



On the fabrication of sausages, or of Open Government and Private Data

Bart van der Sloot

Researcher at the Institute for Information Law, University of Amsterdam, the Netherlands (b.vandersloot@uva.nl).

Abstract: Governments have become increasingly open and transparent over the last few years. Originally, this trend was largely based on the desire to give citizens access to governmental information, so that state policies and regulatory practices could be controlled and debated. The right to access of governmental data was directly linked with democratic values such as autonomous citizenship, public debate and control on governmental power. In the beginning of this century, emphasis has shifted to a new ground for requiring transparency, namely the re-use of public sector information. Re-use of governmental data by third parties is mostly executed by market parties with commercial interests. The principles of open government and data re-use specifically conflict with intellectual property and privacy rights. This article analyses the tension between open government policies and the protection of personal information from a legal perspective. Finally, it assesses whether and if so, how the two principles can be reconciled.

Keywords: Privacy, PSI, data protection, re-use, transparency, democracy, autonomy

Governmental transparency has been a thorny issue for quite a while. All too often, states have the tendency to be secretive about their policies and actions, fearing critique from citizens and journalists alike, at the same time dreadful that information might fall in the wrong hands. However, the reluctance to open up can only partially be explained by the fear that transparency could give rise to increased critique on and control over governmental actions; first and foremost, states seem to fear a decline of their information monopoly. The Wikileaks-affair and the discomfort openly shown by governments all over the world may be seen as only the latest illustration of this attitude. In their defense, politicians of all stripes and creeds claim that there is no need for more governmental transparency and subsequent control on the legislative process, since it is not so much the process of law and policy making that counts, but the outcome of it. As Otto von Bismarck once remarked: 'Je weniger die Leute wissen, wie Würste und Gesetze gemacht werden, desto besser schlafen sie!' According to him, law making was like fabricating sausages; one may enjoy the outcome of the process, but apatite might be lost if aware of the process and the ingredients used. Moreover, politicians claim that there is no need for full transparency and continuous control, since citizens have the power to vote away their representatives every four years or so.

However, as the democratic principle has become increasingly dominant in the current political paradigm, especially in western societies, states are nevertheless forced to open up. Openness and transparency stimulate the public debate, enable an increasing control on governmental power and facilitates knowledge and autonomy of citizens. Besides access to governmental information, a relatively new principle functions as a catechism for transparency: the re-use of governmental information. Non-profit organizations may re-use governmental data, but it is mostly pursued for commercial purposes. Businesses utilize the commercial potential of the public sector information left unused by governmental institutions.

Generally, open government policies based on access to information or re-use of information may conflict with two legal rights. Firstly, there may rest intellectual property rights on the information contained in public sector documents and databases. Secondly, open government policies may come into conflict with privacy legislation and the protection of personal data. Since

governmental data often refer to or incorporate information about the private lives of citizens, their activities, income, health, criminal activities and so on, enclosing such data may have consequences for the privacy and private lives of citizens. To refer the earlier used metaphor: one of the main ingredients of the sausage of governing is personal information.

This article will focus primarily on the second conflict, assessing the tension between open government policies and the protection of privacy in the European framework. First, it will provide some general outlines of the right to access to governmental data and it will proceed by analyzing the most important aspects of the right to re-use of public sector information. Thirdly, the study will shortly describe the possible conflict of transparency policies with intellectual property rights. The major focus of this study lays, fourthly, on the analysis of the tension between open government practices and the right to privacy. It will be shown that there are several obstacles with regard to the access to and the re-use of public sector information. Finally, this paper proceeds by presenting currently proposed solutions to these problems and presents a new one, namely to introduce citizens' personal privacy settings with which citizens register their preferences regarding re-use of public sector information containing their private data.

1. Access to governmental information

Article 10 of the European Convention on Human Rights, containing the right to freedom of speech, holds that everyone has the right to freedom of expression and that this right shall include the freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. Other than equivalent articles in other international documents regarding fundamental rights, it does not contain a right to information. For example, in article 19 of the International Covenant on Civil and Political Rights the right to freedom of expression includes the freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice. Likewise, article 13 of the American Convention on Human Rights holds that the right to freedom of thought and expression includes the freedom to seek, receive, and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing, in print, in the form of art, or through any other medium of one's choice.

For a long time, the European Court for Human Rights, the ultimate interpreter of the European Convention, has stuck to the text rather strict, only granting a right to information where this derived from the respect of other human rights mentioned in the convention, for example the right to a fair process. However, since recently, this has changed. The European Union adopted a Charter of Fundamental Rights, in which article 11 holds that the right to freedom of expression shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. Also, since 2009, the European Court for Human Rights has repeatedly accepted a right to information by citizens with regard to their national states, linked to personal and social interests.¹

The most important instruments for obtaining access to governmental information however remain Member States' national legislation on this point.² Although roughly speaking, northern and western European countries seem to provide for a more encompassing right to information than the southern and eastern countries, all countries increasingly grant their citizens this right. What is important for the current research is that the laws and case law regarding access to governmental information have in common a reference to fundamental rights of the citizen and to democratic principles relating to control on governmental power.

¹ ECHR 14 April 2009, *Társaság a Szabadságjogokért vs Hungary*, appl. no. 37374/05. ECHR 26 May 2009, *Kenedi vs Hungary*, appl. no. 31475/05.

² See however also: Regulation (EC) No 1049/2001 of the European Parliament and the Council. of 30 May 2001 regarding public access to European Parliament, Council and Commission documents.

