs depends on whether creative choices are being made. Authorship follows work and co-authorship issues may arise for AI developers/programmers only in special cases. Related rights may offer protection to "authorless" AI outputs in distinct cases (audio, video). Albeit outside the remit of this Study, it was noted that, in general, the need for new rights related to intellectual property (e.g. for "authorless" AI output) should be based on economic evidence and so far there are no indications of (pressing) industry needs.

In the **question and answer session**, there was further discussion of what should qualify as work as some research has shown that there is not always a causal link between human intervention and the output and that there are some grey areas, illustrated by cases where the human intervention in creativity may be exaggerated. It was noted that this is a general copyright problem. Many works can be described in terms of a combination or iteration of unprotected ideas and, if enough ideas are combined, there may be enough to qualify as a 'work'.

Again referring to the creation phases, advances in deep neural networks reflect that sometimes it is not known how inspiration leads to creation, but this is equally true for the human creative process. The focus of enquiry should not be on whether or not a machine acted in a 'çreative' manner, but whether creative choices by human beings were made and expressed. It was also noted that the national courts tend to apply the test of 'originality' to AI outputs in different ways, which may create problems with a lack of legal certainty. This is indeed a general problem (cf. Dutch supreme court case on the protection of daily conversation recorded and transcribed).

Regarding the three phases of creation, the use of the word 'redaction' in the report will be reviewed, in particular with reference to software.

Some clarification was sought with reference to "authorless" AI outputs remaining unprotected only in cases where no related right is available (Draft Final Report section 5.7 (p141). While there are related rights for recording of sound and video, there is a lack of symmetry in legislation where there are no equivalent rights for alphanumerical information.

It was noted that input-related issues were outside the terms of reference of this Study.

Presentation of the legal review – Patent law (Professor Daniel Gervais, IViR)

The presentation covered inventorship, ownership, novelty, inventiveness and disclosure in the context of European patent law. It was noted that the scope of the Study is limited by its terms of reference.

Inventorship issues were identified not least that the need to name the inventor is a formal EPO requirement and that a machine cannot be named as the inventor (while noting that there are appeals pending on the DABUS case). Member State laws have substantive requirements for inventorship and may vary (noted here, the Shemtov study). Related to this is the moral right of the inventor with the link between the human holder of the right and the output (again, an area most likely to be dealt with at Member State level). A machine cannot currently invent throughout the whole process without any human intervention (there is no documented case to be found to date of a purely AI-generated invention with no human causal link) but it is likely to happen in the future, although it may be many years in the future.

On **ownership**, the Study proposes a roadmap on who owns AI-assisted output. Looking for a human causal link, the development and production chain is often long with many who may claim ownership (e.g. when learning, one reads many books but should the authors of those have some ownership of subsequent creative outputs?). In machine learning, ownership can result from the human input to the definition of the question, the objectives, selection, interventions, final selection, etc.

There are many human roles to consider and an example from the case studies was presented. In the IKTOS case study (pharmaceuticals), there is human intervention in many areas including: (a) the preparation/definition of the initial dataset (for ML); (b) the setting up of specific objectives for the ML process; (c) the preselection of compounds; (d) interventions by engineers during the iterative trial and error process; and (e) final selection of the compounds. Additionally, there may be (f) programming the AI platform (if applicable) and (g) programming the AI "system".

Ownership issues are typically not resolved by EPC/EPO. In Member States, AI is not an employee or legal person and therefore does not have "successor in title" rights. There is no Member State in which AI can be claimed to have ownership or inventorship (as AI cannot claim 'personhood').

To defeat the claim of **novelty** in a patent application, a single item of the state of the art must contain the elements of a claim in the application and enable the POSITA ('person of ordinary skill in the art') to "practice the technical teaching which is the subject of the document, taking into account also the general knowledge at that time in the field to be expected" of them. There are possible changes to this: quantitative change to examination process (as AI can process more data) and qualitative change (as AI processes data differently than humans). This may lead to a "cat and mouse" situation between applicants and patent offices (where the applicant have more powerful machines to identify novelty than the offices have to detect prior art).

An invention is considered to involve an inventive step if, having regard to the state of the art, it is not obvious to the POSITA. However, the POSITA is a legal fiction, a "composite entity". The specification should show that there is a technical problem that the invention solves. AI may change the three steps (determining the "closest prior art", establishing the "objective technical problem" to be solved, considering whether or not the claimed invention, starting from the closest prior art and the objective technical problem, would have been obvious to the skilled person").

AI systems may be used to identify the relevant prior art (the dataset(s) to be accessible to these AI systems, the fields covered) and a machine may have to be told (by some means) to what fields it should limit itself. That may influence notion of "common general knowledge". It may also lead to a heightened standard being required to measure and assure 'inventiveness'.

Disclosure has to be conducted in a manner sufficiently clear and complete for it to be carried out by a POSITA, who may avail themselves of "contemporaneously available common general knowledge and the literature referred to in the description". If there is a key role for the AI in the production, it may be questioned how the AI carried out its role. This is a move away from AI as a 'black box' and towards a demand for the role of AI in production to be explainable.

One option could be to have a 'deposit' for such information, akin to the Budapest International Microorganism Deposit System which has been helpful when pursuing patents involving biological material. One issue would be what should be deposited and whether that deposit would be confidential. For example, deposits could include defined algorithms, training models, training data or other (any or all of which may be confidential). It can be done but it is not simple and it is not clear if it should be done.

The case of pharmaceuticals is one in which AI plays many roles that can augment human capabilities and accelerate certain steps (e.g. compound identification, clinical study parameters). Humans are still very much part of processes – but there may be new roles to be understood (e.g. AI programmers).

In conclusion, naming a human inventor obligation remains as formal requirement. At Member State level, there are substantive requirements but no practical conflicts are currently being seen. If and when they arise, they may lead to or require legal decisions about who is the human and who can claim the status of AI-assisted invention. Novelty and inventiveness analyses will change as more and more AI tools are used to identify and process prior art but this is rather a matter for patent offices to address. Criteria interpretation may also evolve in case law. A possible deposit for disclosure requirement could be considered

In the **question and answer session**, it was suggested that there needed to be a distinction between process claims (problem) and product (no problem) claims in the disclosure requirements paragraphs of the report. It was clarified that the aim here is not to draw firm distinction but to say that there may be more difficulty for processes than products (as a process may be a block box in its entirety). For a product, it is easier to make a claim of novelty. The text will be reviewed to clarify this point.

On the issue of there always being human input but that its relevance to inventorship is disputable, it was noted that the result is that companies cannot patent inventions where they cannot identify the human input. (One case study company for this Study, Benevolent AI, reported having to change their strategy to accommodate this.) Responding to this point, it was reminded that this is a Study about EU law and not about the laws of all Member States, nor about laws outside of the EU. The related reference to the situation being unlikely to change for decades will be reviewed.

The EPO requires the identification of a human inventor (and although sometimes a 'straw person' could be enough, there are national laws where this may not work). There are several aspects not covered in this Study, including coverage of all national laws and their application, the debate on expanding legal definitions, and the lack at international level, of a definition of an inventor (although an invention is defined under TRIPS). It was noted that the EPO has come up with inventor requirement by default (although, for example, in Cyprus and Monaco inventors do not need to be natural persons). No documented cases were found by the Study where there is no human in the process, although in time there may be such cases. The inventorship/invention language in the Study will be reviewed.

Referring to the state-of-the-art, it is not only technology that evolves but also terminology and legalities and thus the term 'inventor' will also evolve and a 'person skilled in the art' may change to be an algorithm.

Regarding sufficiency of disclosure and the deposit issue, it was observed that this is not done for traditional software patents. Any proposal should consider whether a deposit system, if needed, could be beneficial to AI alone or should be for all software, and should focus on any positive difference that would result (given that the functionality is the important aspect and is not readable from the source code). A further study on the interesting topic of deposits could be useful.

The session concluded with a question about the role of the inventors moral right in patent law. There is not abundant literature on this. As more AI is added in the future, the situation will become more complicated and is likely to become a matter for national laws and national courts. So far, such cases do not yet exist but it is expected that they will in the future.

Close of meeting

At the end of the workshop, the contributors, presenters and moderators were all thanked for their inputs. It was noted that all participants are welcome to send comments over the following few days. The comments from the workshop will influence and be integrated into report as appropriate to the terms of reference of the Study. Some may be used to highlight controversial issues that were not able to be covered in the Study that can be the basis for future work. In due course, the Study will be published by the European Commission and shared with the workshop participants and others who contributed to the Study.

The European Commission ended by highlighting some of its future studies:

- DG CNECT copyright unit will run a study on copyright and new technologies (looking at how companies use AI in the creative industries and the impact of scenarios, including economic impacts that were excluded from this Study)

- DG GROW is launching a study on the Trade Secret Directive

In addition, the Commission EC will put forward an action plan this Autumn to further improve the working of the intellectual property Framework.

The European Parliament will also shortly publish its report on AI and intellectual property.

Agenda of the workshop

JIIP
Joint Institute for Innovation Policy

Trends and Developments in Artificial Intelligence - Challenges to the IPR Framework

Final workshop – Agenda

European Commission - DG CNECT
Unit CNECT.DDG1.A.1 - Robotics and Artificial Intelligence

10th September 2020, 14:30 – 17:30 CET

By Go To Webinar

Agenda Item	Time
Introduction and welcome (EC and JIIP)	14:30 – 14:45
Presentation of the state-of-the-art (Christian Hartmann, JIIP-Joanneum) Q&A	14:45 – 15:10
Presentation of the legal review – Copyright law (Bernt Hugenholtz, IViR) Q&A (Moderator: João Quintais, IViR)	15:10 – 16:15
Presentation of the legal review – Patent law (Daniel Gervais, IViR) Q&A (Moderator: João Quintais, IViR)	16:15 – 17:20
Wrap-up – next steps	17:20 – 17:30
Close of meeting	17:30

This Study has been carried out under Service Contract No. LC-01188842 within Framework Contract SMART 2018/0052 for the European Commission DG CNECT — Directorate-General for Communications Networks, Content and Technology (Unit A.1 - Robotics and Artificial Intelligence)

Participants, AI-IPR webinar workshop, 10 September 2020

Project Officer for this Study

- Yordanka IVANOVA (DG CNECT, European Commission)

Study Team

- P. Bernt HUGENHOLTZ, IViR, University of Amsterdam
- João Pedro QUINTAIS, IViR, University of Amsterdam
- Daniel GERVAIS, IViR, University of Amsterdam
- Christian HARTMANN, Joanneum Research/JIIP
- Jacqueline ALLAN, Joint Institute for Innovation Policy (JIIP)

