



European
Commission

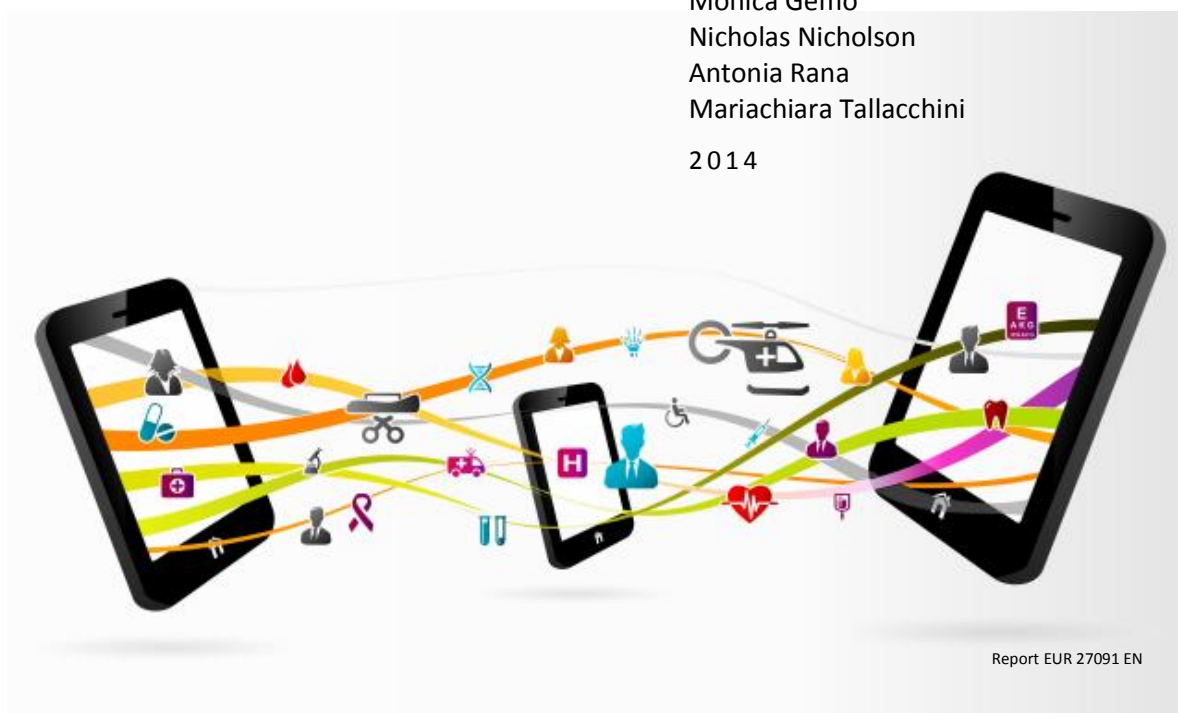
JRC SCIENCE AND POLICY REPORTS

OPEN DATA IN HEALTH: how knowledge may generate trust

*Report on the workshop
held at JRC, Ispra, Italy
18th November 2014*

Estefania Aguilar Moreno
Monica Gemo
Nicholas Nicholson
Antonia Rana
Mariachiara Tallacchini

2014



Report EUR 27091 EN

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information

Monica Gemo

Address: Joint Research Centre, Via Enrico Fermi 2749, TP 361, 21027 Ispra (VA), Italy

E-mail: monica.gemo@jrc.ec.europa.eu

Tel.: +39 0332 783081

<https://ec.europa.eu/jrc/>

<http://ipsc.jrc.ec.europa.eu/>

This publication is a Science and Policy Report by the Joint Research Centre of the European Commission.

Legal Notice

This publication is a Science and Policy Report by the Joint Research Centre, the European Commission's in-house science service. It aims to provide evidence-based scientific support to the European policy-making process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

Cover image © madpixblue - Fotolia.com

JRC 92431

EUR 27091 EN

ISBN 978-92-79-45348-9

ISSN 1831-9424

doi: [10.2788/421226](https://doi.org/10.2788/421226)

Luxembourg: Publications Office of the European Union, 2014

© European Union, 2014

Reproduction is authorised provided the source is acknowledged.

Printed in Italy

Table of contents

The Joint Research Centre	3
Introduction	5
The workshop and the dialogue among participants	9
Recommendations	13
1. General recommendations on Open Data in health and trust	13
2. Recommendations on making Open Data and privacy complementary	18
3. Recommendations on the different roles for institutions and the media	19
The agenda	21
Abstracts and presentations	22
Participants	103

The Joint Research Centre

Joint Research Centre

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle. Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.



Institute for Protection and Security of the Citizen

The Institute for the Protection and Security of the Citizen (IPSC) is one of the seven institutes of the European Commission's Joint Research Centre (JRC). IPSC is an applied research and development institute, aimed at analyzing, modelling and developing new security applications in support to EU policies.

Digital Citizen Security Unit

The mission of the Digital Citizen Security Unit is to strengthen trust and security of the European Citizen in a sustainable and inclusive ICT-based European society by scientific research on how emerging Information and Communication Technologies will impact on the security and privacy of citizens' daily life. In the balance between European security needs and fundamental citizen rights, the unit works on risk mitigation, on cyber security, data protection, privacy and other ethical considerations, and on the associated legal and regulatory frameworks.

Institute for Health and Consumer Protection

The Institute for Health and Consumer Protection (IHCP) is another of the seven institutes of the European Commission's Joint Research Centre (JRC). IHCP provides scientific and technical support to EU policies in the areas of food, consumer products, chemicals and public health, with particular emphasis on the protection of interests and health of European citizens.

Open Data in Health: how knowledge may generate trust

Report on the workshop held at the
JRC - Ispra (Italy), 18th November 2014

Introduction

In the process of European integration, trusted relations between institutions and citizens have been highlighted as a key-element in establishing reliable conditions for citizens when dealing with institutions. The passage from primarily direct and human-based relations to mostly digitalised interactions has even increased the need for trust, not only in relation to the technical aspects of security measures (e.g. protection of data in the web), but also in connection with the normative issues of transparency, accountability, openness, accessibility, etc. Here, the expression “building trust” refers to the process of identifying and implementing the necessary requirements to frame and maintain confident digital interactions: this process needs some dedicated reflections and actions, and should be shaped as a continued iterated process.

Since the origins of the scientific method and of the State under the rule of the law, knowledge, and knowledge production and exchange, have been depicted as essential elements of trustworthy institutions and of trusted public decision-making. This connection between knowledge and trust has been fundamental in establishing the ethics and integrity of science within the scientific communities. In knowledge and technology-based societies the creation and sharing of knowledge represents a major path towards generating and maintaining trusted relations between institutions and citizens.¹

In this context Open Data reveals a special relevance. The philosophy behind the concept has been long established within the scientific community. Before becoming a matter of institutional concern, Open Data as “open knowledge” was rooted in the practice and the ethos of the scientific community. Researchers perceived the benefit of openness and of sharing of data in their activities. Already in the

¹ Y. Benkler, and H. Nissenbaum. 2006. Commons-based Peer Production and Virtue. *The Journal of Political Philosophy*, 14,.4, 394-419.

mid-20th century, in his classical elaboration of the normative structure of science and the scientific community, sociologist Robert Merton was supporting the idea of common good applied to knowledge as part of the “ethos” of science, while showing the practical benefits of open scientific data (and the need for absence of intellectual property rights in research).²

However, the expression "Open Data" in itself is recent and has gained popularity with the rise of the Internet. Open Data is, according to a widely accepted definition—also subtended to the EU policies and legislation—,³ data that can be freely used, reused and redistributed by anyone—subject only, at most, to the requirement to attribute and share alike.