2. Re-use of governmental information

Standard text A new paradigm for governmental transparency has emerged. Already in the year 2000, the total value of the European Public Sector Information (PSI) was estimated to be around 68 billion euro annually.³ To ensure that at least a part of this potential is utilized, the re-use of public sector information is encouraged in Europe through the PSI Directive of 2003.⁴ Although it does not entail any obligation for public sector organizations to disseminate such information, most Member States have incorporated such a principle in their national laws. It is important to distinguish between the right of access to governmental information and the right to re-use.⁵ While the rationale behind rules ensuring access to governmental information are mainly linked to democracy and control on governmental power, re-use is primarily important for commercial interests.⁶

With regard to the re-use of public sector information, three concepts are important: (1) it must regard re-use, (2) of information (3) in the hands of the public sector. All three concepts are defined broadly in the PSI Directive. A public sector body means the state, regional or local authority, bodies governed by public law and associations formed by one or several such authorities or one or several such bodies governed by public law. If they are not financed by a public authority, then they may still qualify as a body governed by public law if they are subject to supervision by those bodies.⁷ A public sector document refers to any content or any part of such content whatever its medium.⁸ Finally, 're-use' is characterized as the use by persons or legal entities of documents held by public sector bodies, for commercial or non-commercial purposes other than the initial purpose within the public task for which the documents were produced.⁹

3. Intellectual property

Both regimes with regard to open government may come into conflict with two legal rights. The first is the right to respect for intellectual property. The second is the right to respect for privacy. Although the second right will be the main focus of this article, this paragraph will shortly outline some of the most important problems with regard to the right to respect for intellectual property. Intellectual property refers to copyright on material and to related rights, which often grant an individual or a company the right to restrict the use of material, demand financial compensation when the material is used by third parties and to set requirements on who, how, why and when the material is used.¹⁰ The three most important legal instruments with regard to intellectual property in the European Union are the Berne Convention for the Protection of Literary and Artistic Works (the Berne Convention), the Agreement on Trade-Related Aspects of Intellectual Property Rights (the TRIPS Agreement) and the Copyright Directive.

Intellectual property is a very broad concept and the right may refer to every production in the literary, scientific and artistic domain, whatever may be the mode or form of its expression, such as

³ European Commission, 'Commercial exploitation of Europe's public sector information', 20 September 2000, p. 6.

⁴ Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information. (PSI Directive).

⁵ Pas & Vuyst (2004), p. 2.

⁶ Goens (2010), p. 399.

⁷ Article 2 § 1 & 2 § 2 PSI Directive.

⁸ Article 2 § 3 PSI Directive.

⁹ Article 2 § 4 PSI Directive.

¹⁰ See also recital 22 of the PSI Directive.

books and other writings, lectures, dramatico-musical works and choreographic, musical compositions, cinematographic works, works of drawing, painting, architecture, sculpture, engraving and photographic works, illustrations, maps, plans, sketches and three-dimensional works relative to geography, topography, architecture or science.¹¹

Since the scope is so broad, governmental information often contains material on which rests the intellectual property of citizens and organizations. This can refer to works that have been deposited with public agencies as a requirement for legal protection, it may refer to architectural information of governmental buildings, to the design of public parks etc. This being so, open government policies may come into conflict with the legal right to respect for intellectual property and the right to restrict the use of material and demand financial compensation for using it. This hurdle for governmental transparency policies is however not the core issue of this article and thus will remain largely untouched. What will be the main focus of this study is the possible conflict of open government policies with the right to privacy and data protection.

4. Privacy and data protection

European Union legislation lays much emphasis on the protection of privacy and the closely related right to data protection. To understand the subtle difference between the two, reference can be made to the Charter of Fundamental Rights of the European Union. Article 7 holds that everyone has the right to respect for his or her private and family life, home and communications. This is usually referred to as the right to privacy. Its equivalent is article 8 of the European Convention on Human Rights, which grants a person the right to respect for his private and family life, his home and his correspondence. Article 8 Charter of Fundamental Rights of the European Union holds that everyone has the right to the protection of personal data concerning him or her, that such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law, that everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified and that compliance with these rules shall be subject to control by an independent authority. This is called the right to data protection. Since there is relatively little case law with regard to the conflict of open government policies with the right to privacy, this article shall mainly refer to the right to data protection. This right is most prominently protected by the European Data Protection Directive.¹²

Since public sector information contains personal data, the distribution of the information for either access or re-use may trigger the applicability of the Data Protection Directive. For example the PSI Directive, but most laws regarding access to information do so too, makes explicit reference to the Data Protection Directive when it states that it will leave intact and in no way affects the level of protection of individuals with regard to the processing of personal data under the provisions of Community and national law and in particular does not alter the obligations and rights set out in Data Protection Directive.¹³ National legislators are required to implement and apply the PSI Directive in full compliance with the principles relating to the protection of personal data in accordance with Data Protection Directive.¹⁴ Finally, the PSI Directive does not contain a definition of personal data, but refers to that of the Data Protection Directive.¹⁵ Both the Commission's Green

¹¹ Article 2 Berne Convention.

¹² Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. (Data Protection Directive).

¹³ Article 1 § 4 PSI Directive.

¹⁴ Recital 21 PSI Directive.

¹⁵ Article 2 § 5 PSI Directive.

Paper¹⁶ and the recent communication on the review of the PSI Directive¹⁷ confirm the full applicability of the Data Protection Directive. This is also the case with the distribution of governmental data for the purpose of access and transparency without re-use.

Hence, it needs to be determined to what extent the Data Protection Directive applies to the access to and the re-use of public sector information. If it does apply, three major categories of obligations will need to be taken into account. First, there must exist a legitimate purpose for processing personal data. Second, when processing personal data, the safeguards prescribed by law must be observed. Thirdly, the rights of the data subject in relation to the transparency principle must be observed.