Invited participants

- Ryan ABBOTT, University of Surrey, Professor of Law and Health Sciences
- Mario BLACK, Intellectual Property Manager, BenevolentAI
- Dan L. BURK, Chancellor's Professor of Law, U. California, Irvine
- Sven BOSTYN, Associate Professor, JUR Centre for Advanced Studies in Biomedical Innovation Law, Faculty of Law, University of Copenhagen
- Peter BLOK, Court of Appeal The Hague and Professor of Patent Law, Utrecht University
- Remy CHAVANNES, Lawyer and Partner, Brinkhof N.V.
- Caroline COLIN (CNECT) (DG CNECT, European Commission)
- Valentin COUTOULY, European Parliament, (Stephane Sejourne's Office)
- Alica DALY, Senior Policy Officer for Artificial Intelligence and Data at WIPO
- Jean-Marc DELTORN, Snr Researcher, Centre International IP Studies (CEIPI)
- Ziga DROBNIC (DG CNECT, European Commission)
- Virginie FOSSOUL (DG GROW, European Commission)
- Stef van GOMPEL, Associate Professor, IViR, University of Amsterdam
- Reto HILTY, Director Max Planck Institute for Innovation and Competition, Munich, and Professor U. Zurich
- Maria IGLESIAS (JRC, European Commission)
- Peter MEZEI, Associate Professor, Faculty of Law, University of Szeged, Hungary
- Krzysztof NICHCZYNSKI (DG CNECT, European Commission)
- Begona Gonzalez OTERO, Max Planck Institute for Innovation and Competition, Munich, Senior Research Fellow
- Heli PIHLAJAMAA, Director of the Patent Law Directorate in the EPO

- Markus RIECK, Patent Attorney, Pharmacist, FUCHS IP, Team member of the Artificial Inventor Project
- Matthias SCHMIDT, German Ministry of Justice and Consumer Protection
- Vytenis SEMETA (DG GROW, European Commission)
- Sheron SHAMUILIA (JRC, European Commission)
- Noam SHEMTOV, Reader in IP and technology Law, Queen Mary, University of London
- Alain STROWEL, Professor at the Saint-Louis University & UC Louvain (Belgium)
- Tatiana SYNODINOU, Assistant professor, Law Faculty of the University of Cyprus
- Ulrike TILL, Director of the Division of Artificial Intelligence (AI) Policy, Office of the Director General, WIPO (World Intellectual Property Organisation)

ANNEX IVB – Interim Expert Workshop Report

SUMMARY REPORT 4 March 2020

The workshop consisted of in-depth clustered interviews held on March 4, 2020, carried out in a workshop-like setting. The activity was divided in two parts. Part I covered challenges posed by AI to EU copyright law. Part II covered challenges to European patent law. Each Part was structured around specific topics and questions, to which a proportionate amount of time was allocated. The discussion was introduced and moderated by team members, and included a brief explanation of each question and relevant issues arising therefrom. The experts were then invited to discuss each question in-depth.

The list of questions is replicated below.

The interviews started with a **preliminary question (1)** on the conceptual or analytical approach to AI outputs. In particular, experts were asked, for analytical purposes, whether it makes sense to distinguish between types of or assistance by the AI system to the created or produced output? One popular distinction is between: (a) AI-*generated* output, i.e. autonomously created or produced by the AI system; and (b) AI-*assisted* output, i.e. created or produced by an AI system with some level of human intervention.

- From the discussion it emerged that most experts considered scenario (a) non-existent, in the sense that all AI outputs have some level of human intervention in the development or application of the AI system. That is to say, experts considered that in different stages of designing, developing or using an AI system or tool that leads to an application or output potentially protected by IPR there is almost inevitably a significant degree of human intervention. It is therefore deemed not very useful to discuss AI-*generated* outputs other than as a boundary case where clearly no IPR protection should be recognised. For the remaining AI-*assisted* output cases – which cover a broad spectrum of cases– the main issue is to determine the degree of human causation, intervention or proximity in relation to the particular output, as that will determine IPR protection.

The discussion then entered Part I, on Artificial Intelligence Challenges to Copyright Law.

The initial question discussed here was about **AI outputs and the notion of "work"** (2). In particular, experts were asked:

(a) Can an AI output qualify as a protected "work"?;

(b) How can we apply the current standard of the Court of Justice of the European Union (CJEU) of "author's own intellectual creation" (originality) resulting from "free and creative choices" to AI outputs?

(c) Is it appropriate to assess originality of AI outputs by drawing a conceptual distinction between different phases of authorship? For example, one might distinguish the following phases:[723]

(i) Conception (creative plan/design/specifications);

(ii) Expression (execution); and

(iii) Redaction (presentation, "curation");

(d) What degree of human intervention (if any) is required for a finding of originality of an AI output during the conception, expression, and redaction phases?

(723) The division into these different stages is inspired by desk research and case law. See e.g. Jane C. Ginsburg & Luke A. Budiardjo, "Authors and Machines", *Berkeley Technology Law Journal*, Vol. 34 (discussing "detailed conception" and controlled execution). See also CJEU case law on orginality, notably the Court's considerations in Case C-145/10 – Painer, paras 87-94.

(e) In particular, can human intervention in the conception and redaction phases alone suffice for a finding of originality?

An important part of the discussion here was about the challenges of examining work and authorship separately. It was accepted that this may be useful for analytical purposes when asking questions, but most experts agreed that these aspects must be jointly analysed. That is to say, Questions 2 and 3 are inextricably linked.

Discussion on questions (a) through (e) were held with the well-known example of the *Next Rembrandt* project[724] as a point of reference. This project involved a deep-learning algorithm trained on data extracted from Rembrandt's oeuvre that helped produce an astonishingly 'Rembrandt-like' portrait. The experts agreed that, most likely, the portrait would qualify as a work protected under EU copyright standards, notwithstanding the important role played in the creative process by the deep learning algorithm. More generally, there was consensus that many outputs currently generated with the aid of machine learning systems would qualify as protected works, since AI outputs are usually the product of (multiple) human interventions reflecting the exercise of creative freedoms, both at the conceptual and redactional stages of the production.

Participants however raised issues with the use of the terms "conception", "expression" and "redaction" as either being to US-centric[725], too vague or in some cases unclear. For example, "conception" could easily be understood as unprotected under the idea /expression dichotomy, whereas "expression" could be considered to always be protected. The "redaction" stage on the other hand is linked to the issue of the legal relevance of selection and arrangement for content leading to its publication/availability, and how much that matters for ascribing copyright protection. It was noted that more precisely named phases with similar meanings to those above would be useful to legally test whether an AI-assisted output qualifies as a protected work. In that sense, it was suggested that the Study team use terminology closer to that employed by the CJEU in *Painer*[726].

On the remaining issues, different issues were raised on how to qualify the "free and creative choices" in this context (e.g. what choices are creative vs functional), types of output with more vs less creative input (Next Rembrandt vs robo-journalism for sporting events), and other relevant aspects for the legal analysis that the Study Team will consider.

The following question was on **Authorship of AI outputs (3),** as well as a specific question on related or *sui generis* rights. The experts were asked:

(a) If an AI output does qualify as a "work", to whom should authorship be allocated? Specific examples to be discussed but possibilities include:[727]

(i) The user of the AI system (most obviously for AI-assisted outputs);

(ii) The "designer(s)" of the AI system (an option e.g. for AI-generated outputs);

(iii) The user and designer as co-authors of a joint work (an option for AI-assisted outputs);

(b) In what scenarios are AI outputs "authorless"? What are the legal consequences?

(c) To what extent can existing related or sui generis rights provide protection for AI outputs?

(724) See 'The Next Rembrandt', *https://www.nextrembrandt.com*.

(725) See e.g. American authors using similar terminology: Bridy, Annemarie, Coding Creativity: Copyright and the Artificially Intelligent Author (July 18, 2011). Stanford Technology Law Review, Vol. 5, pp. 1-28 (Spring 2012) (discussing the notion of conception in US case law on authorship); Jane C. Ginsburg & Luke A. Budiardjo, "Authors and Machines", *Berkeley Technology Law Journal*, Vol. 34 (discussing conception and execution, as well as expression); Yanisky-Ravid, Shlomit and Velez-Hernandez, Luis Antonio, Copyrightability of Artworks Produced by Creative Robots, Driven by Artificial Intelligence Systems and the Concept of Originality: The Formality - Objective Model, Minnesota Journal of Law, Science & Technology, Forthcoming (discussing "conception" and the originality standard in US copyright law).

(726) http://curia.europa.eu/juris/document/document.jsf?docid=115785&doclang=EN

(727) These examples are partly inspired by Jane C. Ginsburg & Luke A. Budiardjo, "Authors and Machines", *Berkeley Technology Law Journal*, Vol. 34, P. 343, 2019. The authors define the designer as the "individual (or set of individuals) who endowed the machine with the training and the creative raw material requisite to the machine's generation of a "creative" output". See *id.*, p. 446.

As noted, these questions were discussed in combination with those under Question 2.

- One insight of the discussion was that most experts considered this to be predominantly a policy question. In the absence of clear legal rules, if copyright protection is recognised for AI outputs and there are no clear rules on authorship, then authorship of those outputs should be determined in a way that satisfies the policy rationale underlying copyright protection.

- Most experts found no argument to grant protection to the developers/designers of AI systems as a rule. This is because in most instances AI systems are akin to general purpose technologies. These systems are functional and used to assist certain users in creating a specific output(s). Such outputs will then likely be subject to choices on selection and arrangement leading to publication/availability. Developers/designers may also rely on protection under other legal regimes (e.g. trademark law). Only in few cases where there is a clear causation link between the AI system as designed and the concrete output as expressed should this possibility of copyright protection for the developer/designer be considered. In all other cases, we must assess the creative process/chain to understand which creative human choices are directly expressed in the output, and as a result worthy of copyright protection. For this assessment, the discussion on Question 2 about the different phases of authorship will be useful.

- On sub-question (b), the consensus is that where an AI output is truly authorless, then it cannot be subject to protection (assuming also the absence of a legal rule attributing ownership without authorship).

- On sub-question (c), there was consensus that if a certain AI-assisted output qualifies as a type of subject matter protected by related rights it may be possible that existing rules are sufficient to attribute ownership of that output to the respective related rights holder (e.g. depending on the specific fact pattern, a sound recording produced with the assistance of an AI system could be owned by the phonogram producer owning or using that AI system and responsible for the production of that phonogram).

Finally, we asked the experts some **policy questions on AI and copyright law challenges** (4). In particular:

(a) What policy arguments might justify introducing a related right for AI outputs? Are these arguments convincing in the context of international and EU copyright law and related fields?

(b) If protection under a related right were introduced for AI outputs, what type and scope of protection would be justified?

- On these questions the discussion revolved around whether we should consider this from the perspective of incentives or disincentives for human creation. In particular, the easy recognition of copyright protection for AI outputs could create a saturation of works that would place an enormous competitive pressure on human authors. Options to curtail this effect were discussed, from taxation to the creation of artificial costs for protection so as to put authors on a level playing field.