The revised EU Directive 2013/37/EU on the re-use of public-sector information emphasises the "vast, diverse, and valuable pool of resources that can benefit the knowledge economy," especially in terms of development of new services based on novel ways of combining and applying such information – with the goals of stimulating economic growth and promoting social engagement.

However, even though Open Data have been framed by European institutions mostly as new field for economic development, the creation, use and reuse of data by citizens also represents a great opportunity to test and improve trust in institutional digital interactions.

Indeed, the very meaning of Open Data lies at the interface between epistemology and democratic theory. Due to its potential for sharing knowledge and knowledge production, Open Data can play a unique role as to the task of shaping trusted digital relations.

The concept refers to a deeply value-laden vision of the human cognitive endeavor, namely the ideal of the universal and boundless sharing of knowledge; moreover, it is a phenomenon radically rooted in, and generated by, digital technologies.

On the one hand, from an epistemological perspective, the open character (especially) of science has been framed and proposed as both an indicator and an evidence itself of valid knowledge; on the other hand, from a political point of view, this openness has been associated to the democratic connotation of the society based on this kind of knowledge.

Openness in Open Data, as the feature highlighting the correct approach to both epistemic and democratic systems, is shared by other "open" movements such as open source and open access. All these trends share specific values in looking at the relations between science and society, encompassing availability and access, reuse and redistribution, and universal participation.⁴

² R.K. Merton. 1968. Science and Democratic Social Structure. In: Social Theory and Social Structure, 604–615, Free Press, New York.

³ See Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, and Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information on the re-use of public sector information. See also the Communication from the Commission to the Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, Towards a thriving data-driven economy, COM(2014) 442 final, Brussels, 2.7.2014.

⁴ <http://opendatahandbook.org/en/what-is-open-data> [Last accessed: 02/12/2014]

First, data must be available as a whole, and at no more than a reasonable reproduction cost, preferably by downloading over the Internet. Also, the data must be available in a convenient and modifiable form. Second, data must be provided under terms that permit reuse and redistribution including the intermixing with other datasets. Third, everyone must be able to use, reuse and redistribute. For example, ‘non-commercial’ restrictions that would prevent ‘commercial’ use, or restrictions of use for certain purposes (e.g. only in education), should not be allowed.

In any case, Open Data cannot be conceived in isolation. The possibilities offered by new data processing technologies and new means of data analytics increase the value of each single dataset in the sense that it can be linked and enriched with other data sources in order to get more valuable information and, in some cases, to re-identify individuals with unpredictable consequences in terms of impact on fundamental rights. It is worth to say that Open Data is also integral part of Big Data so any analysis on the value and potential of Open Data practices need to be seen in the light of this.

Open Data has several applications in the domain of state reforms and e-government. Amongst these, data related to health seems to possess a privileged role in connecting knowledge and trust between institutions and citizens.

Health is a highly political domain, not only in terms of social welfare allocations, but also for its connections to the environment, environmental politics, and industrial policies. Environment and health are intertwined domains where knowledge production by citizens (citizen science) has often complemented, when not confronted, official knowledge. Moreover, significant changes are happening in how scientists and citizens relate and become “partners” in performing research (e.g. in genomic research). Finally, health represents a major sector where citizens are willing to become more knowledgeable and empowered in order to make better informed, autonomous, and personalized decisions. It is definitely interesting to see how, through several web-mediated initiatives, health data have become a way for citizens and researchers to argue against certain legal restrictions (for instance as to the sharing of genetic information) and to introduce new rights (e.g., access to raw genetic data as a new (moral) individual right.⁵

However, and at the same time, even though citizens’ interactions with health-related institutions—and with institutions in general—are increasingly mediated by digital technologies, often citizens cannot easily process and reuse health-related information because of the high skills and technological means required. In this respect, the assumption that trust can be simply strengthened by opening up the health data sector cannot be taken for granted: relations amongst knowledge, trust, and technologies have to be better explored and rethought.

⁵ J.E. Lunshof, G.M. Church, B. Prainsack. 2014. Raw Personal Data: Providing Access. *Science*, 24 January, 343, 6169, 373-374.

A key point, here, concerns the importance of how data are used, and especially aggregation or combination of data for developing new services and applications—without which both economic and social opportunities are likely to be missed.

The combination or linkage of data sets is the *sine qua non* of building and enhancing information on which to make better decisions. In this regard, a clear distinction has to be made between data, information, and knowledge. In the most rudimentary sense, data refers to hard and tangible facts/measurements captured at specific points in time. The fusion and combination of relevant but different data sets provides an information base which can be analysed and weighed. This process develops a knowledge base on which inferences can be drawn and decisions made. The distinction is a crucial one when talking about health data. Data is the basis of information; if only a portion of the data is available, it can lead to biased or even wrong information.

Despite the potential for Open Data to contribute to re-founding institutional trust, several issues need to be clarified and addressed. A full agreement about how Open Data should be normatively defined and what health data means does not exist. While the EU is still completing a normative framework for Open Data, a harmonised vision of it is not yet in place. Gaps in national legislations and, lack of standardization in Open Data requirements, quality, and availability are undermining its value in enabling institutional trust. Also, the focus and the relevance of Open Data have been primarily associated with its economic and commercial value, while its role in re-establishing citizens' trust towards institutions—beyond the rhetoric of mere transparency—has not been adequately investigated. For instance, the fact that the existing policies and laws on privacy and Open Data are not convincingly harmonised, not only diminishes the effectiveness of public action, but can also instigate distrust. Moreover, in order to fully exploit Open Data as a source for trust, two further elements are crucial, namely that the necessary security measures are set up, and that adequate learning and skills are provided to citizens to empower them.

The workshop and the dialogue among participants

The JRC project on Trust in Digital Interactions (TRUDI) deals with the construction and renewal of confident and trusted relationships between institutions and citizens, addressed as a major and urgent issue to be solved. In this context the JRC organised a Workshop on “Open Data in Health: how knowledge may generate trust” on November 18, 2014. The workshop aimed to investigate some general issues surrounding Open Data in the EU normative perspective, reflect on institutional and civic imaginaries about Open Data, and identify more promising Open Data models to trigger new processes for trust between institutions and citizens in health matters. The issues and questions that the workshop aimed to explore were the following. What is the state-of-the-art on Open Data in the European policies and legal documents? Is Open Data mostly perceived as an economic opportunity? Is its civic dimension adequately highlighted? Are Open Data policies and initiatives addressing and meeting citizens’ needs properly? Which trade-offs are currently existing between Open Data and Privacy? Does the current situation actually reflect citizens’ readiness to engage in the governance of their health data? Is open “raw” data enough for building trust amongst institutions and citizens? How technical security measures can—if they can— help overcome these tensions? What is needed to adequately empower citizens in using Open Data in the health sector?