4.1. Applicability

The applicability of the directive is triggered when (1) "personal data" are (2) "processed" under the authority of the (3) "controller" of the personal data on the (4) territory of the European Community. With regard to the first requirement, personal data are defined under the directive as any information relating to an identified or identifiable natural person (the data subject). An identifiable person is someone who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.¹⁸ The Working Party 29,¹⁹ the advisory institution regarding privacy and data protection in the European Union,²⁰ has elaborated on four elements of personal data: 'any information', 'relating to', 'an identified or identifiable' and 'natural person'.²¹ Both objective and subjective information, i.e. facts and opinions, would fall under the concept of 'information'; the form in which it is kept is irrelevant. Information may relate to a person either qua content, if information refers to a person, qua purpose, if the information is used to evaluate or influence personal behavior, or qua result, if the consequence of data processing is that a person might be treated or looked upon differently.²²

Personal data may either be directly identifiable, such as a name, or indirectly, such as a telephone number or a combination of non-directly identifiable information, such as age and address. 'Even ancillary information, such as "the man wearing a black suit" may identify someone out of the passers-by standing at a traffic light.'²³ To determine whether a person is identifiable, all means likely and reasonably to be used either by the controller, or by any other person to which the information is disseminated, to identify a person should be taken into account.²⁴ This means that it is not necessary that a person is *de facto* identified by someone, but that this is reasonably possible. Furthermore, the criterion is not that the controller of the information should be able to identify a person, but that third parties, being either persons or companies that have access to the information, are able to identify individuals.

¹⁶ European Commission, 'Public sector information: a key resource for Europe. Green Paper on public sector information in the information society', COM(1998) 585, p. 16.

¹⁷ Commission of the European communities, 'Communication from the commission to the European Parliament, the Council, The European Economic and Social committee and the Committee of the Regions. Re-use of Public Sector Information – Review of Directive 2003/98/EC –', SEC(2009) 597, p. 8. C. Corbin, 'EC Communication on the PSI PSI Directive: PSI re-use stakeholder reaction', European PSI Platform – Topic Report No .3.

¹⁸ Article 2 § A Data Protection Directive

¹⁹ <http://ec.europa.eu/justice/policies/privacy/workinggroup/index_en.htm>.

²⁰ Recital 65 & article 29-30 Data Protection Directive.

²¹ Article 29 Working Party, 'Opinion 4/2007 on the concept of personal data (WP 136)', Brussels, 20 June 2007.

²² WP 136, p. 10.

²³ WP 136, p. 13.

²⁴ Recital 26 Data Protection Directive

Given the general scope of the definition of personal data, many documents will contain personal data.²⁵ Under the Data Protection Directive, there is a special category of so called sensitive data, which are personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership and data concerning health or sex life.²⁶ It is not uncommon for public sector bodies to hold such information.²⁷

'States maintain records spanning an individual's life from birth to death, including records of births, marriages, divorces, professional licenses, voting information, worker's compensation, personnel files (for public employees), property ownership, arrests, victims of crime, criminal and civil court proceedings, and scores of other information. Federal agencies maintain records pertaining to immigration, bankruptcy, social security, military personnel, and so on. These records contain personal information including a person's physical description (age, photograph, height, weight, eye color); race, nationality, and gender; family life (children, marital history, divorces, and even intimate details about one's marital relationship); residence, location, and contact information (address, telephone number, value and type of property owned, description of one's home); political activity (political party affiliation, contributions to political groups, frequency of voting); financial condition (bankruptcies, financial information, salary, debts); employment (place of employment, job position, salary, sick leave); criminal history (arrests, convictions, traffic citations); health and medical condition (doctors' reports, psychiatrists' notes, drug prescriptions, diseases and other disorders); and identifying information (mother's maiden name, Social Security number).²⁸ In conclusion, much of the public sector information will contain both ordinary and sensitive personal data. Therefore, the first criterion will be satisfied.

Secondly, for the Data Protection Directive to apply, the personal data must be processed. The concept of data processing is defined very broadly as any operation or set of operations which is performed upon personal data, whether or not by automatic means, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction.²⁹ In short, almost anything that can be done with personal data falls within this all-encompassing definition.³⁰ The Working Party observes in its opinion on the re-use of public sector information and the protection of personal data, that '[] the disclosure to third parties of personal data collected and held by public sector bodies is to be considered as processing of personal data, given that the definition of processing includes a *disclosure by transmission* with the consequence that the material conditions that govern the processing of personal data have to be observed.'³¹ This is also applicable on the disclosure of information without the purpose of re-use. Thus, there are three stages of processing the personal data. First when the data is originally gathered by the public sector organization, secondly when the data is transferred from the public sector organization to a third party and thirdly when that third party uses the gained data for its own purpose.

Thirdly, the obligations under the directive apply to the controller of the personal data. The controller is defined as the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data.³² On him lie all the obligations under the directive.³³ The public sector organization

²⁵ Judgment of the European Court of Justice C-101/2001 of 06.11.2003 (Lindqvist), § 27.

²⁶ Recital 33 & article 8 Data Protection Directive.

²⁷ WP 83, p. 4. Pas & Vuyst (2004), p. 1.

²⁸ Solove (2002), p. 1139.

²⁹ Article 2 § B Data Protection Directive.

³⁰ Goens (2010), p. 402.

³¹ Article 29 Working Party, 'Opinion 7/2003 on the re-use of public sector information and the protection of personal data - Striking the balance - (WP 83)', Brussels, 12 December 2003, p. 4.