- There was consensus among experts that there is no clear economic rationale for granting copyright or related rights protection to AI outputs. There is no obvious market failure in relation to the development of AI systems or resulting outputs. In the absence of such market failure, additional recognition of protection beyond what would result from the regular application of copyright law rules appears unjustified. This would for instance speak against providing related rights protection beyond the closed catalogue currently recognised in EU law (i.e. against a solution like that which was implemented for the right in unpublished works in the Term Directive[728]).

(728) This is a reference to the regime to the publisher's right in the publication of previously unpublished works as prescribed by Article 4 of Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (codified version). A proposal in this direction is advanced by A Ramalho, Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems (June 13, 2017). Available at SSRN: https://ssrn.com/abstract=2987757 or http://dx.doi.org/10.2139/ssrn.2987757.

In Part II of our in-depth clustered interviews we discussed Artificial Intelligence Challenges to Patent Law.

The first question in this part referred to **Inventorship and ownership of AI outputs (5)**. Because it is outside the scope of this project, these questions explicitly do not consider the hypothetical scenario where AI systems have legal personhood and can therefore be considered as inventors. Experts were asked in particular:

(a) To what extent is human inventorship a requirement for patentability? (N.B. in the sense of a human creative or intelligent conception of the invention, or contribution thereto) Or is the requirement merely that a human is designated as the inventor?[729]

(b) Are inventor's (moral) rights indicative of a substantive or formal requirement of human inventorship?[730]

(c) What degree of human contribution/intervention is required for there to be human inventorship in an AI output?

(d) What is the effect of automated claim broadening in the assessment of inventorship?

(e) If human inventorship is required, either in the conception or merely as designation, to whom is inventorship of an AI output allocated?

(f) Can the approach followed for biotech inventions, where patents can be granted to non-deterministic outputs, apply to the assessment of inventive step or non-obviousness of AI outputs?[731]

(g) If the requirement of inventorship is met, to whom does the right over a European patent on an AI output belong?

- The consensus among experts was that under the EPC human inventorship is a purely formal requirement for patentability. In other words, pursuant to recent EPO decisions (discussed by the experts) it is sufficient to have a human inventor designated on the patent application. The main reason is that the notion of inventorship is not harmonised. It was difficult to find a common denominator when drafting the EPC. Therefore, the issue of inventorship is one to be ultimately decided by national law.

- Most experts were sceptical that an AI system can be considered the inventor, as that would require such a system to itself define the problem, in addition to identifying the solution. In any event, even if it were the case that an AI system could be an inventor (in the sense of a purely AI-*generated* output), the EPC appears to advance only a formal requirement of inventorship – the designation of the human inventor. Provided that requirement is met, it does not seem to matter under the EPC whether the invention at issue is AI-generated or AI-assisted.

- In this context, there was some discussion on the potential applicability to AI outputs of the rules and legal fictions for biotech inventions. The particular example of hybridomas technology was considered a good analogy to AI-*assisted* inventions.

- Most experts considered there is an important difference between the European patent system and the European copyright system. Namely, in the latter personality/human authorship considerations have a more central role, whereas they are ancillary to European patent law. This leaves greater margin to consider an AI system as an inventor, or at least to accommodate a lesser degree of human intervention/contribution/causation to an AI-assisted output.

- In light of the above, most experts were not ready to conclude that inventor's (moral) rights in the EPC are indicative of a substantive requirement of human inventorship (sub-question b).

(729) See Article 81 and Rule 19 EPC, as well as Article 90 EPC.
(730) See Articles 58 to 62 EPC.
(731) See also Question 6(d).

- As regards the degree of human contribution/intervention required for there to be human inventorship in an AI output (sub-question c), most experts noted that this will be a matter for national law to decide (e.g. German and UK case law on co-inventorship point as decisive factors the setting the problem and deciding which tools to use, whereas in France the situation is much less clear). However, it was also noted that what matters is whether or not the result at issue is non-obvious; if it is, the AI output will be patentable. This illustrates yet another difference between copyright and patent law: in patent law, invention and inventor are not as closely linked as work and authorship are in copyright law.

- The experts considered sub-question d) of lesser relevance, as it basically relates to an abstraction of the exercise of claiming, not making new inventions.

- As for sub-question e), two key points were made by the experts. First, it is up to national law to govern this question in most cases; it is not a matter that can be solved by the EPC. That being said, it was also noted that it is necessary to make a distinction here between invention and inventive step, and to identify the different context in which the AI system was used when developing the output/invention. In certain configurations, both the developer of the AI tool and its user can be co-inventors. But, for instance, if the tool at issue is generally available and not a pure black box, then it is part of the state of the art. In all cases, it must be assessed whether the AI system/tool uses is sufficient or necessary cause for the output/invention.

- Regarding sub-question f), the consensus was that the approach followed for biotech inventions, where patents can be granted to "non-deterministic" outputs[732], can apply to the assessment of inventive step or non-obviousness of AI outputs. It was noted that the term "non-deterministic" is imprecise here, and it should be clarified what is meant in this context, i.e. black-box AI systems in relation to which it is difficult or impossible for humans to explain with precision the causal link between specific inputs and outputs.

- Finally, on sub-question g), the answer under the EPC is that the "applicant" will be the owner of the right over a European patent on an AI output. Problems between applicant and inventor should be addressed by national law. There are not sufficient cases on entitlement to invention to indicate that this is a significant enough problem to merit further harmonisation. Importantly, there was no consensus that this in an AI specific problem; it might become more prominent of an issue in the context of AI outputs, but there is not enough data on this.

The subsequent questions focused on **Inventive Step and Industrial Applicability of AI outputs (6)**. Experts were asked the following questions:

(a) How should the requirement of inventive step and industrial applicability be interpreted in relation to AI outputs? Do they apply in the same way as to human-made inventions? If not, what are the differences?

(b) How should the concept of "person having ordinary skill in the art" (POSITA) be interpreted in the assessment of AI outputs? Can the concept of POSITA is this context include other AI Systems?

(c) How should prior art be defined in the assessment of AI outputs? What are the relevant fields of technology? Can the relevant prior art include other AI outputs?

(d) Can the approach followed for biotech inventions, where patents can be granted to non-deterministic outputs, apply to the assessment of inventive step or non-obviousness to AI outputs?[733]

- There was debate on application of the inventive step to AI outputs. Most experts argued that for AI-assisted inventions the examination would not imply significant changes to the current problem-solution approach followed by the EPO. You must identify and assess the state of the art and the claimed invention, and then

(732) On the notion of "non-deterministic" see Iglesias Portela, M., Shamuilia, S. and Anderberg, A., Intellectual Property and Artificial Intelligence - A literature review, EUR 30017 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-14178-5 (online), doi:10.2760/2517 (online), JRC119102., p.8 ("another issue might arise from the fact that 'deep learning technologies are non-deterministic: they involve some randomised initialisation. Therefore, even the same training data and the same neural network architecture might lead to slightly different performance of machine learning'."), referring to WIPO, Standing Committee on the Law of Patents, Thirtieth Session Geneva, June 24 to 27, 2019, 'BACKGROUND DOCUMENT ON PATENTS AND EMERGING TECHNOLOGIES', available at: https://www.wipo.int/edocs/mdocs/scp/en/scp_30/scp_30_5.pdf.

(733) See also Question 5(f).

ask the key question: would (not "could") the POSITA realise this invention? This implies a necessary link between prior art and invention.

- It was noted that AI systems link inputs and outputs. The inventive step looks at the necessary link between input and output. Just noting that an AI system could arrive at a particular solution is not sufficient to reject the invention for being obvious. What matters is whether the AI would actually reach that particular invention. In the context of AI, solutions emerge by minimising a loss function (optimisation) to reach a space of possibilities. It is a random process from the start point to the end point, making it difficult to establish a causal link between prior art and claimed invention that leads to an assessment of obviousness.

- Although the POSITA is a composite entity, it was also clear from the discussion that the inventive step must be subject solely to human assessment. AI systems may be used to identify the relevant prior art – and the EPO is doing this – but the assessment of the inventive step is made by a human. The result is that beyond a certain level of complexity of the claimed AI-assisted invention, EPO examiners will have difficulty in establishing the causal link required to a finding of obviousness. It will therefore be rare for applications for AI-assisted outputs/inventions to be rejected on the grounds that they fail the inventive step requirement. Most likely, opposition proceedings will become the backstop for the system to reject patents on grounds of obviousness.

- On industrial applicability, the consensus is that AI-assisted outputs do not pose a new and unique problems for the patent system at this stage.

The following questions were about **Sufficiency of Disclosure and AI outputs (7).** According to Article 83 EPC, an application must disclose the invention in a manner sufficiently clear and complete for it to be carried out by a POSITA. In this context, experts were asked the following:

(a) How should the disclosure requirement be examined in the case of AI outputs?[734]

(b) For AI outputs, is it suitable to assess sufficiency of disclosure based on concepts of reproducibility and plausibility of the claimed invention?

(c) Is a system for the deposit for algorithms used in the development or generation of the AI output – similar to the existing system of deposit of biological material – suitable in this context?[735]

- Most experts considered that AI-assisted product patents do not pose unique disclosure issues but that AI-assisted process claims might. Issues may result from the black-box nature of AI systems, which make it challenging to provide a sufficiently clear and complete disclosure for the invention to be carried out by a POSITA. One potentially problematic case identified was that of diagnostics, where it may be challenging for the applicant to explain the results obtained when these come from the operation of an AI system.

- In these problematic cases a system for the deposit for algorithms used in the development or generation of the AI output (e.g. deposit the training data) might not suitable. The reason is that such a deposit will most likely be insufficient to adequately explain the claims.

- NB: there are currently efforts ongoing at the EPO to find workable solutions to address sufficiency of disclosure in relation to AI outputs.

The final questions were on **policy issues** relating to **AI and patent law (8)**. Experts were asked:

(a) Is the current European patent framework, namely the EPC, suitable to address the challenges posed by AI outputs? If not, what changes (interpretation and reform) should be proposed?

(734) See also Rule 42 EPC on content of the description, especially paras (c) and (e), as well as the Guidelines on Sufficiency of Disclosure F.III.1, fourth paragraph.
(735) See also Rules 31 and 33 EPC.

(b) Assuming legal uncertainty or no legal protection for (certain) AI outputs as inventions under the current legal rules, is there a normative justification under the European patent system (namely the EPC) to grant such outputs protection?

- On these questions most experts were sceptical that there was any justification to change the current rules of the EPC. There seems to be no market failure in the development of AI systems. At most, certain rules may in specific cases be difficult to apply to AI outputs and, where that is the case, it may be justified to make minor adjustments. One such rule where experts admitted there could be a problem is in relation to sufficiency of disclosure for certain AI outputs. However, even then there was no consensus whether a deposit system (similar to the existing system of deposit of biological material) would be of use, for the reasons noted above.
- Beyond this, experts discussed the potential risks that the use of AI systems may pose for the future of patent law, either in terms of over-patenting and in terms of defensive publication and increase of the prior art and public domain. But the lack of a clear market failure justifying intervention was again pointed out as a main reason to conclude that the European Patent framework, namely the EPC, is currently suitable to address the challenges posed by AI outputs.