The workshop encompassed three sessions, each asking a different set of questions, namely:

- 1 - What is the current state-of-the-art on Open Data? What does Open Data in health mean?
- 2 - European policies on Privacy and Open Data: conflicting or complementary?
- 3 - Existing portals for Open Data, institutional duties, citizens’ expectation: which roles for whom? Which role for the media?

In each session—chaired by a European Commission officer with expertise in the related field—different visions on Open Data, from the institutional, the academic/research, and the civil society perspectives were illustrated.

The first set of questions, mostly related to the meaning and implications of Open Data in the relations between institutions and citizens. The issues were explored through the illustration of two cases of Open Data.

In its presentation Rob Hagendijk (Amsterdam University, NL) described the socio-technical imaginaries related to the data on rare diseases and the relevance of the European normative framework for rare diseases as the adequate and trusted scale for collecting and sharing health information. Indeed, despite their manifested anti-Europe feelings, citizens in the Netherland expressed the need for a Europeanization of knowledge about rare diseases, thus revealing that an important element of trust exists towards EU institutions.

The Research director of the French Ministry of Health, Franck von Lennep, illustrated the French Open Database in the health sector and the Report on Open Data en Santé published by the Commission on

Open Data (July 2014). Even though the commitment of the French administration towards Open Data in health started 10 years ago, the Report represents the beginning of a new policy plan aimed at implementing the idea of “démocratie sanitaire” within the context of renewed trusted relations between institutions and citizens in the State under the rule of the law (Etat de droit).

The second session addressed the controversial relations between privacy and Open Data through two complementary presentations, given by Manuel Garcia Sanchez, from the Spanish Authority for Data Protection, and by Magnus Stenbeck, epidemiologist from the Swedish Karolinska Institute.

Sanchez expressed concern towards the impact that Open Data can have on privacy, showing how these impacts are still largely unknown and unpredictable. In Sanchez’s opinion, a culture of privacy should be disseminated amongst researchers, while several technical and non-technical measures should be adopted in order to cope with uncertainties and threats posed by unleashed data.

Stenbeck illustrated the longstanding Swedish policy in support of collecting health data through registries, and argued that the proposed European Regulation has the potential to threaten this public health framework.

Besides their different perspectives as data protection officer and epidemiologist, however, both speakers agreed that Open Data should come together with privacy. Indeed, even though the new proposed Regulation on data protection seems to pose some challenges to registries on health and diseases (especially in the field of tumor registries), privacy should find its place within the concept and practice of Open Data.

Finally, the third session primarily focused on the role of the media in improving Open Data usability for citizens and therefore in empowering them.

Antonia Rana (scientific officer at the JRC) presented an articulated and detailed analysis of Open Data primarily from the perspective of their quality, showing how the quality aspects may be crucial in connecting knowledge and trust. Also, Rana explored some existing portals on Open Data, comparing them in terms of usability by citizens.

Rosy Battaglia (also on behalf of Guido Romeo, both data journalists) touched on several issues related to how usable Open Data can play a relevant role for citizens as tools to protect their health and the environment.

Following the presentations, the Chairs started and moderated the discussion, allowing participants to the workshop to express their views. The Workshop aimed at drafting some suggestions and recommendations to be offered to the European Commission.

The discussion following the first session focused on the definition of Open Data and the need for a clear univocal language as a matter of transparency and democracy. Several comments also addressed the issue of citizens’ participation and involvement in defining the agenda and the priorities for Open Data in health.

In the second session, the discussion touched on how Open Data and privacy can properly relate. Directive 2013/37/EU on the re-use of public-sector information makes reference to the current data protection Directive (95/46/EC) and clearly bounds the scope of Open Data in relation to it. In this sense

one could indeed say – at least from a superficial view – that the European policies on privacy and open data are complementary. However, participants agreed this complementarity still needs to be achieved.

In the last two years the JRC started supporting the work of the European Network of Cancer Registries in view of the EU's priorities in tackling cancer. Cancer registries have to collate data from a number of different sources before deriving the full set of information forming the basis on which cancer epidemiological research can be conducted. The type of data-protection framework on which Europe decides may have profound implications on this process. How can Europe ensure a sensible data-protection framework that is proportionate to the real risks posed to sensitive personal data? Concern was expressed that, in the worst case scenario, many current useful epidemiological tools could be severely compromised and even lost altogether; moreover, that exciting future developments in the field – possible via the link-up of different data sources — may be entirely thwarted. However, other views were expressed outlining that it is possible to properly deal with the needs of epidemiological research while ensuring full respect of individual's fundamental rights. In that sense, researchers have to work with data protection practitioners and bioethicists in order to cope with tools allowing full transparency, adequate information and consent management for individuals while fulfilling personal data processing needs in research.

In the third and final session, the quality and usability of knowledge were addressed, especially in relation to knowledge production by citizens. Several examples were made of initiatives led by citizens and scientists who, by making use of ICT and health data, have joined with the goal of protecting environmental health in highly polluted contexts. These situations illustrate how knowledge and trust are closely connected in the relations between citizens and institutions.

Institutions should be committed to produce good data as part of their obligation to empower citizens and to implement their rights. However, most participants agreed on the role of data journalists as mediators of knowledge and as controllers of institutional behaviour.

As a whole, the workshop showed that Open Data, and Open Data in Health represent a set of complex issues that needs to develop in order to reveal its different aspects, its potential and unforeseen implications. Only the progressive unfolding of the process of releasing and reusing data, though with some forms of oversight and protection, is the way forward.

Recommendations

This section summarises reflections and recommendations emerged and shared through the discussions by participants. They reflect the division in three Groups organised around the three thematic sessions of the Workop.

The discussion in Group 1 focused on how to create more trust between civil society, industry, and the government.

Group 2 focused on the seemingly conflicting policies of Open Data and privacy, and on the need to overcome the existing gaps.

Group 3 focused on the different roles that social structures other than institutions, especially the media, should perform to improve usability of Open Data by citizens.

Some recommendations remain intentionally overlapping, thus revealing and highlighting the connectedness of the different issues. These overlaps have been here signalled through cross-references.

1. General recommendations on Open Data in health and trust

13

1.1. A clear and harmonized definition of “Open Data” and the accompanying Open Standards (see 3.1)

Reference to Open Data should be made only when some minimal requirements for Open Standards⁶ are met and when the data conforms to “the” Open Data definition (see 2.1, 2.2 and 3.1).

As said, according to the definition adopted by the EU, but representing a shared broad understanding of Open Data, this is “data that can be freely used, reused and redistributed by anyone—subject only, at most, to the requirement to attribute and share alike.” However, several official documents have characterized Open Data in a precise and detailed way. According to the UK White Paper on Open Data (2013), for instance, Open Data, namely “(q)ualitative or quantitative statements or numbers that are assumed to be factual, and not the product of analysis and interpretation,”⁷ is precisely classified depending on their degree of openness and has to meet certain requirements.