³² Article 2 § D Data Protection Directive.

disseminating the information will be qualified as the controller at the time of gathering, analyzing, using and disseminating the information. The third party will also qualify as the controller of the personal data when gaining the public sector information. Consequently, it too has to fulfill all the obligations under the directive.³⁴

Fourthly and finally, the data protection rules apply when (1) processing is carried out in the context of the activities of an establishment of the controller on the territory of the Member State or (2) when the controller is not established on Community territory and for purposes of processing personal data makes use of equipment situated on the territory of a Member State, unless such equipment is used only for purposes of transit through the territory of the Community.³⁵ There is no doubt that public sector organizations would fall under the first category.³⁶ To fall under the first category, the third parties must be or have an establishment on the territory of a Member State, the latter implies the effective and real exercise of the business activity.³⁷ The legal form of such an establishment is not the determining factor in this respect. When a single controller is established on the territory of several Member States, he must ensure that each of the establishments fulfils the obligations imposed by the national law applicable to its activities.³⁸ Furthermore, the third party should process the data in the context of his everyday business activities.³⁹ If the first category would not apply, the second one would if the controller is not established on Community territory, but makes use of equipment, automated or otherwise, situated on the territory of the Member State, unless the equipment is used only for purposes of transit through Community territory, which would seldom be the case.⁴⁰

The purpose of the directive is to guarantee data subjects an adequate level of protection, wherever the controller is established.⁴¹ If the third party would not fall under one of the two categories mentioned above, then the public sector organization as the controller is under special obligations to ensure that the rules under the Directive are respected by the third party. For example, if the third country the data is transferred to does not have a proper data protection regime, it may not proceed with the dissemination.⁴²

In conclusion, given the fact the even indirectly identifiable information and a phrase like the man wearing a black suit may qualify as personal data, most public sector information will contain personal data, especially when different data sets are seen in relation to each other. The data is processed at three stages, namely at the moment the data is initially gathered and used by the public sector organization, at the moment it disseminates the information to third parties and when the third party is using the information for its own purposes. Both the governmental authority and the third party will have to fulfill the conditions under the Data Protection Directive, as they both qualify as the controller of the data. Only seldom will third parties not fall under the territorial scope of the directive and if so, the governmental organization will be under special obligations to guarantee that the rules under the directive are respected by the third party.

The following paragraphs will assess three major categories of obligations with regard to data processing: first, the required legitimate purpose, secondly, respecting the safeguards spelled out

³³ Article 29 Working Party, 'Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169)', Brussels, 16 February 2010.

³⁴ WP 169, p. 33.

³⁵ Article 4 Data Protection Directive. Sub b refers primarily to embassies and will be mostly irrelevant with regard to the question of re-use of public sector information.

³⁶ Article 29 Working Party, 'Opinion 8/2010 on applicable law (WP 179)', Brussels, 16 December 2010.

³⁷ European Court of Justice, 4 July 1985, Case C-168/84, (*Berkholz*).

³⁸ Recital 19 Data Protection Directive.

³⁹ WP 179, p. 14.

⁴⁰ WP 179, p. 23.

⁴¹ Recital 18 & 20 Data Protection Directive.

⁴² Article 25 Data Protection Directive.

by the directive and finally, respecting the rights of the data subject in connection with the transparency principle.

4.2. Legitimate purpose

The directive holds that non-sensitive personal data may only be processed on a legitimate basis. The directive mentions six ways to do so. A data controller may process personal data if the data subject has unambiguously given his consent, when it is necessary for the performance of a contract, a legal obligation, a public task carried out in the public interest or to protect the vital interest of the data subject. Finally, data processing is allowed when the interests served by the processing weigh higher than the interests of the data subject.⁴³ The public sector organization will usually have gathered information about citizens in the course of a legal obligation or when fulfilling a task carried out in the public interest.⁴⁴ Therefore, it has a legitimate purpose for processing large quantities of personal data, even more so, since data processing in the light of public security activities are excluded from the scope of the directive.⁴⁵

As a controller, the third party receiving the public sector information must also fulfill one of the six circumstances mentioned in the directive. Usually, access to and re-use of the information will not be necessary for the performance of a contract, a legal obligation, a public task carried out in the public interest or to protect the vital interest of the data subject. To clarify, the access to governmental data for purposes of democratic control is only seldom legitimized by the public interest since the directive refers to a 'task', which entails a legal obligation. Getting the consent of the different persons of whom personal data is contained in the information may be a laborious process, since the consent must be given freely, on specific terms and must be given on an informed basis.⁴⁶ Even if a third party was willing to undergo this process, it would be questionable if it would be able to contact each and every data subject of whom information is contained in the obtained data.

The most likely legitimate purpose to apply is the so called balancing provision, with which the interest of the controller or the third party to which the data is disseminated is balanced with the interest of the data subject, especially with regard to the respect for his fundamental rights to privacy and data protection.⁴⁷ Both the right to privacy and the right to data protection are fundamental and core human rights, contained in among others the European Convention of Human Rights and the Charter of Fundamental Rights of the European Union. If re-use is requested by third parties with an aim at profit or serving business activities, there will be no fundamental right at stake and the re-use would thus not be legitimate. Only if another fundamental right is served by the re-use of public sector information containing personal data, most commonly the right to freedom of speech, will there be a situation in which the two interests must be balanced. This balance may be different in the case of access to governmental data, serving democratic principles and the right to freedom of speech. In any case, the balance must be struck on a case-by-case basis.⁴⁸

A further complicating fact is that the Directive provides for a separate regime with regard to the processing of sensitive data, which is prohibited unless the data subject has given his explicit consent, if the process is necessary to comply with employment law or to protect the vital interests

⁴³ Recital 30-32 & article 7 Data Protection Directive.

⁴⁴ WP 83, p. 5. Recital 9 PSI Directive.

⁴⁵ Article 3 Data Protection Directive.