QUESTIONNAIRE

The in-depth clustered interviews were carried out in a workshop-like setting. This was divided in two parts. Part I covered challenges posed by AI to EU copyright law. Part II covered challenges to European patent law. Each Part was structured around specific topics and questions, to which we allocated a proportionate amount of time. The discussion was guided by team members, which provided a brief explanation of each question and relevant issues arising therefrom. The experts were then invited to discuss each question in-depth.

Preliminary question on conceptual or analytical approach to AI outputs

1. **For analytical purposes, does it make sense to distinguish between types of or assistance by the AI system to the created or produced output? One possible distinction is the following**:

    a) AI-generated output, i.e. autonomously created or produced by the AI system;

    b) AI-assisted output, i.e. created or produced by an AI system with some level of human intervention.

In the following we refer to “AI output(s)” as covering both scenarios above (generated and assisted), and the discussion in Parts I and II should to the extent possible or useful reflect the nuances of these scenarios.

Part I: Artificial Intelligence Challenges to Copyright Law

2 AI outputs and the notion of “work”

a) Can an AI output qualify as a protected “work”?

b) How can we can we apply the current CJEU standard of “author’s own intellectual creation” (originality) resulting from “free and creative choices” to AI outputs?

c) Is it appropriate to assess originality of AI outputs by drawing a conceptual distinction between different phases of authorship? For example, one might distinguish the following phases:

- *Conception* (creative plan/design/specifications);
- *Expression* (execution); and
- *Redaction* (presentation, “curation”).

d) What degree of human intervention (if any) is required for a finding of originality of an AI output during the *conception*, *expression*, and *redaction* phases?

e) In particular, can human intervention in the *conception* and *redaction* phases alone suffice for a finding of originality?

3 Authorship of AI outputs

a) If an AI output does qualify as a "work", to whom should authorship be allocated? Specific examples to be discussed but possibilities include:[736]

- The user of the AI system (most obviously for AI-assisted outputs);
- The "designer(s)" of the AI system (an option e.g. for AI-generated outputs);
- The user and designer as co-authors of a joint work (an option for AI-assisted outputs)

b) In what scenarios are AI outputs "authorless"? What are the legal consequences?

c) To what extent can existing related or sui generis rights provide protection for AI outputs?

4 Policy Questions

a) What policy arguments might justify introducing a related right for AI outputs? Are these arguments convincing in the context of international and EU copyright law and related fields?

b) If protection under a related right were introduced for AI outputs, what type and scope of protection would be justified?

Part II: Artificial Intelligence Challenges to Patent Law

5 Inventorship and ownership of AI outputs

The following questions do not consider the hypothetical scenario where AI systems have legal personhood and can therefore be considered as inventors.

a) To what extent is human inventorship a requirement for patentability? (N.B. in the sense of a human creative or intelligent conception of the invention, or contribution thereto) Or is the requirement merely that a human is *designated* as the inventor?

- See Article 81 and Rule 19 EPC, as well as Article 90 EPC.

b) Are inventor's (moral) rights indicative of a substantive or formal requirement of human inventorship?

- See Articles 58 to 62 EPC.

c) What degree of human contribution/intervention is required for there to be human inventorship in an AI output?

d) What is the effect of automated claim broadening in the assessment of inventorship?

e) If human inventorship is required, either in the conception or merely as designation, to whom is inventorship of an AI output allocated?

(736) These examples are inspired by Jane C. Ginsburg & Luke A. Budiardjo, "Authors and Machines", *Berkeley Technology Law Journal*, Vol. 34, P. 343, 2019. The authors define the designer as the "individual (or set of individuals) who endowed the machine with the training and the creative raw material requisite to the machine's generation of a "creative" output". See *id.*, p. 446.

f) Can the approach followed for biotech inventions, where patents can be granted to non-deterministic outputs, apply to the assessment of inventive step or non-obviousness of AI outputs?

- See also Question 6(d)

g) If the requirement of inventorship is met, to whom does the right over a European patent on an AI output belong?

6 Inventive Step and Industrial Applicability of AI outputs

a. How should the requirement of inventive step and industrial applicability be interpreted in relation to AI outputs? Do they apply in the same way as to human-made inventions? If not, what are the differences?

b. How should the concept of "person having ordinary skill in the art" (POSITA) be interpreted in the assessment of AI outputs? Can the concept of POSITA is this context include other AI Systems?

c. How should prior art be defined in the assessment of AI outputs? What are the relevant fields of technology? Can the relevant prior art include other AI outputs?

d. Can the approach followed for biotech inventions, where patents can be granted to non-deterministic outputs, apply to the assessment of inventive step or non-obviousness to AI outputs?

i. See also Question 5(f).

7 Sufficiency of Disclosure and AI outputs

According to Article 83 EPC, an application must disclose the invention in a manner sufficiently clear and complete for it to be carried out by a POSITA.

a. How should the disclosure requirement be examined in the case of AI outputs?

- See also Rule 42 EPC on content of the description, especially paras (c) and (e), as well as the Guidelines on Sufficiency of Disclosure F.III.1, fourth paragraph.

b. For AI outputs, is it suitable to assess sufficiency of disclosure based on concepts of reproducibility and plausibility of the claimed invention?

c. Is a system for the deposit for algorithms used in the development or generation of the AI output – similar to the existing system of deposit of biological material – suitable in this context?

- See also Rules 31 and 33 EPC.

8 Policy Questions

a. Is the current European Patent framework, namely the EPC, suitable to address the challenges posed by AI outputs? If not, what changes (interpretation and reform) should be proposed?

b. Assuming legal uncertainty or no legal protection for (certain) AI outputs as inventions under the current legal rules is there a normative justification under the European patent system (namely the EPC) to grant such outputs protection?

ANNEX V – Legislation

INTERNATIONAL

- Paris Convention for the Protection of Industrial Property (as amended on September 28, 1979) (**Paris Convention**)
- Berne Convention for the Protection of Literary and Artistic Works, opened for signature Sept. 9, 1886, 828 U.N.T.S. 221 (**Berne Convention**)
- International Convention for the Protection of Performers, Producers of Phonograms and Broadcasting Organisations, adopted 26 October 1961 [entry into force: 18 May 1964] # (**Rome Convention**)
- Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, 1869 U.N.T.S.299 (**TRIPS Agreement**)
- WCT WIPO Copyright Treaty, Dec. 20, 1996, 2186 U.N.T.S. 121 (**WCT**)
- WIPO Performances and Phonograms Treaty, 1996 O.J. (L 89) (**WPPT**).
- Patent Law Treaty, adopted at Geneva on June 1, 2000 (**PLT**).
- Regulations under the Patent Law Treaty (as in force from January 1, 2006) (**PLT Regulations**)
- Patent Cooperation Treaty, done at Washington on June 19, 1970, amended on September 28, 1979, modified on February 3, 1984, and on October 3, 2001 (as in force from April 1, 2002) (**PCT**)
- UN General Assembly, *Universal Declaration of Human Rights*, 10 December 1948, 217 A (III), available at: https://www.refworld.org/docid/3ae6b3712c.html (**Universal Declaration of Human Rights**)

EUROPEAN (NON-EU)

Conventions, Protocols, Implementing Regulations

- Convention on the Grant of European Patents (European Patent Convention) of 5 October 1973 as revised by the Act revising Article 63 EPC of 17 December 1991 and the Act revising the EPC of 29 November 2000 European Patent Convention (**EPC**)[737].
- Implementing Regulations to the Convention on the Grant of European Patents of 5 October 1973 as adopted by decision of the Administrative Council of the European Patent Organisation of 7 December 2006 and as last amended by decision of the Administrative Council of the European Patent Organisation of 27 March 2020 (**EPC Implementation Regulations**)
- Protocol on Jurisdiction and the Recognition of Decisions in respect of the Right to the Grant of a European Patent (Protocol on Recognition) of 5 October 1973. Online: https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ma4.html.(**EPC Protocol on Recognition**)

EPO EPC Guidelines

- Guidelines for Examination in the European Patent Office (November 2019) **(Guidelines)**

(737) The new text of the Convention adopted by the Administrative Council of the European Patent Organisation by decision of 28 June 2001 (see OJ EPO 2001, Special edition No. 4, p. 55) has become an integral part of the Revision Act of 29 November 2000 under Article 3(2), second sentence, of that Act.

EUROPEAN UNION (EU)

Regulations

- Regulation (EC) No 864/2007 of the European Parliament and of the Council of 11 July 2007 on the law applicable to non-contractual obligations (**Rome II Regulation**);

- Regulation (EU) No 1215/2012 of the European Parliament and of the Council of 12 December 2012 on jurisdiction and the recognition and enforcement of judgments in civil and commercial matters (recast) (**Brussels Regulation**).

- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (**General Data Protection Regulation**).

Directives

- Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society **(InfoSoc Directive**)

- Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (**Database Directive**)

- Directive 2006/116/EC of the European Parliament and of the Council of 12 December 2006 on the term of protection of copyright and certain related rights (codified version) (**Term Directive**)

- Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2004 on the enforcement of intellectual property rights (as corrected in OJ L 157, 30.4.2004) (**Enforcement Directive**)

- Directive 2006/115/EC of the European Parliament and of the Council of 12 December 2006 on rental right and lending right and on certain rights related to copyright in the field of intellectual property (codified version) (**Rental and Lending Rights Directive**).

- Directive 2009/24/EC of the European Parliament and of the Council of 23 April 2009 on the legal protection of computer programs (Codified version) (Computer Programs Directive).;

- Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ L 157, 15.6.2016, p. 1–18 (**Trade Secrets Directive)**

- Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC (**CDSM Directive**).

ANNEX VI – Policy Documents

EUROPEAN PARLIAMENT

- EP, Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), paras 136 – 137 (Interoperability, access to code and intellectual property rights).
- [EP JURI, 'Explanatory Statement', European Parliament 2017]
- EP, 'Resolution on a comprehensive European industrial policy on artificial intelligence and robotics', (2018/2088 (INI)), 12 February 2019,
- DRAFT REPORT on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)), Committee on Legal Affairs of the EU Parliament, (Rapporteur: Stéphane Séjourné), 24.4.2020, https://www.europarl.europa.eu/doceo/document/JURI-PR-650527_EN.html?redirect.
- European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)

EUROPEAN COMMISSION

- EC, 'Artificial Intelligence for Europe', Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2018) 237 Final, 25 April 2018.
- EC, 'Coordinated plan on Artificial Intelligence', Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2018) 795 Final, 7 December 2018
- EC, 'Trends and Developments in Artificial Intelligence – Challenges to the Intellectual Property Rights Framework', Shaping Europe's digital future, SMART 2018/0052, March 2019 ("Call for Tender").
- EC, 'Building Trust in Human-Centric Artificial Intelligence', Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, COM (2019) 168 final, 8 April 2019.
- EC, European strategy for data, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Study Guidelines 2020, COM (2020) 66 final, 19 February 2020.
- EC, CNECT/2020/OP/0009 — Study on Copyright and New technologies: Copyright Data Management and Artificial Intelligence — Smart 2019/0038
- EC, 'White paper On Artificial Intelligence – A European approach to excellence and trust', COM (2020) 65 final, 19 February 2020.