The classification follows Tim Berners Lee’s system of quality stars, where the optimal situation is defined by the assignment of five stars: openly licensed, openly accessible, structured, open format, URIs for entities, linked.

⁶ UK Cabinet Office, Open Standards Principles, September 2013, Available at <https://www.gov.uk/government/publications/open-standards-principles/open-standards-principles#open-standard--definition> (Last accessed: 10 December 2014).

⁷ HM Government, Open Data. White Paper. Unleashing the Potential. Presented to Parliament by the Minister of State for the Cabinet Office and Paymaster General by Command of Her Majesty, June 2012, 7.

As to standards, Open Data has to be accessible at no more than the cost of reproduction; be in a digital, machine readable format for interoperability with other data; and be free of restriction on use or redistribution. Common standards are essential in order to achieve the main goal of Open Data. Indeed, not all data in the public sector is standardised in quality and is equally accessible; and “(a) lack of common standards is a barrier that can make it difficult for users to scrutinise activity or generate added value.”⁸ According to this vision of the roles and responsibilities of institutions, citizens should not be exposed to increased, unjustified costs due to the specific digital choices made by institutions.

1.2. Need for an open governance of Open Data and the need for an inclusive process (see 3.2)

The process for generating data should be as inclusive as possible, encompassing different narratives and all relevant actors, researchers, government and industry, but citizens (patients) and data journalists as well. Data can be very complicated to be collected, summarised and interpreted, and experts may not be aware of what health data are needed by citizens: focus groups, hackathons, surveys, etc. may be used to find this out.

This open engagement is likely to raise awareness and self-reflexivity about the role of Open Data in changing the way certain interactions (e.g., doctor/patient relations) are now taking place. The fact that data is open to everybody has a strong meaning as to a radical democratization of medicine—what the French document on Open Data in Health has called “*démocratie sanitaire*.”⁹

1.3. Knowledge and data should derive from a legitimate process

Attention to the process through which data are produced has been stressed as a major factor in building trust. Knowledge may generate trust when knowledge appears legitimate not just by reference to its being scientifically valid, but because it is produced through a fair, democratic and inclusive process.

Building legitimacy also requires transparency of the process making the assumptions underlying an Open Data project explicit and readily accessible. Disease registries should be built with more attention to an inclusive, democratic process, namely a process where different voices are taken into account and that remains transparent through all its phases.

Moreover, trust and the legitimacy of the process are enhanced through multiple correlated data coming from different sources. Indeed, in this way institutions no longer have a privileged position in

⁸ UK Cabinet Office, Open Standards Principles, September 2013, cit.

⁹ Ministère des Affaires Sociales et de la Santé, Commission open data en santé, Rapport, 9 juillet 2014. Available at: http://www.drees.sante.gouv.fr/IMG/pdf/rapport_final_commission_open_data-2.pdf (Last accessed: 11 December 2014).

producing valid knowledge to be used in decision-making, but they become part of an increased transparent process of both knowledge production and trust production (see 3.6).

1.4 Open Data and open source

In order to transparently achieve the goal of openness in data, open source software may become a necessary requirement. This is software where the rights are granted to access and modify the source code, and to use, reuse and redistribute the software, with no royalty or other costs.

As some documents have made clear, Open Data, Open Standards, and in some cases also Open Source should be seen as component of the same process towards transparency and shared production of knowledge.

1.5. Provide the data and then proceed to improve them

It is not always necessary to have ideal conditions: provide the data first, and learning will follow. Release the data, give room to the utilization and the debate, and allow interpretations unfold,¹⁰ while providing mechanisms for oversight and alertness in order to prevent potential negative outcomes (see 2.2 and 2.3).

Indeed, release of data should be preceded by a careful and detailed impact analysis.¹¹

This does not mean (see 1.6) that the quality of data is not relevant: on the contrary, the goal must focus on achieving the five star ranking— data should be curated in terms of formats and accessibility, without cleaning it, i.e. original data.

Go beyond discussion of principles and discuss the specific technical as well as social and normative solutions—with experts, civil society representatives, etc. to reach an accepted compromise solution and also to build legitimacy.

Even though it is not clear in advance what society wants to know, and whether citizens are interested in the “open” data or are they looking for “open” answers, the unfolding of the process will clarify these issues.

1.6. Data, metadata, and quality

¹⁰ This recommendation remained controversial as some participants disagreed with it. They proposed instead to integrate the recommendation with the following statement: “Premature dispatch of not validated data can undermine any subsequent analysis and interpretation, the latter also prone to bias, confounding and uncertainty like any other process of inference from raw data. The need for a prompt release of data and a direct use and interpretation of it should be always confronted with the specific situation and context, in particular considering the level of complexity of the data itself and the fact that proposed hypotheses have to be tested through the data.”

¹¹ It needs to be taken into account that, once a dataset is release, little control remains in the hands of the data owner.

On the data portal, include explanation of how the data had been produced (metadata) so as to make it easier to use/interpret it correctly.

Quality assurance shall be pursued at each level of the data life cycle: production, registration, storage, distribution, analysis, interpretation, and shall be applied to metadata as well. Example of quality assurance in the field of cancer registration can be shared in other fields. Quality assurance of data and the usability of the platform should benefit of adequate investments (see 3.3 and 3.5).

1.7. Create participatory and inclusive platforms where the process of generating data becomes transparent and empower users/citizens with “rights in design” (see 3.1, 3.3. and 3.5)

Data (and, more broadly, knowledge) should be shared through participatory and collaborative platforms after having brought together all relevant stakeholders. The notion of a digital platform refers to both the software and hardware of a site for the provision of Open Data and the promotion of the active role for stakeholders. It may include social media and citizens online communities.

Moreover, the process should also take into account how individuals are empowered. Citizens should be entitled to having “rights in design,” namely they should be given the right to influence and/or control some features of digital architectures and to decide how they want to interact with a system (e.g. access to raw data). Multiple user interfaces or pathways should be provided for data sets, as well as different levels of aggregation or disaggregation.

1.8. Make platforms user-friendly

It is important to ensure the quality/user-friendliness of the platform itself. Ultimately, society cares about data especially in relation to the new answers to problems that data can make possible. However, it is essential that those who use the data can provide their own interpretations and can communicate them. The same data may be interpreted in different ways (for instance, different people may have a different approach to the choice of confidence intervals and therefore to inference).

To this end, involve and engage all parties, from scientists to journalists to industry.

Explain technical solutions; explain that full anonymisation may not always be a possible, adequate or desirable solution (see 2.1).

1.9. Need to (re)define the research needs: what the society wants to be addressed (even though citizens may not be aware of their preferences when simply asked)

Attention should be paid more to applied, specific knowledge and data related to citizens’ needs. Digital platforms should provide space for citizens’ needs and agendas, not only for researchers’ interests. Citizens should be allowed to add personal stories to the platforms.

1.10. Sensitive vs. Open Data

Open Data systems need to compromise in order to achieve legitimacy. Indeed, the individual providing the data may have a different understanding of her data in terms of privacy than a scientist looking at them (see 2.1).