⁴⁶ Article 2 § H Data Protection Directive.

⁴⁷ Recital 30 Data Protection Directive.

⁴⁸ WP 83, p. 5.

of the data subject. The processing of sensitive data is also legitimate if carried out in the course of legitimate activities with appropriate guarantees by a non profit organization with a political, philosophical, religious or trade-union aim, on the condition that the processing relates solely to the members of the body or to persons who have regular contact with it in connection with its purposes and that the data are not disclosed to a third party without the consent of the data subjects. Lastly, processing of sensitive data is allowed when they are manifestly made public by the data subject.⁴⁹

The governmental organizations gathering sensitive data will not have to comply with these rules if they are processed in relation to activities concerning public security in the broad sense of the term⁵⁰ or if processing relates to health care issues or any other issues relating to the public interest as laid down in law.⁵¹ Commonly, third parties will however not have any obligations under the national employment law, nor have the vital interest of the data subject at hart. The data is not made public by the data subject nor will it have given its consent to the re-use of his personal data.⁵² Finally, the access to and re-use of public sector information will usually not be done by a non profit organization processing data that relates solely to the members of the body or to persons who have regular contact with it in connection with its purposes.⁵³ Thus it is questionable whether the third party would have a legitimate purpose when processing sensitive personal data contained in the public sector information. Only if the processing of personal data is carried out solely for journalistic purposes or the purpose of artistic or literary expression, controllers may deviate from the rules regarding the required legitimate purpose.⁵⁴ A journalistic purpose is not restricted to activities by journalistic media, but may also be pursued by individuals or companies. Sometimes, revealing certain governmental policies and possible missteps may well qualify as journalistic even if done by ordinary citizens and businesses. For such purposes, there is no need for a legitimate purpose for the processing of personal data as listed in the directive.

4.3. Safeguards

Next to the obligation with regard to the legitimate purpose for data processing, the directive spells out several safeguards.⁵⁵ First of all, processed personal data must be accurate and kept up to date; every reasonable step must be taken to ensure that data which are inaccurate or incomplete, having regard to the purposes for which they were collected, are erased or rectified.⁵⁶ This means that the governmental organization possessing the personal data of the subjects must ensure that the information is kept up to date and corrections are made where necessary. Moreover, they have to inform a third party to whom they have distributed public sector information containing personal data when they are aware of the fact that such data is inaccurate or outdated. The third parties as data controllers must also fulfil these obligations.

Furthermore the directive encompasses certain so called data minimisation principles; personal data must be adequate, relevant and not excessive in relation to the purposes for which they are

⁴⁹ Recital 33-36 & article 8 § A Data Protection Directive

⁵⁰ Article 3 Data Protection Directive.

⁵¹ Article 8 § B & § C Data Protection Directive.

⁵² The term 'explicit' consent, in stead of 'unambiguous' consent which is used in relation to the processing of ordinary data, is meant as an aggravating condition. Since it was concluded that this condition would not provide processing legitimacy with regard to ordinary data, it will most certainly not do so with regard to sensitive data.

⁵³ The directive will not apply to the processing of personal data is done by a natural person in the course of a purely personal or household activity. Article 3 Data Protection Directive.

⁵⁴ Recital 17, 37 & article 9 Data Protection Directive.

⁵⁵ Recital 28 Data Protection Directive.

⁵⁶ Article 6 § 1 sub D Data Protection Directive.

collected⁵⁷ and may be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected.⁵⁸ This means that the governmental organizations as well as the third parties will need to make sure that the gathered data is necessary, proportional and the subsidiarity principle is respected. Moreover, if governmental organizations disseminate information to third parties, they are under the obligation to make sure that the third parties will fulfill their obligations in this respect.

Moreover, personal data may only be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes.⁵⁹ This means that both the governmental organization and the third party requesting public sector information for either access or re-use must have a specific, explicit and legitimate purpose for processing personal data. Furthermore, the governmental organization must check whether the third party fulfills its obligation when it disseminates the information and must see to it that the purpose for processing by the third party is not incompatible with his own reasons for processing the data.⁶⁰ What incompatibility means precisely is not apparent from the directive.⁶¹ It is however clear that the prohibition creates an enormous problem for re-use of public sector information, since normally, the data is gathered by the government to serve legal obligations and the public interest and the re-using party will not. The Working Party 29 furthermore emphasizes: 'If personal data are to be re-used for commercial purposes, this secondary purpose may be considered as incompatible and thus the information not be disclosed.'⁶² Therefore, in general terms, only when third parties' goals with regard to processing relate to the original purpose or when the processing of data is executed for historical, statistical or scientific purposes⁶³ will they fulfil their obligations in this respect. With regard to access to governmental information in relation achieving and safeguarding democratic ideals this may be different, such this purpose is not incompatible with the gathering of information by governmental organisation. Both see to fulfilment of governmental tasks. Furthermore, further processing of data for historical, statistical or scientific purposes shall not be considered as incompatible. This is especially relevant in the case of access to governmental information, without re-use purposes.

Finally the directive holds that the controller must implement appropriate technical and organizational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access.⁶⁴ This means that the public sector body disseminating the data must ensure that the data is not used for unlawful purposes or processed in unjust ways. This would be especially a problem with regard to governments that have launched a website whereupon they publish governmental documents, since these documents and the information contained in them are then out of their control. With regard to distribution of data to individual third parties, this may be different since the governmental organization can and must check for what purposes the third party would process the data and how. Of course third parties are also under the obligation to fulfill their obligations with regard to the security principle.

4.4. Transparency & Rights

⁵⁷ Article 6 § 1 sub C Data Protection Directive.

⁵⁸ Article 6 § 1 sub E Data Protection Directive.