ANNEX VII – Table of Cases

Court of Justice of the European Union (CJEU)

- Case C-46/02 Fixtures Marketing Ltd v Oy Veikkaus Ab (2004) ECLI:EU:C:2004:694 (***Fixtures Marketing Ltd v Oy Veikkaus Ab)***
- Case C-203/02 The British Horseracing Board Ltd and Others v William Hill Organisation Ltd. (2004) ECLI:EU:C:2004:695 (***British Horseracing Board and others***)
- Case C-338/02 Fixtures Marketing Ltd v Svenska AB (2004) ECLI:EU:C:2004:696 (***Fixtures Marketing Ltd v Svenska AB***),
- Case C-444/02 Fixtures Marketing Ltd v Organismos prognostikon agonon podosfairou AE (OPAP), (2004) ECLI:EU:C:2004:697 (***Fixtures Marketing Ltd v OPAP***).
- Case C-05/08 Infopaq International v Danske Dagblades Forening (2009) ECLI:EU:C:2009:465 (***Infopaq***)
- Case C-393/09 Bezpečnostní softwarová asociace - Svaz softwarové ochrany v Ministerstvo kultury (2010) ECLI:EU:C:2010:816 (***BSA***)
- Case C-145/10 Eva-Maria Painer (2011) ECLI:EU:C:2011:798 (***Painer***)
- Joined Cases Football Association Premier League Ltd and Others v QC Leisure and Others (C-403/08) and Karen Murphy v Media Protection Services Ltd (C-429/08) (2011) ECLI:EU:C:2011:631 (***Premier League***)
- Case C-604/10 Football Dataco Ltd and Others v Yahoo! UK Ltd and Others (2012) ECLI:EU:C:2012:115 (***Football Dataco***)
- Case C-277/10 Martin Luksan v Petrus van der Let (2012) ECLI:EU:C:2012:65 (***Luksan***)
- Case C-355/12 Nintendo Co. Ltd and Others v PC Box Srl and 9Net Srl (2014) ECLI:EU:C:2014:25 (***Nintendo***)
- Case C-572/13 Hewlett-Packard Belgium SPRL v Reprobel SCRL (2015) ECLI:EU:C:2015:750 (***Reprobel***)
- Case C-310/17 Levola Hengelo BV v Smilde Foods BV (2018) ECLI:EU:C:2018:899 (***Levola Hengelo)***
- Case C-476/17 Pelham GmbH and Others v Ralf Hütter and Florian Schneider-Esleben (2019) ECLI:EU:C:2019:624 (***Pelham***)
- Case C-469/17 Funke Medien NRW GmbH v Bundesrepublik Deutschland (2019) ECLI:EU:C:2019:623 (***Funke Medien***)
- Case C-683/17 Cofemel – Sociedade de Vestuário SA v G-Star Raw CV (2019) ECLI:EU:C:2019:721 (***Cofemel***)
- Case C-833/18 SI and Brompton Bicycle Ltd v Chedech / Get2Get (2020) ECLI:EU:C:2020:461 (***Brompton Bicycle***)

Advocate General (AG) Opinions

- AG Trstenjak, Opinion in Eva-Maria Painer v Standard VerlagsGmbH, Axel Springer AG, Süddeutsche Zeitung GmbH, SPIEGEL-Verlag Rudolf AUGSTEIN GmbH & Co KG, and Verlag M. DuMont Schauberg Expedition der Kölnischen Zeitung GmbH & Co KG C-145/10) [2011] (Painer)
- AG Szpunar, Opinion in Pelham GmbH v Hutter (C-476/17) EU:C:2018:1002; [2019], (Pelham)

NATIONAL COURTS

Austria

- Österreichischen Obersten Gerichtshofs vom 6. Juli 2004 – 4 Ob 133/04v , ZUM-RD 2005, 11

Belgium

- Court of Brussel, 23 May 2017 (Diplomatic card v. Forax), I.R.D.O., 2017, 204 [functional and technical specifications protected as part of computer program].
- Cour d'Appel Bordeaux, 31 January 2005, No. 03/05512

France

- TGI Paris, 6 July 1970, RIDA 190 (1970)(affaire Paris Match);
- Cass. Ass. Plen., 7 March 1986: two cases: Atari and Williams Electronics, R.I.D.A. 1986, no. 129, 136
- Paris 8e ch., 17 June 1988, D. 1989, somm. 44, obs. Colombet
- TGI Paris 5 July 2000, No. 97/24872 (Matt Cooper v. Ogilvy and Mather),
- Cass. Civ. I, 8 jan, 2002, R.I.D.A. 2002, no. 193, 321 (Père-Lachaise cemetery map).
- TGI Paris, 3e ch., 6 Oct. 2009, R.I.D.A. 2010, no. 226, 506 [aerial photographs not copyright protected for lack of originality].
- Cass. 1re civ., 15 janv. 2015, n° 13-23.566, D. 2015, p. 206 et p. 2215

Germany

- LG Berlin 30 May 1989, case 16 O 33/89, GRUR 1990, 270 (Satellite photo).
- OLF Hamburg, June 29, 1995, §997, ZUM-RD 217.
- OLG Dusseldorf, Feb. 13, 1996, 1997 GRUR 49
- LG Hamburg 4 April 2003, case O 515/02, ZUM 2004, 675;
- OLG Hamm, aug. 24, 2004, 2004, ZUM 927
- Court of Appeal (OLG) Karlsruhe 14 April 2010, case 6 U 46/09, GRUR-RR 2010, 234
- See BGH 06.10.2016 - I ZR 25/15, World of Warcraft I, GRUR 2017, 266
- LG Berlin, Urteil vom 20.06.2017 - 16 O 59/16, ZUM 2017, 955.
- BGH 20 December 2018, I ZR 104/17 [photographs of public domain paintings protected by related right];
- KG Berlin, Urt. v. 16.1.2020 – 2 U 12/16.Kart, GRUR 2020, 280.
- BGH 30 april 2020 - I ZR 139/15 - Afghanistan Papiere II

The Netherlands

- Kluwer v. Lamoth, Supreme Court of the Netherlands, June 1, 1990, Nederlandse Jurisprudentie 1991, 377

- Zonen Endstra v. Nieuw Amsterdam (2008) ECLI:NL:HR:2008:BC2153, HR 30.05.2008, NJ 2008, 556
- Dutch Supreme Court (Hoge Raad) 19 January 2018, ECLI:NL:HR:2018:56, NJ 2018/237 (Diplomatic Card v. Forax)

UK

- Express Newspapers v Liverpool Daily Post [1985] 1 W.L.R. 1089, 1093
- Football Dataco Ltd & Ors v. Stan James plc & Ors and Sportradar GmbH and Anor [2013] EWCA Civ 27

EPO

- Case G 0003/92 (13 June 1994). Online: https://www.epo.org/law-practice/case-law-appeals/recent/g920003ex1.html
- EPO Technical Board of Appeal, case T 939/92 (AgrEvo), 12 September 1995. Online: https://www.epo.org/law-practice/case-law-appeals/recent/t920939ex1.htmlhttps://www.epo.org/law-practice/case-law-appeals/pdf/t920939ex1.pdf.
- EPO Boards of Appeal, case T 641/00 (Two Identities/Comvik), 26 September 2002. Online: https://www.epo.org/law-practice/case-law-appeals/recent/t000641ep1.html
- EPO Legal Board of Appeal, The Trustees of Dartmouth College, 4 February 2004, Case J2/01, r.2.6
- EPO Boards of Appeal, case T 1329/04 (Factor-9/John Hopkins), 28 June 2005. Online: https://www.epo.org/law-practice/case-law-appeals/recent/t041329eu1.html.
- EPO decision of 27 January 2020 on EP 18 275 163

UKIPO

- [UKIPO, Formalities Manual (Online Version), Chapter 3: The Inventor, Sections (3.01 - 3.38) (last updated: October 2019), https://www.gov.uk/guidance/formalities-manual-online-version/chapter-3-the-inventor#ref3-02.]
- UKIPO, Case BL O/741/19 (4 December 2019), online: https://bit.ly/2zCPdYm

USPTO

- USPTO, In Re Application No 16/524,350 ("FlashPoint IP Ltd."), Decision on Petition, 22 April 2020. Online: https://bit.ly/2YMINyo.
- USPTO, Compendium of U.S. Copyright Office Practices § 101 (3d ed. 2017). Online: https://www.copyright.gov/comp3/
- USPTO, Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation, Federal Register Docket No. PTO-C-2019-0038 (Document Number 2019-23638) 30 October 2019, https://www.federalregister.gov/documents/2019/10/30/2019-23638/request-for-comments-on-intellectual-property-protection-for-artificial-intelligence-innovation.
- USPTO, Public Views on Artificial Intelligence and Intellectual Property Policy (October 2020), https://www.uspto.gov/sites/default/files/documents/USPTO_AI-Report_2020-10-07.pdf.

WTO

- United States - Section 211 Omnibus Appropriations Act of 1998 - Report of the Appellate Body, WT/DS176/AB/R (02/01/2002),

ANNEX VIII – References for Legal Analysis

Abbott, Ryan. "Everything Is Obvious by Ryan Abbott." *UCLA Law Review* 66, no. 2 (2019): 2–52. "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law." *Boston College Law Review* 57, no. 4 (2016): 1079–1126. https://doi.org/10.2139/ssrn.2727884.

AI-HLEG. "A Definition of AI: Main Capabilities and Disciplines." High-Level Expert Group on Artificial Intelligence (AI HLEG)., April 8, 2019. https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top.

Baldus, Oliver. "A Practical Guide on How to Patent Artificial Intelligence (AI) Inventions and Computer Programs within the German and European Patent System: Much Ado about Little." *EIPR* 41, no. 12 (2019): 750–54.

Bently, Lionel, Brad Sherman, Dev Gangjee, and Philip Johnson. *Intellectual Property Law*. Fifth Edition. OUP, 2018. https://global.oup.com/ukhe/product/intellectual-property-law-9780198769958?cc=nl&lang=en&.

Bloch-Wehba, Hannah. "Access to Algorithms." *Fordham Law Review* 88 (2020). https://papers.ssrn.com/abstract=3355776.

Blok, Peter Hendrik. "The Inventor's New Tool : Artificial Intelligence : How Does It Fit in the European Patent System?" *European Intellectual Property Review* 39, no. 2 (2017): 69–73.

Bonadio, Enrico, and Luke McDonagh. "Artificial Intelligence as Producer and Consumer of Copyright Works: Evaluating the Consequences of Algorithmic Creativity." *Intellectual Property Quarterly* 2 (2020): 112–37.