Explain that some data become useless once they are anonymized. If Open Data are not used then they are useless, if they are used then the ultimate goal can be questioned.

Find compromises on anonymisation to always ensure the usefulness of Open Data. Adoption of technical and by-design measures when dealing with sensitive records (e.g. medical information) can help ensure research aims while granting full respect of individual rights enshrined in the EU legal framework.¹²

1.11. Post-research traceability

Keep track of research studies (and research hypotheses behind them) which are performed through Open Data. Explain how the data was produced to use it better; assure that the analysis, interpretation and reporting will be done responsibly and in a competent and quality-assured way.

As to research data, scientists should be committed towards the effort of preparing the data to be released for sharing, of documenting the research process, and making it publicly available.

1.12. Pay attention to the relevance of (different) scales in Open Data

While some public issues can be better discussed and shared at local/national level (cultural features specific to a national health system), other topics, such as rare diseases, seem to require a wider scale, especially at the European level.

Scale is also relevant in the issue of aggregation vs. disaggregation. This should always be an option built into any data system. Important aspects can easily get lost in aggregation or can be harder to see in disaggregation.

1.13. Open Data requires a different understanding and framing of owning

In the context of Open Data the concept of “belonging” can often be found to refer to the free use of data instead of the expression “property rights.” This broader approach to “to whom data belongs” requires a deeper understanding and a normative framework.

There is also the need to address technological dimensions of Open Data – attention to e-commerce and the need for incentives for Open Data (to improve trust of people providing their data and allowing them full control of their data).

¹² A variety of views were expressed as to how prioritize needs and how compromise on privacy, and the recommendation remained controversial.

2. Recommendations on making Open Data and privacy complementary

2.1. Provide Open Data “with” privacy (see 1.8 and 1.10)

Open Data and privacy are not and should not be presented as opposing, conflicting issues. Indeed, they should be seen as complementary and a culture of complementarity, as well as mechanisms for this end, should be set in place. A culture of Open Data and privacy should be supported, disseminated and implemented, while individual and community rights should be balanced.

Provide traceability information about data use and modification.

There should not be a dichotomy between consent and no consent. However, is informed consent necessary in every case? And what about transparency?

There is the need to facilitate the consent process (especially to take into account the aspect of evolving needs of data, e.g. for epidemiological use), without overburdening the research work.

A graduated consent mechanism should also be made possible in order to ease the research work while ensuring usefulness of the data. Moreover, building research on trust by individuals also has an impact on data quality and, consequently, on increasing the added value of data. The cancer registry example is one of the cases in which new methods need to be developed, as cases missing for lack of consent or other reasons might compromise the usefulness of the registry .

18

2.2. Need for a distinction between public and Open Data (see 1.1)

Some documents have clearly distinguished public data and Open Data. Open Data in the public sector refers to data made available to citizens in open forms; public data has been defined as “anonymized, non-core reference data on which public services are run and assessed, on which policy decisions are based, or which is collected or generated in the course of public service delivery.”¹³

However, terminologies are not univocal,¹⁴ and a clarification is needed, not only in general (see 1.1), but specifically in the health domain.

2.3 Freedom of reuse and privacy

According to Open Data philosophy and policy, users should be free to use the data for whatever purpose or not. Even though this open approach belongs to the theory of Open Data, this is not always the case in practice.

¹³ HM Government, Open Data. White Paper, cit., p.8.

¹⁴ See V. Mayer-Schonberger, K. Cukier. 2013. *Big data: a revolution that will transform how we live, work and think*. London: John Murray.

With reference to privacy and data protection, some possible criteria for distinctions can be proposed: personal information; high consequence/impact; risk-based approach.

Moreover, should we collect all possible data we think we need—allowing an “appetite for data”—or not? Is there room for some “restricted use” in Open Data? Is this still Open Data? Publicly available doesn’t mean unrestrained (e.g. authentication procedures) (see 1.5).

2.4. Provide data together with preventive measures ensuring security and integrity (see 1.5)

Transparency and traceability of data—for example, by providing logs of modifications and use—as well as integrity of data, from source to use and application of data, should be provided.

Integrity is connected to the quality of data: for instance, when you know that data is going to be published, you can try to influence it.

Think about preventative measures (e.g. incentives for anonymisation), and think preventatively about potential purposes – e.g. anticipating the consequences of misuses? What type of license? Who will enforce it?

A risk-based approach is needed to security of data, especially for data controllers – important for providers and users of data to understand their responsibilities.

3. Recommendations on the different roles for institutions and the media

3.1. Open Data needs to be underpinned by Freedom of Information legislation(s) and practices to make public institutions remaining accountable through time (see 1.1)

Freedom of information laws (FOI laws) are aimed to provide access by the general public to data held by national governments. They establish an individual "right-to-know" towards government-held information. However, not all European States have adopted and/or implemented legislative acts to grant their citizens FOI laws. Also, what is the enforceability of the existing FOI laws? Which rights do citizens have to contest the full compliance of governments?

3.2. Joint commissions with different stakeholders should be established to identify priorities/topics/areas/data to be published (see 1.2)

Tensions still exist between what public institutions decide to publish and what the public needs/wants. Moreover, the gaps should be filled between data and information (and how to proceed from the former to the latter).

Forms of partnerships and training are needed (e.g. joint projects with journalists, statistical methodology specialists, NGOs for data analysis and to get to new questions).

Digital technologies could also help in finding out which information is more requested in order to speed the process of making it available (typically, health, environment, and government). However, this cannot be the only driver in delivering data sets.

3.3. Ensuring everybody the opportunity to comment and to see the comments is important

The need also exists for feedbacks from users—and not just from data producers— about the quality and usability of data.

3.4. Data should be made available in different formats in order to ensure communication

Different audiences require publication in different formats: both specialists and non-experts may require raw data, but processed information—explaining who processed the data and how it has been processed— should be made also available to citizens.

Often open portals are quite sophisticated from the technical point of view, but they are not user-friendly. Explanatory tools and tutorials should be prepared to help users.

3.5. Trust also depends on the quality of data and on the process for data production (see 1.6)

Trust is also dependent on qualities such as certification, reputation, and monitoring performed both by publics and specialists.

As to institutions, building and maintaining trust depends on a variety of factors in generating, preserving, and delivering data. The level of skill revealed by institutions in dealing with knowledge and technologies is an essential component of their being trusted by citizens.

3.6 (Data) journalists should be also be accountable and act in a transparent way to remain accountable as a trusted vigilant party in assessing the quality of institutional data

As data journalists may become a trusted party in analyzing institutional behavior and performance, and in providing feed-backs to citizens, requirements of accountability and transparency also apply to them (see 1.3).