⁵⁹ Article 6 § 1 sub B Data Protection Directive.

⁶⁰ WP 83, p. 7-8.

⁶¹ Kuner (2007), p. 100.

⁶² WP 83, p. 9.

⁶³ Recital 29 Data Protection Directive.

⁶⁴ Recital 46 & article 17 Data Protection Directive.

Account should also be taken of the transparency principle and the rights of the data subject. The transparency principle requires that in cases of collection of data from the data subject, the controller must provide the data subject with at least his identity, the purposes of the processing for which the data are intended and the recipients or categories of recipients of the data.⁶⁵ This means that the governmental organization must take all reasonable steps to ensure that every individual data subject is informed of the fact that his personal data is distributed to third parties. It is difficult to see how a governmental organization disseminating public sector information to third parties would see to it that every data subject is adequately informed of this matter. Exceptions exist when processing is executed in relation to security issues, important economic or financial interest of a Member State or when the protection of the data subject or of the rights and freedoms of others prevail.⁶⁶ Freedoms of others may also include the freedom of third parties. However, this exception is not primarily created to restrict the right to data protection, but to further ensure it, as the directive gives as example that Member States may specify that access to medical data may be obtained only through a health professional.⁶⁷ Thus, it is unlikely that Member States may restrict the obligation to transparency to protect the interests of the third parties. It therefore remains difficult to see how the governmental organizations will fulfill their obligation to transparency. Furthermore, the third parties are under a similar obligation.⁶⁸ This means that the third party receiving the public sector information containing personal data must make sure that it informs every data subject of his identity, the purposes for processing etc., except when processing the data with a journalistic purpose. Furthermore, data controllers are lifted from this obligation if data is processed for statistical, historical or scientific research, when the provision of such information proves impossible or would involve a disproportionate effort.⁶⁹ It could be so that third parties requesting public sector information for access or re-use purposes could fall under this exemption, since transparency would mean an unbearable burden.

Furthermore, every data subject has the right to access, which means that it has the right to obtain from the controller a confirmation as to whether or not data relating to him are being processed.⁷⁰ Moreover, the data subject has the right to demand the rectification, erasure or blocking of data the processing of which does not comply with the provisions of the Data Protection Directive, in particular because of the incomplete or inaccurate nature of the data. Subsequently, he is entitled to a notification to third parties to whom the data have been disclosed of any rectification, erasure or blocking, unless this proves impossible or involves a disproportionate effort.⁷¹ This means that both the public sector organization and the third party in their capacities as controllers must fulfill the request of data subjects to rectify, correct or block data. Furthermore, the governmental authority would need to notify all third parties it has disseminated the personal data to of such requests, since they too would need to comply with such a request, having possession of the personal data as well.

Finally, the data subject has a right to object, which means that at least in the cases that the legitimization for processing is found in serving the public interest or in the balance of different interest, which as explained earlier would be the ground most likely to be invoked in the case of the access to or re-use of public sector information, the data subject has the right to object at any time on compelling legitimate grounds relating to his particular situation to the processing of data relating to him. Where there is a justified objection, the processing instigated by the controller may no longer involve those data. Moreover, the directive holds that the data subject has the right to object to the processing of personal data relating to him which the controller anticipates being

⁶⁵ Article 10 Data Protection Directive.

⁶⁶ Recital 43 & article 13 Data Protection Directive.

⁶⁷ Recital 42 Data Protection Directive.

⁶⁸ Article 11 Data Protection Directive.

⁶⁹ Recital 39-40 & article 11 § 2 Data Protection Directive.

⁷⁰ Recital 41 & article 12 § A Data Protection Directive

⁷¹ Article 12 § B & § C Data Protection Directive.

processed for the purposes of direct marketing, to be informed before personal data are disclosed for the first time to third parties or used on their behalf for the purposes of direct marketing and to be expressly offered the right to object free of charge to such disclosures or uses.⁷² This right could mean an obstacle especially for re-use of public sector information by third parties that have found the legitimacy of processing in the balance of interests.

5. New approaches

There are several legal obstacles with regard to open government policies in the light of the Data Protection Directive. What is apparent is that the more classical basis for such policy, linked to democracy and constitutional values has little trouble to live up to the requirements of that directive. This is different with the case of re-use of governmental information for commercial interests. The governmental organizations will have difficulty to see to it that data is not further processed in a way inconsistent with the original purpose, abide the security obligation and respect the transparency principle. Likewise, re-using third parties will have trouble respecting the transparency principle and the rights of the data subject, but maybe most importantly, they will have difficulty satisfying their obligation to have a legitimate purpose with regard to the data processing. Since the purposes of the initial data processing and the purposes for which the data are further processed, the possibilities to re-use public sector information are severely limited. This paper proceeds by shortly presenting two radical solutions to these problems, meaning either prohibiting re-use of public sector information or ignoring the data protection principles when doing so. Then the possibility of data anonymization is tested. Finally, a new solution is presented, namely to introduce personal privacy settings regarding the re-use of citizens' personal data by third parties.

5.1. Radical solutions

On the one hand, as most of the governmental documents will contain personal data and as it will be difficult for both the governmental authority and the re-using party to abide to every obligation spelled out in the Data Protection Directive, a total prohibition of the re-use of public sector information might be the most feasible solution. However, this solution might not be the most satisfying one, because it would entail going back to square one. It would leave the economical potential of the European public sector information unutilized.