Boyden, Bruce E. "Emergent Works." *The Columbia Journal of Law & the Arts* 39, no. 3 (June 20, 2016): 377–94. https://doi.org/10.7916/jla.v39i3.2077.

Bridy, Annemarie. "Coding Creativity: Copyright and the Artificially Intelligent Author." *Stanford Law Review* 5, no. 25 (July 18, 2011): 1–28. "The Evolution of Authorship: Work Made by Code." *Columbia Journal of Law & the Arts* 39 (2016): 395–401.

Butler, Timothy L. "Can a Computer Be an Author - Copyright Aspects of Artificial Intelligence." *Hastings Comm. & Ent.L.J.* 4 (1982).

Calo, Ryan. "Artificial Intelligence Policy: A Primer and Roadmap" 51 (2017): 399–435.

Chiou, Theodoros. "Copyright Lessons on Machine Learning: What Impact on Algorithmic Art?" *JIPITEC* 10, no. 3 (February 21, 2020). https://www.jipitec.eu/issues/jipitec-10-3-2019/5025.

CIPA. "Patenting Inventions Created Using an AI System: A CIPA Discussion Paper." CIPA - Chartered Institute of Patent Attorneys, 2020. https://www.cipa.org.uk/_resources/assets/attachment/full/0/260456.pdf.

CLSPA, Alexandra Bensamoun, and Joëlle Farchy. "Mission du CSPLA sur les enjeux juridiques et économiques de l'intelligence artificielle dans les secteurs de la création culturelle." CLSPA - Conseil Supérieur de la Propriété Littéraire et Artistique, 2020. https://www.culture.gouv.fr/Sites-thematiques/Propriete-litteraire-et-artistique/Conseil-superieur-de-la-propriete-litteraire-et-artistique/Travaux/Missions/Mission-du-CSPLA-sur-les-enjeux-juridiques-et-economiques-de-l-intelligence-artificielle-dans-les-secteurs-de-la-creation-culturelle.

Coche, Eugenie. "Heks'nkaas or the 'Fifty Shades of Taste' Explained by the CJEU through EU Copyright Law." *EIPR*, 2019, 173–80.

Craig, Carys J., and Ian R. Kerr. "The Death of the AI Author." *Osgoode Legal Studies Research Paper*, March 25, 2019. https://doi.org/10.2139/ssrn.3374951.

David Silver, Demis Hassabis. "AlphaGo Zero: Starting from scratch." Deepmind, October 18, 2017. /blog/article/alphago-zero-starting-scratch.

Davidson, Mark J., and P. Bernt Hugenholtz. "Football Fixtures, Horseraces and Spinoffs: The ECJ Domesticates the Database Right." *EIPR*, no. 3 (2005): 113–18.

Delipetrev, B., G. De Prato, F. Martínez-Plumed, E. Gomez, M. López Cobo, and S. Samoili. "AI Watch : Defining Artificial Intelligence : Towards an Operational Definition and Taxonomy of Artificial Intelligence." JRC Working Papers. AI Watch. Joint Research Centre, February 27, 2020. http://op.europa.eu/en/publication-detail/-/publication/6cc0f1b6-59dd-11ea-8b81-01aa75ed71a1/language-en.

Deltorn, Jean-Marc, and Franck Macrez. "Authorship in the Age of Machine Learning and Artificial Intelligence." In *The Oxford Handbook of Music Law and Policy*. Rochester, NY: OUP, 2018. https://doi.org/10.2139/ssrn.3261329.

Denicola, Robert. "Ex Machina: Copyright Protection for Computer-Generated Works." *Rutgers University Law Review* 69 (2016): 251–87.

Derclaye, Estelle. "The CJEU Decision in Brompton Bicycle – A Welcome Double Rejection of the Multiplicity of Shapes and Causality Theories in Copyright Law." *Kluwer Copyright Blog* (blog), June 25, 2020. http://copyrightblog.kluweriplaw.com/2020/06/25/the-cjeu-decision-in-brompton-bicycle-a-welcome-double-rejection-of-the-multiplicity-of-shapes-and-causality-theories-in-copyright-law/.

Dhariwal, Prafulla, Heewoo Jun, Christine Payne, Jong Wook Kim, Alec Radford, and Ilya Sutskever. "Jukebox: A Generative Model for Music." *ArXiv:2005.00341 [Cs, Eess, Stat]*, April 30, 2020. http://arxiv.org/abs/2005.00341.

Dietz, Adolf. "Le Concept d'auteur Selon Le Droit de La Convention de Berne." *RIDA* 155 (1993).

Drexl, Josef, Reto Hilty, Francisco Beneke, Luc Desaunettes, Michèle Finck, Jure Globocnik, Begoña Gonzalez Otero, et al. "Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, October 8, 2019. https://doi.org/10.2139/ssrn.3465577. "Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Law Perspective." *Max Planck Institute for Innovation & Competition* 11, no. 3 (October 8, 2019). https://doi.org/10.2139/ssrn.3465577.

Ebrahim, Tabrez. "Automation & Predictive Analytics in Patent Prosecution: USPTO Implication & Policy." *Georgia State University Law Review* 35, no. 4 (June 1, 2019). https://readingroom.law.gsu.edu/gsulr/vol35/iss4/5.

Eechoud, Mireille van. "Along the Road to Uniformity - Diverse Readings of the Court of Justice Judgments on Copyright Work." *JIPITEC* 3, no. 1 (May 15, 2012). http://www.jipitec.eu/issues/jipitec-3-1-2012/3322.

Ertel, Wolfgang. *Introduction to Artificial Intelligence*. Undergraduate Topics in Computer Science. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-58487-4. "Neural Networks." In *Introduction to Artificial Intelligence*, edited by Wolfgang Ertel, 245–87. Undergraduate Topics in Computer Science. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-58487-4_9. "Reinforcement Learning." In *Introduction to Artificial Intelligence*, edited by Wolfgang Ertel, 289–311. Undergraduate Topics in Computer Science. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-58487-4_10.

European Patent Office, Yann Ménière, Ilja Rudyk, and Lucas Tsitsilonis. "Patents and Self-Driving Vehicles." EPO, 2018. https://www.epo.org/service-support/publications.html?pubid=177#tab3.

Fabris, Daniele. "From the PHOSITA to the MOSITA: Will 'Secondary Considerations' Save Pharmaceutical Patents from Artificial Intelligence?" *IIC - International Review of Intellectual Property and Competition Law* 51, no. 6 (July 1, 2020): 685–708. https://doi.org/10.1007/s40319-020-00953-8.

Ficsor, Mihály. *The Law of Copyright and the Internet: The 1996 WIPO Treaties, Their Interpretation and Implementation*. Oxford, New York: Oxford University Press, 2002.

Fleming, Nic. “How Artificial Intelligence Is Changing Drug Discovery.” *Nature* 557, no. 7707 (May 30, 2018): S55–57. https://doi.org/10.1038/d41586-018-05267-x.

Frank Pasquale. *The Black Box Society*, 2016. https://www.hup.harvard.edu/catalog.php?isbn=9780674970847.

Frankel, Susy, and Daniel J. Gervais. *Advanced Introduction to International Intellectual Property*. Cheltenham, UK ; Northampton, MA, USA: Edward Elgar Pub, 2016.

Fromm, F. “Der Apparat Als Geistiger Schopfer.” *GRUR*, 1964.

Fu, Jia-Chen. “Project 523 Behind Tu’s Nobel Prize Win for Malaria.” *The Conversation*, October 6, 2015. https://www.usnews.com/news/articles/2015/10/06/project-523-behind-tus-nobel-prize-win-for-malaria.

Gervais, Daniel. “Is Intellectual Property Law Ready for Artificial Intelligence?” *GRUR International* 69, no. 2 (February 1, 2020): 117–18. https://doi.org/10.1093/grurint/ikz025. “The Protection under International Copyright Law of Works Created with or by Computers.” *IIC - International Review of Intellectual Property and Competition Law* 5 (1991): 629–60.

Gervais, Daniel J. “The Machine As Author.” *Iowa Law Review* 105 (March 25, 2019): 19–35. “The Protection of Databases.” *Chicago-Kent Law Review* 82, no. 3 (2007). https://papers.ssrn.com/abstract=1116643.

Ginsburg, Jane. “The Concept of Authorship in Comparative Copyright Law.” *DePaul Law Review* 52, no. 4 (June 1, 2003): 1063.

Ginsburg, Jane C. “People Not Machines: Authorship and What It Means in the Berne Convention.” *IIC - International Review of Intellectual Property and Competition Law* 49, no. 2 (January 29, 2018): 131–35. https://doi.org/10.1007/s40319-018-0670-x.

Ginsburg, Jane C., and Luke Ali Budiardjo. “Authors and Machines.” *Berkeley Technology Law Journal* 34, no. 2 (October 21, 2019): 343–448. https://doi.org/10.2139/ssrn.3233885.

Gompel, Stef van. “Creativity, Autonomy and Personal Touch. A Critical Appraisal of the CJEU’s Originality Test for Copyright.” In M. van Eechoud (ed.), *The Work of Authorship*. Amsterdam: Amsterdam University Press, 2014.

Gonzalez Otero, Begoña, and Joao Pedro Quintais. “Before the Singularity: Copyright and the Challenges of Artificial Intelligence.” *Kluwer Copyright Blog* (blog), 10 2018. http://copyrightblog.kluweriplaw.com/2018/09/25/singularity-copyright-challenges-artificial-intelligence/?doing_wp_cron=1594818263.1780009269714355468750.

Grimmelmann, James. “Copyright for Literate Robots.” *Iowa Law Review* 101 (January 2016): 657–81. https://doi.org/10.31228/osf.io/z38qm. “There’s No Such Thing as a Computer-Authored Work — And It’s a Good Thing, Too.” *Columbia Journal of Law & the Arts* 39 (May 4, 2017): 403–16. https://doi.org/10.31228/osf.io/rk8cm.

Grubb, Philip W., Peter R. Thomsen, Tom Hoxie, and Gordon Wright. *Patents for Chemicals, Pharmaceuticals, and Biotechnology*. Sixth Edition. Oxford, New York: Oxford University Press, 2016.

Guadamuz, Andrés. “Do Androids Dream of Electric Copyright? Comparative Analysis of Originality in Artificial Intelligence Generated Works.” *Intellectual Property Quarterly* 2 (June 5, 2017): 169–86.

Hamon, Ronan, Henrik Junklewitz, and Ignacio Sanchez Martin Jose. “Robustness and Explainability of Artificial Intelligence.” EUR - Scientific and Technical Research Reports. Publications Office of the European Union, 2020. JRC119336. https://publications.jrc.ec.europa.eu/repository/handle/111111111/58835.

Hattenbach, Ben, and Joshua Glucoft. “Patens in an Era of Infinite Monekys and Artificial Intelligence.” *Stanford Technology Law Review* 19, no. 2 (October 2015): 32–51.

