Informatica e diritto

Rivista internazionale

diretta da
Costantino Ciampi

Fascicolo 1-2, 2011, ESI, Napoli, 550 p.

Open Data e riuso dei dati pubblici

Open Data and Re-use of Public Sector Information

a cura di

Daniela Tiscornia

Il quadro giuridico

Un nuovo concetto di *Open Government*

Il panorama internazionale e le buone pratiche

I linguaggi, gli strumenti e le applicazioni

Appendice

Public Sector Information & Data Protection: A Plea for Personal Privacy Settings for the Re-use of PSI

BART VAN DER SLOOT*

1. INTRODUCTION

Already in the year 2000, the total value of the European public sector information (PSI) was estimated to be around 68 billion euro annually[1]. 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[2], although it has to be said that it does not entail any obligation for public sector organizations to disseminate such information. It is important to distinguish between the right of access to governmental information and the right to re-use[3]. The *ratio* behind rules ensuring access to governmental information is mainly linked to democracy and control on governmental power; re-use is primarily important for commercial interest[4].

With regard to the re-use of PSI, 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 regarding re-use of public sector information. A public sector body means the state, regional or local authorities, bodies governed by public law and associations formed by one or several such authorities or one or several such bodies governed by

* The Author is Researcher at the Institute for Information Law, University of Amsterdam, specialized in privacy.

[1] EUROPEAN COMMISSION, *Commercial Exploitation of Europe's Public Sector Information*, 20 September 2000, p. 6.

[2] 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).

[3] J. PAS, B. DE VUYST, *The Use and Re-use of Government Information from an EU Perspective*, Proceedings of the 37th Hawaii Int. Conf. on System Sciences, 2004, p. 2.

[4] D. GOENS, *The Exploitation of Business Register Data from a Public Sector Information and Data Protection Perspective: A Case Study*, in "Computer Law & Security Review", 2010, n. 26, p. 399.

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[5]. A public sector document refers to any content or any part of such content whatever its medium[6]. 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[7].

Thus, the scope of the PSI Directive is very broad[8]. Since public sector information also contains personal data, the distribution of the information for re-use may trigger the applicability of the Data Protection Directive[9]. The PSI Directive 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[10]. National legislators should implement and apply the PSI Directive in full compliance with the principles relating to the protection of personal data in accordance with the Data Protection Directive[11]. Finally, the PSI Directive does not contain a definition of personal data, but refers to that of the Data Protection Directive[12]. Both the Commission's Green Paper[13] and the recent communication on the review of the PSI Directive[14] confirm the full applicability of the Data Protection Directive.

[5] See art. 2, par. 1 and 2, PSI Directive.

[6] See art. 2, par. 3, PSI Directive.

[7] See art. 2, par. 4, PSI Directive.

[8] See art. 1, par. 2, PSI Directive; art. 3, PSI Directive enumerates some exceptions.

[9] 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).

[10] See art. 1, par. 4, PSI Directive.

[11] See Recital 21, PSI Directive.

[12] See art. 2, par. 5, PSI Directive.

[13] 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.

[14] 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*, in "SEC", 2009, n. 597, p. 8; C. CORBIN, *EC Communication on the PSI Directive: PSI Re-use Stakeholder Reaction*, European PSI Platform, Topic Report n. 3.

2. Data Protection Directive

In this part, it will be determined to what extent the Data Protection Directive applies to the re-use of public sector information. If it does apply, three major categories of obligations will need to be taken into account. Firstly, there must exist a legitimate purpose for processing personal data. Secondly, 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.

2.1. 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[15]. The Working Party 29[16], the advisory institution regarding privacy and data protection in the European Union[17], has elaborated on four elements of personal data: 'any information', 'relating to', 'an identified or identifiable' and 'natural person'[18]. 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[19]. A natural person is a living physical person.

Personal data may either be directly identifiable, such as a name, or indirectly, such as a telephone number or a combination of non-directly identifi-

[15] See art. 2, par. A, Data Protection Directive.

[16] See http://ec.europa.eu/justice/policies/privacy/workinggroup/index_en.htm.

[17] See Recital 65 and art. 29-30, Data Protection Directive.

[18] See Article 29, Working Party, Opinion 4/2007 on the Concept of Personal Data (WP 136), Brussels, 20 June 2007.

[19] WP 136, p. 10.

able 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"[20]. 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[21]. 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 either the controller or third parties that have access to the information are able to identify individuals, which is especially relevant in the case of re-use.

Given the general scope of the definition of personal data, many documents will contain personal data[22]. 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[23]. It is not uncommon for public sector bodies to hold such information[24].

"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); em-

[20] WP 136, p. 13.