On the other end of the spectrum, one could opt for a total release of public sector information, without being hindered by privacy and data protection concerns, somewhat like the less strict American model. While primarily focusing on prosperity and profit by the privacy sector, this model is based on open and unrestricted access to public sector information at no more than the cost of search and duplication.⁷³ In contrast, the current European model takes into account fundamental rights to privacy and data protection, which play a much less important role in the United States. 'One should bear in mind that in Europe, contrary to the U.S., access and use or re-use are considered as conceptually different activities. In Europe access is considered as a matter of human rights, while use and re-use as an activity based mainly on the principles of competition and intellectual property laws. In the U.S. the commercialization of public sector information is not seen as a separate issue. Access and re-use are considered to be part of the same right. Furthermore, while in the U.S. cultural traditions favor commercialization of public information and commercialization is a well-established practice, the question raised there is whether commercial

⁷² Article 14 Data Protection Directive.

⁷³ Aichholzer (2004), p. 526.

activities by the public sector are appropriate and if public sector bodies should be allowed to commercialize their own information. The opposite presumption is the norm in Europe, where the question is rather if commercial exploitation of public information can be justified at all. In brief, one could state that the United States federal system on accessing government-generated information is a system that basically assumes all government held data to be public asset, by which, as a consequence, any person can access it *and* use it.⁷⁴ Waiving away rights to privacy and data protection is not only in sharp contrast with the European legal tradition and culture, it could also have consequences for the autonomy of citizens, undermine the democratic process and stimulate criminal activities, such as misuse of personal data and identity theft.⁷⁵ This being the case, both radical solutions do not seem very satisfying.

5.2. Anonymisation

A third solution for the tension between the re-use of public sector information and the rights to privacy and data protection may be found in anonymization techniques. The Data protection Directive holds that the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable⁷⁶ and the Working Party 29 holds that '[w]ith a view to avoiding the disclosure of personal data in the first place, such should be excluded where the purpose of the re-use can be fulfilled with the disclosure of personal data rendered anonymous in such a way that the data subject is no longer identifiable.'⁷⁷

This option would perhaps be most satisfying if feasible. It would on the one hand ensure that public sector information can be re-used and on the other protect the privacy of the citizens. It is however questionable whether this project could succeed. First of all, the scope of the concept of personal data under the Data Protection Directive is all-encompassing: it not only refers to sensitive data, but also to ordinary data, not only to information by which a person is identified, but also to data by which a person could reasonably be identified by anyone obtaining the data, not only to direct identifiable information, but also to indirect identifiable information, etc. Since even a phrase like 'the man wearing a black suit' may identify someone, this means that a total anonymization process with regard to public sector information would be almost impossible.⁷⁸ Moreover, even if this process would be fully carried through, the remaining value of the public sector information would be close to nil: 'Data can be either useful or perfectly anonymous but never both.'⁷⁹ Furthermore, it is important to stress that although a specific data set might not reasonably be used to identify a person, this might be different if various datasets are combined and integrated and computer programs and innovative algorithms are used to distil personal profiles from them. Finally, although new and more effective techniques for anonymisation succeed each other rapidly, the same might be said for re-identification techniques wherewith it is possible to undo a large part of the attempt to anonymise the data.⁸⁰ In conclusion, a successful anonymization process would be Sisyphean task and even if successfully deployed, the value of the public sector information would decrease drastically.

⁷⁴ Pas (2004), p. 9 (references deleted). Pas & Vuyst (2004), p. 2

⁷⁵ For the more fundamental value of privacy see: Roessler (2005).

⁷⁶ Recital 26 Data Protection Directive.

⁷⁷ Wp 83, p.4.

⁷⁸ <<http://patientprivacyrights.org/wp-content/uploads/2010/08/The-Case-for-Informed-Consent.pdf>>.

⁷⁹ Ohm (2010), p. 1704.

⁸⁰ See among others: Malin & Sweeney (2004). Hay & Miklau & Jensen & Towsley & Weis (2008).

5.3. Personal privacy settings

A new solution, proposed in this paper, would be to let everyone register their own privacy settings with the government, for example by registering an account on the website through which the government distributes public sector information. Since consent under the Data Protection Directive means any freely given specific and informed, unambiguous or explicit indication of one's wish, this would entail that the citizen must have explicitly filled out a (digital) form, to register for the re-use of his personal data; the default setting may not be an opt-out model. Secondly, consent needs to be specific. This means that the data subject must have consented to a particular re-use. This requirement could be taken into account by letting the data subject choose to whom he would like his personal data distributed: fellow citizens, companies, non profit organizations, other governments etc. This also means that he may distinguish between purposes for which his personal data is re-used, for example between commercial and non commercial purposes. The data subject may be offered the opportunity to distinguish between territories he wants his data to be distributed to, for example indicating that only third parties that have an established in his country of origin may use his data, that the data should remain in the European Union or that if the data is distributed to third countries, which ones he trusts. Maybe most importantly, he must be given the opportunity to select what kind of information he would like third parties to use. For example, he might distinguish between indirect and direct identifiable information and between ordinary and sensitive personal data. It might be feasible to prohibit the re-use of sensitive personal data, since this could hinder a person's privacy and autonomy to an intolerable extent. It would also be preferable if the citizens might choose from which database they would like their personal data to be distributed.

Thirdly, the data subject needs to be informed. This means that the public sector must distribute information about what data of the citizen is kept by the governmental organization. It must also register the businesses that have requested access to certain data. Thus the citizens may inform themselves of which businesses have which data for what purposes. A notification system could be linked to a mail system so that the citizen is notified every time that a party has started to re-use his personal data. As a safety catch, it must be possible for a data subject to request the third party to stop re-using his personal data, even though the re-use would fall between the parameters set by the citizen himself. A citizen could also be required to reconfirm his privacy settings periodically, so as to ensure that the settings continue reflecting his will.⁸¹

For this project to succeed, it would be necessary to determine which documents contain what information of whom. With regard to names, dates of birth, home addresses, criminal activities and health related information this might be relatively easy; geographical information might be linked to the home address in certain periods of time, working space etc. Statistical information might be linked to groups in which citizens are classified; for example, if a document contains information on elderly people above the age of 65, then every citizen born before 1946 is referred to. With regard to other, less easily identifiable information, it would be the government's task to make sure that the documents are scanned and the painful process of indirect identification is conducted. The costs for this process should be paid by the re-using parties.