Hattenbach, Ben, and Gavin Snyder. “Rethinking the Mental Steps Doctrine and Other Barriers to Patentability of Artificial Intelligence.” *The Columbia Science and Technology Law Review* 19, no. 2 (May 15, 2018): 313–39.

Hewitt, Stephen. “Protection of Works Created by the Use of Computers.” *New L.J.* 133 (1983).

Hilty, Reto, Jörg Hoffmann, and Stefan Scheuerer. “Intellectual Property Justification for Artificial Intelligence.” *Oxford University Press*, no. 5 (February 11, 2020): 1–29.

Hugenholtz, P. Bernt. “Neighbouring Rights Are Obsolete.” *IIC - International Review of Intellectual Property and Competition Law* 50, no. 8 (October 1, 2019): 1006–11. https://doi.org/10.1007/s40319-019-00864-3.

Hugenholtz, P. Bernt, Mireille van Eechoud, Stef van Gompel, and Natali Helberger. “The Recasting of Copyright & Related Rights for the Knowledge Economy.” Report to the European Commission, DG Internal Market. Amsterdam: IViR, 2006. https://papers.ssrn.com/abstract=2018238.

Hughes, Rose. “EPO Refuses ‘AI Inventor’ Applications in Short Order - AI Inventor Team Intend to Appeal.” *The IPKat* (blog), December 22, 2019. http://ipkitten.blogspot.com/2019/12/epo-refuses-ai-inventor-applications-in.html.

Iglesias Portela, Maria, Sheron Shamuilia, and Amenda Anderberg. “Intellectual Property and Artificial Intelligence - A Literature Review.” Luxembourg: European Commission, December 17, 2019. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC119102/intellectual_property_and_artificial_intelligence_jrc_template_final.pdf.

Kaminski, Margot. “The Right to Explanation, Explained.” *Berkeley Technology Law Journal* 34 (January 1, 2019). https://doi.org/10.15779/Z38TD9N83H.

Kaminski, Margot E. “Authorship, Disrupted: AI Authors in Copyright and First Amendment Law.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, December 12, 2017. https://papers.ssrn.com/abstract=3086912.

Kim, Daria. “‘AI-Generated Inventions’: Time to Get the Record Straight?” *GRUR International* 69, no. 5 (May 1, 2020): 443–56. https://doi.org/10.1093/grurint/ikaa061.

Kluwer-Patent-Blogger. “CIPA Calls for Clarity about Artificial Intelligence and Patents.” *Kluwer Patent Blog* (blog), June 5, 2020. http://patentblog.kluweriplaw.com/2020/06/05/cipa-calls-for-clarity-about-artificial-intelligence-and-patents/. “EPO: A Machine Cannot Be an Inventor.” *Kluwer Patent Blog* (blog), January 29, 2020. http://patentblog.kluweriplaw.com/2020/01/29/epo-a-machine-cannot-be-an-inventor/.

Knight, Will. “The Dark Secret at the Heart of AI.” *MIT Technology Review*, April 11, 2017. https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/.

Knutson, Kaelyn. “Anything You Can Do, AI Can’t Do Better: An Analysis of Conception as a Requirement For Patent Inventorship And A Rationale For Excluding AI Inventors.” *Cybaris An Intell. Prop. L. Rev* 11, no. 2 (January 1, 2020). https://open.mitchellhamline.edu/cybaris/vol11/iss2/2.

Kraetzig, Viktoria. “Does Copyright Law Have to Balance Fundamental Rights beyond the Written Exceptions? Unfortunately, the German Federal Supreme Court Has Recently Left This Question Open.” *Kluwer Copyright Blog* (blog), May 28, 2020. http://copyrightblog.kluweriplaw.com/2020/05/28/does-copyright-law-have-to-balance-fundamental-rights-beyond-the-written-exceptions-unfortunately-the-german-federal-supreme-court-has-recently-left-this-question-open/.

Kubat, Miroslav. *An Introduction to Machine Learning*. Springer International Publishing, 2015. https://doi.org/10.1007/978-3-319-20010-1.

Kummer, Max. *Das Urheberrechtlich Schützbare Werk*. Stämpfli, 1968.

Landes, William M., and Richard A. Posner. *The Economic Structure of Intellectual Property Law*. First Edition. Cambridge, Mass: Belknap Press: An Imprint of Harvard University Press, 2003.

Lauber-Rönsberg, Anne, and Sven Hetmank. “The Concept of Authorship and Inventorship under Pressure: Does Artificial Intelligence Shift Paradigms?” *Journal of Intellectual Property Law & Practice* 14, no. 7 (July 1, 2019): 570–79. https://doi.org/10.1093/jiplp/jpz061.

Lauder, Ed. "The Japan Patent Office Deploying AI to Screen Patent Application - AI Business." *Informa - AI Business*, April 24, 2017. https://aibusiness.com/document.asp?doc_id=760272&site=aibusiness.

Lim, Daryl. "AI & IP: Innovation & Creativity in an Age of Accelerated Change." *Akron Law Review* 52 (2018): 813–75.

Luginbuehl, Stefan. "Patent Protection of Inventions Involving Artificial Intelligence." In *Transition and Coherence Ii Intellectual Property Law*, 2020.

Martin Senftleben; Laurens Buijtelaar. "Robot Creativity: An Incentive-Based Neighbouring Rights Approach." *European Intellectual Property Review*, 2020.

McLaughlin, Michael. "Computer-Generated Inventions." *American University Washington College of Law*, January 7, 2018, 1–32. https://doi.org/10.2139/ssrn.3097822.

Melendez, Carlos. "Is There a Future for Off-the-Shelf AI?" *InfoWorld*, December 12, 2017. https://www.infoworld.com/article/3241885/is-there-a-future-for-off-the-shelf-ai.html.

Mezei, Péter. "From Leonardo to the Next Rembrandt – The Need for AI-Pessimism in the Age of Algorithms." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, May 4, 2020. https://papers.ssrn.com/abstract=3592187.

Microsoft reporter. "The Next Rembrandt: Recreating the Work of a Master with AI." *Blurring the Lines between Art, Technology and Emotion: The Next Rembrandt* (blog), 04 2016. https://news.microsoft.com/europe/features/next-rembrandt/.

Milde, Karl F. Jr. "Can a Computer Be and Author or an Inventor." *Journal of the Patent Office Society* 51, no. 6 (1969): 378–406.

Newitz, Annalee. "An AI Wrote All of David Hasselhoff's Lines in This Bizarre Short Film." *Ars Technica*, April 25, 2017. https://arstechnica.com/gaming/2017/04/an-ai-wrote-all-of-david-hasselhoffs-lines-in-this-demented-short-film/.

Newman, Daniel. "Why AI As A Service Will Take Off In 2020." *Forbes*. Accessed July 19, 2020. https://www.forbes.com/sites/danielnewman/2020/01/07/why-ai-as-a-service-will-take-off-in-2020/.

Niklas, Jędrzej, and Lina Dencik. "European Artificial Intelligence Policy: Mapping the Institutional Landscape." Working Paper DATAJUSTICE project. DATAJUSTICE, 2020.

Nordemann, Jan Bernd. "AIPPI: No Copyright Protection for AI Works without Human Input, but Related Rights Remain." *Kluwer Copyright Blog* (blog), November 21, 2019. http://copyrightblog.kluweriplaw.com/2019/11/21/aippi-no-copyright-protection-for-ai-works-without-human-input-but-related-rights-remain/.

OECD. *Artificial Intelligence in Society*. Paris: OECD, 2019. https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm.

Osha, J.P., G. Henry, R. Nack, and L. Shen. "AIPPI Resolution on Copyright in Artificially Generated Works." Resolution. London, 2019.

Pihlajamaa, Heli. "Legal Aspects of Patenting Inventions Involving Artificial Intelligence Summary of Feedback by EPC Contracting States." Presentation. EPO, 2020.

Pila, Justine, and Paul Torremans. *European Intellectual Property Law*. Second Edition. Oxford, New York: Oxford University Press, 2019.

Porter, Jon. "OpenAI's MuseNet Generates AI Music at the Push of a Button." *The Verge* (blog), April 26, 2019. https://www.theverge.com/2019/4/26/18517803/openai-musenet-artificial-intelligence-ai-music-generation-lady-gaga-harry-potter-mozart.

Ramalho, Ana. “Patentability of AI-Generated Inventions: Is a Reform of the Patent System Needed?” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, February 15, 2018. https://doi.org/10.2139/ssrn.3168703. “Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems.” *Journal of Internet Law* 21, no. 1 (June 13, 2017): 12–25. https://doi.org/10.2139/ssrn.2987757.

Ramalho, Ana Quintela Ribeiro Neves. “Originality Redux: An Analysis of the Originality Requirement in AI-Generated Works.” *AIDA*, 2019, 2–18.

Reda, Julia. “German Federal Supreme Court Defends Press Freedom in Two High-Profile Copyright Cases, No Resolution of Sampling Dispute.” *Kluwer Copyright Blog* (blog), May 1, 2020. http://copyrightblog.kluweriplaw.com/2020/05/01/german-federal-supreme-court-defends-press-freedom-in-two-high-profile-copyright-cases-no-resolution-of-sampling-dispute/.

Reinbothe, Jörg, and Silke Von Lewinski. *The WIPO Treaties on Copyright: A Commentary on the WCT, the WPPT, and the BTAP*. Oxford, New York: Oxford University Press, 2015.

Ricketson, Sam. “The 1992 Horace S. Manges Lecture - People or Machines: The Bern Convention and the Changing Concept of Authorship.” *Columbia-VLA Journal of Law & the Arts* 16, no. 1 (1992 1991): 1–38.

Rockwell Anyoha. “The History of Artificial Intelligence.” *Science in the News* (blog), August 28, 2017. http://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/.

Ryan Abbott. “The Artificial Inventor Project.” *WIPO Magazine*, December 2019. https://www.wipo.int/wipo_magazine/en/2019/06/article_0002.html.

Sample, Ian. “Powerful Antibiotic Discovered Using Machine Learning for First Time.” *The Guardian*, February 20, 2020, sec. Society. https://www.theguardian.com/society/2020/feb/20/antibiotic-that-kills-drug-resistant-bacteria-discovered-through-ai.

Samuelson, Pamela. “Allocating Ownership Rights in Computer-Generated Works.” In *Symposium Cosponsored by University of Pittsburgh Law Review and The Software En on The Future of Software Protection*, 1185–1228. Pittsburgh, Pennsylvania, USA: University of Pittsburgh Press, 1986. “Reconceptualizing Copyright’s Merger Doctrine.” *Journal of the Copyright Society of the U.S.A.* 63 (2016). https://papers.ssrn.com/abstract=2763903.

Schuster, W. Michael. “Artificial Intelligence and Patent Ownership.” *Washington and Lee Law Review* 75, no. 4 (February 19, 2019): 1945.

Selbst, Andrew D., and Solon Barocas. “The Intuitive Appeal of Explainable Machines.” *Fordham Law Review* 87 (2018). https://doi.org/10.2139/ssrn.3126971.