[21] See Recital 26, Data Protection Directive.

[22] Judgment of the European Court of Justice C-101/2001of 06.11.2003 (Lindqvist), par. 27.

[23] See Recital 33 and art. 8, Data Protection Directive.

[24] WP 83, p. 4; J. PAS, B. DE VUYST, *The Use and Re-use of Government Information from an EU Perspective*, cit., p. 1.

ployment (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)"[25]. In conclusion, much of the public sector information will contain both ordinary and sensitive personal data. Therefore, the first criterion of applicability will usually be satisfied in the case of re-use of PSI.

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[26]. In short, almost anything that can be done with personal data falls within this all-encompassing definition[27]. 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"[28]. Thus, there are three stages of processing the personal data. Firstly when the data is originally gathered and used 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

[25] D.J. SOLOVE, *Access and Aggregation: Public Records, Privacy and the Constitution*, in "Minnesota Law Review", 2002, n. 86, p. 3.

[26] See art. 2, par. B, Data Protection Directive.

[27] D. GOENS, *The Exploitation of Business Register Data from a Public Sector Information and Data Protection Perspective: A Case Study*, cit., p. 402.

[28] See 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.

determines the purposes and means of the processing of personal data[29]. On him lie all the obligations under the directive[30]. The public sector organization disseminating the information will be qualified as the controller at the time of gathering, analyzing, using and disseminating the information. The re-using third party will also qualify as the controller of the personal data when gaining the public sector information for re-use. Consequently, it too has to fulfill all the obligations under the directive[31].

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[32]. There is no doubt that public sector organizations would fall under the first category[33]. To fall under the first category, the re-using third parties must have an establishment on the territory of a Member State, which implies the effective and real exercise of the business activity[34]. 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[35]. Furthermore, the third party should process the data in the context of his everyday business activities[36]. 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[37].

[29] See art. 2, par. D, Data Protection Directive.

[30] See Article 29, Working Party, Opinion 1/2010 on the Concepts of "Controller" and "Processor" (WP 169), Brussels, 16 February 2010.

[31] WP 169, p. 33.

[32] See art. 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.

[33] See Article 29, Working Party, Opinion 8/2010 on Applicable Law (WP 179), Brussels, 16 December 2010.

[34] European Court of Justice, 4 July 1985, Case C-168/84, (Berkholz).

[35] See Recital 19 Data Protection Directive.

[36] WP 179, p. 14.

[37] WP 179, p. 23.

The purpose of the directive is to guarantee data subjects an adequate level of protection, wherever the controller is established[38]. If the re-using 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 re-using 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[39].

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 re-using party is using the information for its own purposes. Both the governmental authority and the re-using 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 by the directive and finally, respecting the rights of the data subject in connection with the transparency principle.

2.2. 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

[38] See Recital 18 and 20, Data Protection Directive.

[39] See art. 25, Data Protection Directive.

data subject[40]. 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[41]. 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[42].

As a controller, the third party receiving the public sector information must also fulfill one of the six circumstances mentioned in the directive. Usually, the 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. 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[43]. 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[44]. 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[45], the Charter of Fundamental Rights of the European Union[46] and the United Nations Universal Declaration of Human Rights[47]. 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 on the side of the controller and the re-use would thus not be legitimate, save some exceptional cases. Only if another fundamental right is served by the re-use of public sector information containing personal data, most

[40] See Recital 30-32 and art. 7, Data Protection Directive.
[41] WP 83, p. 5 and Recital 9, PSI Directive.
[42] See art. 3, Data Protection Directive.
[43] See art. 2, par H, Data Protection Directive.
[44] See Recital 30, Data Protection Directive.
[45] The European Convention on Human Rights, Rome 4 November 1950, article 8.
[46] Charter of Fundamental Rights of the European Union, (2000/C 364/01), art. 7-8.
[47] United Nations Universal Declaration of Human Rights, December 10, 1948, art. 12.

commonly the right to freedom of speech, will there be a situation in which the two equal interests must be balanced. This 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 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 term[50] or if processing relates to health care issues or any other issues relating to the public interest as laid down in law[51]. This allows them to legally gather and process large quantities of sensitive data. 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 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 re-using third party would

[48] WP 83, p. 5.

[49] See Recital 33-36 and art. 8, par. A, Data Protection Directive.

[50] See art. 3, Data Protection Directive.

[51] See art. 8, par. B and par. C, Data Protection Directive.

[52] The term "explicit" consent, instead 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. See art. 3, Data Protection Directive.

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].