This also relates to the last point, namely that of profit-sharing.⁸² The problem with the described model would be that there is no incentive for citizens to opt-in to the re-use scheme. This problem could be overcome by granting citizens a percentage of the profit made by the third parties through the re-use of their personal data or a lump sum set by the government, depending on type of data and the period and purpose of processing. This creates an incentive to register and ensures a

⁸¹ See in analogy: Working Party 29, 'Opinion 2/2010 on online behavioural advertising (WP 171)', Brussels, 22 June 2010, p. 18.

⁸² There already have been lawsuits in which citizens got compensation money. *Kehoe v Fidelity Federal Bank & Trust*, (4 S.Ct. 1612 (Mem.)), 21 F.3d 1209 126, 2005. *Burdon* (2009).

tailor-made model for every individual citizen. Citizens that are not that interested with their personal data may offer their every personal data for re-use by whatever company, located in whatever jurisdiction for whatever purpose. Doing so, it could be possible that in the future they would make a reasonable profit with their data. Citizens that have more and stronger privacy concerns might choose not to register their personal privacy settings or allow re-use only to a limited extent.⁸³

6. Conclusion

Open government policies may either relate to serving democratic values such as autonomy, public debate and control on state power, but is increasingly based on commercial interests. This research has shown that the second basis is problematic in the light of privacy legislation. As governing is like the making of a sausage and one of the main ingredients is private data, when third parties use the ingredients for a different recipe, personal information is ill protected since the governmental safeguards no longer apply. Since the PSI Directive holds the Data Protection Directive to be fully applicable on the re-use of public sector information, both the governmental organizations and the re-using parties are under a number of obligations, a good part of which they will have difficulty to fulfill. Good solutions are very few and far between: both a total prohibition and allowing re-use without conditions are unsatisfying because they do not take into account the economic potential of the information respectively the value of data protection and privacy. Anonymization would be the best solution, since it would diminish the privacy-aspects of re-use and would still ensure that the documents might be re-used. However, since the concept of personal data is so big, this might be a Sisyphean task. Even if the governmental organizations would succeed, the data would presumably have lost most of its value. A new solution is suggested in this paper, namely to let every citizen register its own personal privacy setting regarding re-use of public sector information. This ensures that the citizen is informed about the re-use taking place, has consented to it and that everyone creates his own tailor made model for re-use. As an incentive, citizens might be rewarded a percentage of the profit made by the re-using parties or a lump sum per time information is re-used.

References

- Aichholzer, G. (2004). Electronic Access to Public Sector Information: Some Key Issues. *Electronic Government, Lecture Notes in Computer Science*, 2004-3183, p. 525-528.
- Burdon, M. (2009). Commercializing Public Sector Information: Privacy and Security Concerns. *IEEE Technology and Society*, 2009-2, p. 34-40.
- Goens, D. (2010). The exploitation of Business Register data from a public sector information and data protection perspective: A case study. *Computer law & security Review*, 2010-26, p. 398-405.
- Hay, M. & Miklau, G. & Jensen, D. & Towsley, D. & Weis, P. (2008). Resisting Structural Reidentification Anonymized Social Networks, *Computer Science Department Faculty Publication Series*, Paper 179, 102-114.

⁸³ When individuals can determine who may use their data - in the expectation to be compensated - this would not discriminate against non-commercial uses, such as journalistic, statistic and scientific usages, since they already fall under a special regime under the Data Protection Directive.

Kuner, C. (2007). *European data protection law: corporate compliance and regulation*. Oxford University Press, Oxford, 2007.

Malin, B. & Sweeney, L. (2004). How (not) to protect genomic data privacy in a distributed network: using trail re-identification to evaluate and design anonymity protection systems. *Journal of Biomedical Informatics*, 2004-37, p. 179–192.

Ohm, P. (2010). Broken promises of privacy: responding to the surprising failure of anonymization. *UCLA law review*, 2010-57, p. 1701-1777.

Pas, J. (2004). The Use and Re-use of Public Sector Information (PSI): Some Legal and Policy Reflections. <http://www.cais-acsi.ca/proceedings/2004/pas_2004.pdf>, p. 1-14.

Pas, J. & Vuyst, de B. (2004). The Use and Re-use of Government Information from an EU Perspective. *Proceedings of the 37th Hawaii International Conference on System Sciences*, p. 1-6.

Roessler, B. (2005). *The value of Privacy*. Polity Press, Cambridge, 2005.

Solove, D. J. (2002). Access and Aggregation: Public Records, Privacy and the Constitution. *Minnesota law review*, 2002-86, p. 1137-1203.

About the Author

Bart van der Sloot

Bart van der Sloot is a researcher at the Institute for Information Law, University of Amsterdam, Netherlands. He specializes in privacy, but is also interested in the liability of internet intermediaries, internet regulation and copyright issues on the web. He is the coordinator of the Amsterdam Platform for Privacy Research (APPR), which incorporates about 50 researchers from the University of Amsterdam, who in their daily research and teaching focus on privacy related issues. They do so from different perspectives, such as law, philosophy, economics, informatics, medicine, communication studies, political science, etc. In October 2012 APPR organizes a four-day, interdisciplinary, international privacy conference in Amsterdam. See: www.apc2012.org and <http://www.appr.uva.nl/appr-en/main.cfm>