The agenda



Open Data in Health: how knowledge may generate trust

JRC - Ispra (Italy) 18 November 2014
Building 36, Room 10

09:00	Welcome, Opening, and Introduction to the workshop – Estefania Aguilar Moreno, Monica Gemo, Mariachiara Tallacchini (JRC - IPSC) (10 min.) Open data ice breaking: Short video of Tim Berners Lee "The year open data went worldwide" (5 minutes)
09:15 - 10:45	Session 1: Open data framework: What is the current state on Open Data? What does Open Data in Health mean? (Chair: Monika Schröder - DG CNET) Rob Hagendijk (University of Amsterdam, The Netherlands) - Socio-technical imaginaries on Open Data and Health: institutions and citizens (tentative title) (25 min) Franck Von Lenep (DREES, France) - The French document on Open Data in Health and the concept of "démocratie sanitaire" (tentative title) (25 min) Chair starting the discussion and general discussion on session 1 (30 min)
10:45 - 11:15	Tea & Coffee break
11:15 - 13:15	Session 2: The European policies on privacy and Open data: conflicting or complementary? (Chair: Nicholas Nicholson – JRC-IHCP) Manuel García Sánchez (Spanish data protection agency, Spain) - Open Data in Health and Privacy perspectives (tentative title) (25 min) Magnus Stenbeck (Karolinska Institute, Sweden) The experiences of registries, Open Data and privacy (tentative title) (25 min) Chair starting the discussion and general discussion on session 2 (30 min)
13:15 - 14:15	Lunch Buffet
14:15 - 15:30	Session 3: Existing portals for Open Data, institutional duties, citizens' expectations: which roles for whom? Which role for the media? (Chair: Massimo Craglia – JRC-IES) Antonia Rana (JRC-IPSC) - Open Data websites and some reflections on raw data, usability and trust (25 min) Guido Romeo (Wired, Italy) - The role of the media and the expectations of the civil society (25 min) Chair starting the discussion and general discussion on session 3 (30 min)
15:30 - 16:00	Working Group/brainstorming session (one group per session)- Identifying the open issues, drafting key-recommendations
16:00 - 17:00	Brainstorming wrap-up from the Chairs, Plenary discussion & Conclusion remarks
17:00	End of day

21



IPSC – G06 Unit - <http://ec.europa.eu/jrc>

Abstracts and presentations

Session 1: Open data framework: What is the current state on Open Data? What does Open Data in Health mean?

Chair: Monika Schröder, DG CONNECT, European Commission, Belgium

Monika.SCHROEDER@ec.europa.eu

Monika is a policy officer in the Knowledge Sharing Unit of the European Commission's Directorate General for Communications Networks, Content and Technology (DG CONNECT), where she develops initiatives to improve cooperation on cross-cutting themes such as competitiveness, sustainability, ethics, behavioural economics and bridging research and policy. Previously Monika worked as a policy developer for the Digital Agenda, the EU policy framework for information and communication technologies (ICT).

22

Rob Hagendijk, International School for Humanities and Social Sciences, University of Amsterdam, The Netherlands

R.P.Hagendijk@uva.nl

Rob Hagendijk PhD is Dean Emeritus of the International School for Humanities and Social Sciences (ISHSS) at the University of Amsterdam (UvA). He has been trained in sociology, political science and philosophy and he has been studying problems of science, technology and society since the early 1970s. After his retirement (2014) he has continued his research work on rare diseases, pharmaceutical innovation and health policy. In this work he collaborates with experts on rare diseases at the academic hospital, AMC. He is a member of the Amsterdam Institute for Social Science Research, the UvA's Political Science Department, and The Netherlands Graduate School for Science, Technology and Modern Culture (WTMC). Internationally he has been President of the European Association for the Study of Science and Technology (EASST), (founding) member of the Harvard-based Science, Technology and Democracy Network (SDN) and collaborative editor of the journal *Science, Technology and Human Values*.

Forthcoming publication: Stephen Hilgartner, Clark Miller, Rob Hagendijk. *Science and Democracy; Making Knowledge and Making Power in the Biosciences and Beyond*. Routledge, Genetics and Society Series, 2015 (March).

Socio-technical imaginaries on Open Data and Health: Rare diseases, experimental policy making and the Europe political project

In 2012 the leaking of a draft proposal to stop reimbursement for medical costs to patients suffering from Pompe or Fabry disease caused public uproar in The Netherlands. The reasons behind the proposal were the staggering costs of health care and the low efficacy of new drugs. The responses from patients, professionals, the media, the general public and assorted politicians were so strong that the draft was quickly withdrawn, but the debate went on for another year. Remarkable in the controversy was not the amount of industry bashing and the distrust and disgust with public health insurers and public health bureaucracy. Remarkable was the importance attached to international, i.e. European collaboration with respect to the development of better and affordable therapies and drugs for rare diseases. And this is all the more remarkable given (a) strong anti-European sentiments among the Dutch; and (b) widespread concerns about 'big data' and commercialisation of privacy sensitive information. To advance European collaboration and success in clinical research and medical technology assessment requires that patient treatment protocols are shared across borders and settings. For that to work smoothly public trust is a key variable. I will argue that rare diseases provides an excellent terrain for European experimental collaboration and learning to build systems that are secure and trusted by the general public and the groups concerned. Yet, its success depends on more than creating and demonstrating technical securitization of health data and privacy. It also relates to broader socio-technical imaginaries pursued by governments and how these resonate with and play on deep seated sentiments and views, ranging from individual autonomy and identity to the political and economic organization of society and its ethical and social implications.

23

Franck von Lenep, Directorate for Research, Studies, Assessment and Statistics at the French Ministry of Social Affairs, Health and Women's rights, France

DREES-DIR@sante.gouv.fr

Franck von Lenep is Director for Research, Studies, Assessment and Statistics at the French Ministry of Social Affairs, Health and Women's rights since 2012. He graduated in economics and statistics from ENSAE (National School for Statistics and Economic Administration), attached to France's National Institute of Economic and Statistical Information (INSEE). Prior to joining the Ministry of Health, he was an advisor to the Minister of Budget and Public Accounts. Among other positions, he also worked for the French National Health Insurance Fund.

The French report on Open Data in Health, the Public Health Bill goals on open data and the concept of "démocratie sanitaire"

The National inter-scheme Information system on health insurance (SNIIRAM) is the main health data base in France. Since 2003, it has gathered billions of claims from the healthcare insurance funds and millions of patient discharges from hospitals. Initially established as an operational decisional database for the health system administrators, it has since been used more and more for public health and health service research purposes. There are growing expectations from civil society to access this data base more easily. A pluralist commission, set up by the Minister of Health in November 2013, released its final

report in July 2014, with recommendations to both develop open data in health and facilitate access to the database while ensuring adequate protection of personal information. The recent Public Health Bill, presented by the Minister of Health in October 2014, to be discussed by the French Parliament next year, includes an important set of rules on health data, which takes these recommendations into account.