2.3. 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, it has to inform a third party to whom it has distributed public sector information containing personal data when it is 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 collected[57] 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

[54] See Recital 17, 37 and art. 9, Data Protection Directive.
[55] See Recital 28, Data Protection Directive.
[56] See art. 6, par. 1, *sub* D, Data Protection Directive.
[57] See art. 6, par. 1, *sub* C, Data Protection Directive.
[58] See art. 6, par. 1, *sub* E, Data Protection Directive.
[59] See art. 6, par. 1, *sub* B, Data Protection Directive.

the third party requesting public sector information for 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 reused 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 purposes[63] will they fulfil their obligations in this respect.

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.

[60] WP 83, pp. 7-8.

[61] C. KUNER, *European Data Protection Law: Corporate Compliance and Regulation*, Oxford, Oxford University Press, 2007, p. 100.

[62] WP 83, p. 9.

[63] See Recital 29, Data Protection Directive.

[64] See Recital 46 and art. 17, Data Protection Directive.

2.4. Transparency & Rights

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 with regard to the transparency principle and the right to access 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 to re-use public sector information. 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 re-using 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 too that it informs every data subject of his identity, the purposes for processing etc.

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 re-use purposes could fall under this exemption, since transparency would mean

[65] See art. 10, Data Protection Directive.

[66] See Recital 43 and art. 13, Data Protection Directive.

[67] See Recital 42, Data Protection Directive.

[68] See art. 11, Data Protection Directive.

[69] See Recital 39-40 and art. 11, par. 2, Data Protection Directive.

an unbearable burden, although it has to be stressed that the directive seems to have primarily historical, statistical and scientific research in mind with regard to this exception.

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.

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 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 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 for re-use of public sector information by third parties that have found the legitimacy of processing in the balance of interests.

[70] See Recital 41 and art. 12, par. A, Data Protection Directive.

[71] See art. 12, par. B and par. C, Data Protection Directive.

[72] See art. 14, Data Protection Directive.

3. Solutions

There are several obstacles with regard to the re-use of public sector information in the light of the Data Protection Directive. 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 are incompatible, the possibility to re-use public sector information is 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 the introduction of personal privacy settings regarding the re-use of citizens' personal data by third parties.

3.1. 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[73]. 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[74]. While primarily focusing on prosperity and profit by the privacy sector, this model

[73] J. Pas, B. de Vuyst, *The Use and Re-use of Government Information from an EU Perspective*, cit., p. 2.

[74] J. Pas, B. de Vuyst, *The Use and Re-use of Government Information from an EU Perspective*, cit., p. 2; J. Pas, *The Use and Re-use of Public Sector Information (PSI): Some Legal and Policy Reflections*, in http://www.cais-acsi.ca/proceedings/2004/pas_2004.pdf, p. 9.

is based on open and unrestricted access to public sector information at no more than the cost of search and duplication[75]. 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. Waiving away these rights 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[76]. This being the case, both radical solutions do not seem very satisfying.

3.2. Anonymization

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 identifiable[77] 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"[78].

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

[75] G. AICHHOLZER, *Electronic Access to Public Sector Information: Some Key Issues*, in "Electronic Government, Lecture Notes in Computer Science", Vol. 3183, 2004, p. 2.

[76] For the more fundamental value of privacy see B. ROESSLER, *The Value of Privacy*, Polity Press, 2005.

[77] See Recital 26, Data Protection Directive.

[78] WP 83, p. 4.

information would be almost impossible[79]. 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"[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.

3.3. 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 ones 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 of those countries 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 differentiate between indirect and direct identifiable information and between ordinary and sensitive personal data, though it might be feasible to prohibit the re-use of sensitive personal data, since this could hinder a per-

[79] See http://patientprivacyrights.org/wp-content/uploads/2010/08/The-Case-for-Informed-Consent.pdf.

[80] P. OHM, *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization*, in "UCLA Law Review", 2010, n. 57, p. 1704. Moreover, there is the problem of re-identifycation.

son's privacy and autonomy to an intolerable extent. Finally, 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 either a percentage of the profit made by the third parties trough the re-use of their personal data or a lump sum set by the government,

[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; M. BURDON, *Commercializing Public Sector Information: Privacy and Security Concerns*, in "IEEE Technology and Society", 2009, n. 2.

depending on type of data and the period and purpose of processing. This creates an incentive to register and ensures a 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 is 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.

4. Conclusion

Most public sector documents will contain personal data, given the fact that the definition thereof is all-encompassing. 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 with 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 value of the documents is retained. However, since the concept of personal data is so big, this might be a Sisyphean task and even if the governmental organizations would succeed, the data would presumable have lost most of its value. A new solution is suggested in this paper, namely to let every citizen register its own personal privacy settings regarding the 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.