Selbst, Andrew D., and Julia Powles. “Meaningful Information and the Right to Explanation.” *International Data Privacy Law* 7, no. 4 (November 1, 2017): 233–42. https://doi.org/10.1093/idpl/ipx022.

Sganga, Caterina. “Say Nay to a Tastier Copyright: Why the CJEU Should Deny Copyright Protection for Taste (and Smells).” *Journal of Intellectual Property Law & Practice* 14, no. 3 (March 1, 2019): 187–96. https://doi.org/10.1093/jiplp/jpy161.

Shemtov, Noam. “A Study on Inventorship in Inventions Involving AI Activity.” Reader in Intellectual Property and Technology Law, Centre for Commercial Law Studies. Queen Mary University of London: EPO, February 2019. http://documents.epo.org/projects/babylon/eponet.nsf/0/3918F57B010A3540C125841900280653/%24File/Concept_of_Inventorship_in_Inventions_involving_AI_Activity_en.pdf.

Silver, David, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, et al. “Mastering the Game of Go with Deep Neural Networks and Tree Search.” *Nature* 529, no. 7587 (January 2016): 484–89. https://doi.org/10.1038/nature16961.

Solon Barocas, Moritz Hardt, Arvind Narayanan. *Fairness and Machine Learning.* fairmlbook.org, 2019. https://fairmlbook.org/.

Spindler, Gerald. “Copyright Law and Artificial Intelligence.” *IIC - International Review of Intellectual Property and Competition Law* 50, no. 9 (November 1, 2019): 1049–51. https://doi.org/10.1007/s40319-019-00879-w.

Vermeulen, Erik, and Wulf Kaal. “How to Regulate Disruptive Innovation: From Facts to Data.” *Jurimetrics: The Journal of Law, Science, and Technology* 57, no. 2 (2017). https://doi.org/10.2139/ssrn.2808044.

Vertinsky, Liza. “Boundary-Spanning Collaboration and the Limits of Joint Inventorship Doctrine.” *Houston Law Review* 55, no. 2 (November 20, 2017): 3887.

Vertinsky, Liza, and Todd M. Rice. “Thinking about Thinking Machines: Implications of Machine Inventors for Patent Law.” *Boston University Journal of Science & Technology Law* 8, no. 2 (2002): 574–613.

Vincent, James. “OpenAI Has Published the Text-Generating AI It Said Was Too Dangerous to Share.” *The Verge* (blog), November 7, 2019. https://www.theverge.com/2019/11/7/20953040/openai-text-generation-ai-gpt-2-full-model-release-1-5b-parameters.

Visser, Derk, Laurence Lai, Peter de Lange, and Kaisa Suominen. *Visser’s Annotated European Patent Convention (EPC).* 2019 Ed. Wolters Kluwer, 2019. https://lrus-prod-active-cd-01.azurewebsites.net.

Vousden, Stephen. “Protecting GUIs in EU Law: Bezpečnostní Softwarová Asociace.” *Journal of Intellectual Property Law & Practice* 6, no. 10 (October 1, 2011): 728–37. https://doi.org/10.1093/jiplp/jpr111.

Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation.” *International Data Privacy Law* 7, no. 2 (May 1, 2017): 76–99. https://doi.org/10.1093/idpl/ipx005.

Wachter, Sandra, Brent Mittelstadt, and Chris Russell. “Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, October 6, 2017. https://doi.org/10.2139/ssrn.3063289.

Wild, Joff. “Artificial Intelligence and the Future of the Patent System | IAM.” *IAM* (blog), July 11, 2018. https://www.iam-media.com/law-policy/artificial-intelligence-and-future-patent-system.

WIPO. “Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence.” WIPO, May 21, 2020. https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=499504. “WIPO Technology Trends 2019: Artificial Intelligence.” World Intellectual Property Organisation, 2019. https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=5982426.

World Economic Forum. “Artificial Intelligence Collides with Patent Law.” White Paper. Geneva: World Economic Forum - Centre for the Fourth Industrial Revolution, April 2018. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwjV3cyDsM_qAhUPC-wKHcG0AmcQFjAAegQIBRAB&url=http%3A%2F%2Fwww3.weforum.org%2Fdocs%2FWEF_48540_WP_End_of_Innovation_Protecting_Patent_Law.pdf&usg=AOvVaw3i-6YADyd3pImW9NZoQ_zbP.

Yanisky-Ravid, Shlomit, and Xiaoqiong (Jackie) Liu. “When Artificial Intelligence Systems Produce Inventions: The 3A Era and an Alternative Model for Patent Law.” *Cardoza Law Review* 39 (March 1, 2017): 2215–63. https://doi.org/10.2139/ssrn.2931828.

ONLINE SOURCES

- AIVA, https://www.aiva.ai/
- AIVA, FAQs, "I I don't understand the terms of License", https://aiva.crisp.help/en/article/i-dont-understand-the-terms-of-license-1wqvh5v/
- AIPPI, 'Inventorship of inventions made using Artificial Intelligence', Study Guidelines 2020, https://aippi.info/wp-content/uploads/2020/02/Q272-SGL-P-2020-Inventorship-of-inventions-made-using-AI.pdf
- Artificial Inventor Project, http://artificialinventor.com.
- Christies, "Is artificial intelligence set to become art's next medium?" (12 December 2018) https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx
- Deep L, https://www.deepl.com/en/translator
- European Commission, Strategy, Shaping Europe's digital future, Policies, "High Level Expert Group on Artificial Intelligence", https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence .
- European Parliament News, Press Release, "Artificial Intelligence: MEPs presented their vision on forthcoming rules" (12 May 2020 https://www.europarl.europa.eu/news/en/press-room/20200510IPR78803/artificial-intelligence-meps-presented-their-vision-on-forthcoming-rules.
- EPO, Artificial Intelligence, https://www.epo.org/news-events/in-focus/ict/artificial-intelligence.html
- EPO, 'World's five largest patent office's agree on joint task force for emerging technologies and AI', EPO News 2019, https://www.epo.org/news-events/news/2019/20190613a.html.
- EPO, Member states of the European Patent Organisation (accessed 18 June 2020). Online: https://www.epo.org/about-us/foundation/member-states.html.
- EPO, National Applications. Online: https://www.epo.org/applying/national.html.
- EPO, European patents and the grant procedure (2016), at p. 16. Online: https://bit.ly/2HpA3pI
- EPO, 'EPO publishes grounds for its decision to refuse two patent applications naming a machine as inventor' (28 January 2020). Online: https://www.epo.org/news-issues/news/2020/20200128.html.
- FiveIPOffices, Report of the IP5 expert round table on artificial intelligence, EPO, Munich, 31 October 2018, https://www.fiveipoffices.org/material/ai_roundtable_2018_report/ai_roundtable_2018_report .
- FiveIPOffices, First IP5 NET/AI task force meeting takes place in Berlin, Germany (17 January 2020), https://www.fiveipoffices.org/news/20200117
- InferKit, https://inferkit.com/
- InferKit, Docs, Text generation, https://inferkit.com/docs/generation
- Iprova Invent First, http://www.iprova.com/about-us/
- Open AI, GPT-2: 1.5B Release (5 November 2019), https://openai.com/blog/gpt-2-1-5b-release/ .
- OpenAI, Musenet, https://openai.com/blog/musenet/.
- OpenAI, Jukebox, https://openai.com/blog/jukebox/
- Talk to Transformer, https://talktotransformer.com.

- The Next Rembrandt, https://www.nextrembrandt.com/.
- OpenAI, Image GPT (17 June 2020), https://openai.com/blog/image-gpt/
- USPTO, Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation, Federal Register Docket No. PTO-C-2019-0038 (Document Number 2019-23638) 30 October 2019, https://www.federalregister.gov/documents/2019/10/30/2019-23638/request-for-comments-on-intellectual-property-protection-for-artificial-intelligence-innovation.
- Storyai, About – Q&A, https://storyai.botsociety.io/about
- Tate, Look Closer, All About Cubism, *What is cubism and why was it so radical?* https://www.tate.org.uk/art/art-terms/c/cubism/all-about-cubism
- Tate, Marchel Duchamp, Fountain, https://www.tate.org.uk/art/artworks/duchamp-fountain-t07573.
- Tate, Marchel Duchamp, Bicycle wheel, https://www.moma.org/learn/moma_learning/marcel-duchamp-bicycle-wheel-new-york-1951-third-version-after-lost-original-of-1913/.
- The Next Rembrandt, https://www.nextrembrandt.com/.
- VPRO, AI Song Contest, FAQ, https://www.vprobroadcast.com/titles/ai-songcontest.html
- Wikipedia, entry 'Twelve-bar blues', https://en.wikipedia.org/wiki/Twelve-bar_blues
- WIPO, Impact of Artificial Intelligence on IP Policy: Call for Comments, https://www.wipo.int/about-ip/en/artificial_intelligence/call_for_comments/.
- WIPO, Impact of Artificial Intelligence on IP Policy: Call for Comments, https://www.wipo.int/about-ip/en/artificial_intelligence/call_for_comments/
- WIPO, Artificial Intelligence and Intellectual Property, https://www.wipo.int/about-ip/en/artificial_intelligence/
- WIPO, WIPO Technology Trends – Artificial Intelligence, https://www.wipo.int/tech_trends/en/artificial_intelligence/
- WIPO, 'WIPO Conversation on IP and AI', Draft issues paper on intellectual property policy and artificial intelligence', *WIPO/IP/AI/2/GE/20/1*, 13 December 2019, https://www.wipo.int/edocs/mdocs/mdocs/en/wipo_ip_ai_ge_20/wipo_ip_ai_2_ge_20_1.pdf
- WIPO, Submissions received on the Draft Issue Paper, https://www.wipo.int/about-ip/en/artificial_intelligence/policy.html#submissions.
- WIPO, WIPO Conversation on Intellectual Property (IP) and Artificial Intelligence (AI): Second Session (9 to 7 July 2020) https://www.wipo.int/meetings/en/details.jsp?meeting_id=55309.
- WIPO, Summary of the Paris Convention. Online: https://www.wipo.int/treaties/en/ip/paris/summary_paris.html (accessed 15.2.2020)
- WIPO, Summary of the Patent Law Treaty (2000). Online: https://www.wipo.int/treaties/en/ip/plt/summary_plt.html.

ANNEX IX – Declaration on the List of Pre-existing Rights

I, Jacqueline ALLAN, representing the Joint Institute for Innovation Policy (JIIP) ('the Contractor'), party to the contract **Trends and Developments in Artificial Intelligence – Challenges to the IPR Framework – SMART 2018/052** warrant that the results are free of rights or claims from creators or from any third parties for any use the contracting authority may envisage and declare that the results do not contain any pre-existing rights to the results or parts of the results or to pre-existing materials as defined in the above-mentioned contract.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications at: https://op.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

Publications Office
of the European Union