**Socio-technical imaginaries on Open
Data and Health:
Rare diseases, experimental policy
making and the EU**

Rob Hagendijk
Universiteit van Amsterdam
r.p.hagendijk@uva.nl

25

Three parts

- On rare diseases and Europeanization
- On sociotechnical imaginaries and trusted institutional arrangements
- On registries, open data and trust

I. Rare diseases

- Incidence ≤ 5 people in 10.000 of pop.
- N of such diseases currently >7000 , moving to 10.000 (more rare diseases)
- More than 30 million people in Europe have a rare disease
- Most rare diseases are chronic, about 80 % have a genetic dimension
- 75% of the diseases manifest in children, of which 30% die $6 <$ years of age
- For most of these diseases there is no cure available
- Drugs and treatment are often very expensive
- Orphan designations have been created in the EU (10 years market exclusivity) to stimulate drug development \rightarrow more market admissions
- Big Pharma is increasingly interested in orphan drugs

26

The Dutch Pompe-Fabry Controversy, 2012-2013

- Public uproar and media politics
 - A plan to stop reimbursing for treatment was leaked
 - Big Pharma framed as the villain
 - But also: health insurance companies
 - The power of patient advocacy: representing misery and suffering \rightarrow media mobilization and politics
- Solution: Europe!
 - Continue reimbursement for another 2 years
 - Promote research to enhance efficacy of drugs and treatment procedure
 - *The future for improved health care is in European/international collaboration*

Building EU Policies for Rare Diseases

- EU health strategy since 1990s
- Research (FP6...) ...> e-infrastructures/ESFRI
- EMA → Orphan drugs designation (1995-)
- EUCERD (experts committee, now High Level Committee on RD)
- ORPHANET – database for researchers, but patients etc- information center, journal, 1997)
- EURORDIS (*rare disease community building: patients umbrella organization, 1997*)
- EUROPLAN (National plans for treating rare diseases, 2008)
- ERN (European reference networks, 2005 -)
- IRDiRC (International RDi Research Consortium, 2011)
- RD-Connect –(2012-2016) - infrastructures and platform for large-scale data management and patient organizations.

27

EURORDIS TV Contact Us Donate

EURORDIS The Voice of Rare Disease Patients in Europe

About EURORDIS About Rare Diseases Rare Disease Policy Orphan Drugs & Treatments Living with a Rare Disease Services to Patients Training Resources News & Events

Home Languages EN FR DE ES IT PT RU

News and Events
Find out the latest news and events from the rare disease community!

Members' news and announcements
Check out the latest news and events from our members

Rare Diseases awareness dates
Specific Rare Diseases days, weeks, months

International rare disease events
Check out the latest international rare disease events

Sign-up now!
Sign up now to make your voice heard! Be part of the campaign to make 2019 the European Year for Rare Diseases!

II. On sociotechnical imaginaries and building trusted institutions

28

Building trustworthy policy arrangements/institutions

- Strategy
 - Promote research
 - Promote drug development
 - Requires community building and mobilization
 - Patient empowerment
 - Strong play on values
 - Controlled integration of private interest in public policy
 - High input legitimacy
 - Output legitimacy...?
- Collaboration, partnership, ownership, participation
- Keep politics and market forces integrated but controlled

Sociotechnical imaginaries - defined

- “collectively imagined forms of social life and social order reflected in the design and fulfillment of nation-specific scientific and/or technological projects...Imaginaries ... at once describe attainable futures and prescribe futures that states believe ought to be attained.” (Jasanoff & Kim, *Minerva* (2009) 47:119–146)
- Co-production of science, technology and society via institutionalization
- Trusting humans, trust in technological arrangements

Sociotechnical imaginaries and trusted institutions

- How to create trusted institutions
 - Making promissory and inclusive narratives that:
 - Stress communality, inclusiveness and a better common future
 - Indicate how social and technical features are to be combined and tied to shared, basic values
 - Combine and align not just legal, moral, social, economic relations but also technological and scientific features of futures
 - Organize accountability, communication structures and inclusiveness proactively
 - Define deliverables and trajectories
- What they say and what they do: input, throughput & output legitimacy
- ST- imaginaries are different, depending on cultural history and political economy

III. Registries as open data

30

Patient registries - research infrastructures

- In Europe there are 614 patient registries on Orphanet website ((Jan. 2014)
- Of these: 77 regional, 446 national, 40 European, 74 global in scope
- 524 public, 34 private not for profit, 49 for profit, 34 undefined
- (In 2011) 490 out of 514 registries were academic, 8 patient organization, 16 private company
- EPIRARE: European Platform RD Registries
- *It is primarily about science, but also about care, support and empowerment!*

European Platform on Rare Diseases Registration – The patients view

- Combine patients' information needs with registries linked to biobanks and population data
- Include direct reporting by patients on various issues/variables: make the system interactive
- a uniform legislative framework for RDPR
- a common, public European registry infrastructure
- Controlled access to metadata, access to original data sources (registries, biobanks) via Platform
- Ethical and legal safeguards should be guaranteed
- Opt-in and opt-out rules should be in place

31

Whose Registries, Whose Open Data?

“These data, ultimately belonging to patients, must be made widely accessible to all stakeholders in a respectful and sustainable framework” (EURORDIS)

Thank you!



Ministère des finances et des comptes publics
Ministère des affaires sociales, de la santé et des droits
des femmes
Ministère du travail, de l'emploi, de la formation professionnelle
et du dialogue social

santé
protection sociale

drees
recherche, études, évaluation - statistiques

The French report on Open Data in Health, the Public Health Bill goals on open data and the concept of "démocratie sanitaire"

Franck von Lennep, Nov 18th 2014

1

33



Ministère des finances et des comptes publics
Ministère des affaires sociales, de la santé et des droits
des femmes
Ministère du travail, de l'emploi, de la formation professionnelle
et du dialogue social

santé
protection sociale

drees
recherche, études, évaluation - statistiques

Context

- The French Government strategy on open data
The French « Administrateur général des données » (chief data officer) www.data.gouv.fr
- A huge database (SNIIRAM-PMSI) gathering billions of claims from the healthcare insurance funds and millions of patient discharges from hospitals
- Civil society claiming easier access to this database
- Technical constraints
- "Loi informatique et libertés" (1978) and CNIL (Commission nationale informatique et libertés) protecting privacy

2


 Ministère des finances et des comptes publics
 Ministère des affaires sociales, de la santé et des droits
 des femmes
 Ministère du travail, de l'emploi, de la formation professionnelle
 et du dialogue social





 recherche, études, évaluation - statistiques



Expectations from facilitated access

- Démocratie sanitaire
- Public health
- Research and innovation
- Public policies efficiency
- Medical practice quality

3

34


 Ministère des finances et des comptes publics
 Ministère des affaires sociales, de la santé et des droits
 des femmes
 Ministère du travail, de l'emploi, de la formation professionnelle
 et du dialogue social

 recherche, études, évaluation - statistiques

SNIRAM-PMSI

Description of the database :

- Encompasses the whole population of France, with detailed codes and chained data for each patient :
 - Ambulatory care reimbursement claims
 - Drugs
 - Hospital stays in private and public sectors
- 3 last years
- + sample (1/100th) over 20 years
- Pseudonymisation : **but reidentification still possible**

4



Ministère des finances et des comptes publics
Ministère des affaires sociales, de la santé et des droits
des femmes
Ministère du travail, de l'emploi, de la formation professionnelle
et du dialogue social

santé
protection sociale

drees
recherche - études - évaluation - statistiques

"Commission open data"

- Pluralist commission
- Nov 2013 – Jun 2014
- Public report with recommendations to both develop open data in health and facilitate access to the database while ensuring adequate protection of personal information
- <http://www.drees.sante.gouv.fr/rapport-de-la-commission-open-data-en-sante,11323.html>

5

35



Ministère des finances et des comptes publics
Ministère des affaires sociales, de la santé et des droits
des femmes
Ministère du travail, de l'emploi, de la formation professionnelle
et du dialogue social


santé
protection sociale

drees
recherche - études - évaluation - statistiques


Projet de loi de santé (Public Health Bill)

- To be discussed by Parliament in 2015
- Among other goals : aims at reinforcing public policies efficiency and health democracy
- Article 47 : Access to health data, largely inspired by « Commission open data » recommendations

6


 Ministère des finances et des comptes publics
 Ministère des affaires sociales, de la santé et des droits
 des femmes
 Ministère du travail, de l'emploi, de la formation professionnelle
 et du dialogue social

santé
protection sociale



 recherche, études, évaluation - statistiques

=> Develop open data


- Develop anonymous databases in health
 - => Focus groups to understand users's needs
 - => Construction of anonymized databases (k-anonymity and l-diversity)
 - => reflexions on more industrialized methods
- Develop samples : towards a 1/10 sample ?

7

36


 Ministère des finances et des comptes publics
 Ministère des affaires sociales, de la santé et des droits
 des femmes
 Ministère du travail, de l'emploi, de la formation professionnelle
 et du dialogue social

santé
protection sociale


 recherche, études, évaluation - statistiques

=> Facilitate access to the database while ensuring adequate protection of personal information

- Access for :
 - Public bodies
 - Public research
 - Private research
- Technical committee with experts appreciates need for requested data
- + Public interest criterion, appreciated by a pluralist commission
- Decision by CNIL
- The bill facilitates datamatching with research data (cohorts, surveys...) or other administrative data (careers, income...)
- Tracability of requests

8

Session 2: The European policies on privacy and Open Data: conflicting or complementary?

Chair: Nicholas Nicholson, Institute for Health and Consumer Protection, DG JRC, European Commission, Italy

Nicholas.NICHOLSON@ec.europa.eu

Nicholas Nicholson has been working with the European Commission since 2002. He is currently supporting EU public-health coordination activities in the fields of registry data and healthcare quality. Prior to joining the Commission, he worked in IT-related functions in fields ranging from public health to the space and mobile telecommunications industries. He holds a degree in Physics and a PhD in Medical Physics.

Manuel García Sánchez, International Department, Spanish Data Protection Authority, Spain
mgs@agpd.es

Manuel García holds a Degree in Economics and a Master Degree on ICT Management. He joined the Spanish Data Protection Authority in 2002 with responsibilities in data protection supervision and technical advice tasks. After a secondment in Brussels, he joined the International Department of the with responsibilities in law enforcement and new technologies issues. He regularly attends meetings of formal working groups at European level dealing with privacy and data protection issues, particularly those linked with the tasks of the Article 29 Working Party.

37

Data Protection and Health Research: Challenges in the age of Big Data

The use of clinical information for health research is becoming easier as far as new technologies as well as legal and social initiatives aiming to facilitate data sharing are developing. However, enabling faster and more productive research in order to ensure better healthcare should not be in detriment of other individual's fundamental rights, notably privacy and data protection.

Open data initiatives will help to foster health research but are not exempted of risks. Research requires data that are rich and usable enough for that purpose probably at the cost of increasing the level of risks such as re-identification or the use of personal data for secondary purposes. These risks may be amplified when considering the integration of open data sources in a more powerful environment: Big Data. Data are precious not only because of their great potential for investigation but also because of its growing economic value.

Where the protection of privacy and data protection is at stake, a balanced approach is needed in order to ensure that the proper mix of technical and procedural safeguards is in place. That entails the need to introduce, as a part of the researcher toolbox, the knowledge and use of techniques aiming to better diagnose the risks at stake as well as the ways to mitigate those risks. In that sense, the use of Privacy by Design and Privacy by Default techniques could be of paramount importance. All of which without ignoring the need for well-defined and tailored use policies.

A new data protection framework is under negotiation with the aim to harmonize rules at EU level as well as to offer new ways to face challenges related to the new ways of data processing, notably the on-line activity. The possible impact on research is still to be defined but it seems that health research will benefit from a specific set of provisions aiming to facilitate personal data processing for health related purposes. Pending the final outcome of the negotiations, we should try not to forget that with great power comes great responsibility.

Magnus Stenbeck, Karolinska Institute, Sweden

magnus.stenbeck@ki.se

Since 2007 preoccupied with building research infrastructure for personal data in health and the social sciences. Former director of the Swedish Database Infrastructure Committee, secretary and expert in two government commissions on the use of administrative registers in academic research

Registry-based health research, open data, and personal integrity

In Europe, 40 women are diagnosed with breast cancer each hour. Fortunately, the vast majority of these women survive the disease. Radiotherapy decreases the risk of breast cancer recurrence but carries severe side effects. In 2013 a team of British, Danish and Swedish researchers published an EC-funded study of heart attack risk among women receiving radiotherapy treatment for breast cancer. The study showed an increased risk for heart attacks with increasing doses of radiation to the heart. The results of the study have been incorporated accordingly into radiation therapy practice. Breast cancer patients were identified through cancer registries and subsequent heart attacks through in-patients registries.

HPV infection is extremely common and around 80% of sexually active women will have been infected before age 50. However, only a small proportion will develop cervical cancer precursor lesions which in their turn may lead to invasive cancer. In many EU member states cervical screening programmes have reduced the number of invasive cervical cancers and deaths by 50% to 80%. Through connecting cancer register data with developed screening methods and research findings, a strong connection between HPV infection and cervical cancer was seen. By development of effective vaccines against cervical cancer and offering HPV vaccination to all young girls in several EU member states, a more profound prevention of cervical cancer is achieved and the health effects will grow with generations to come.

These are just two examples (of many) on how access to medical records and personal identification numbers on entire populations has made it possible to follow many individuals over time and establish important medical research findings that have been implemented throughout Europe and the world and have saved thousands of lives. This is the real benefit of the population health registers and the common personal ID number existing in the Scandinavian countries since more than 50 years.

Data availability has to be balanced against the important objective to protect individual integrity when it comes to controlling one's own personal information. This is regarded as a basic human right throughout the world. The current proposals for new European data protection legislation seeks to protect individuals from commercial and other misuse of personal information, a problem which is growing along with the technical possibilities to spread information rapidly and widely.

The measures discussed have however paid little attention to the need to use personal information to the benefit of the health and longevity of current and coming generations in Europe and the world.

Yesterday's patients donated knowledge to today's patients, and today's patients donate knowledge to tomorrow's patients. A break in the chain of donation of information would have a devastating effect on current and future patient care and disease prevention. Such a break may occur if the proposed European data protection regulation does not take the needs of health related research into consideration.

Not only research findings, but also the data on which these are based, need to be freely accessible to scientific research. But contrary to the open policy when it comes to publication of research findings, access to personal data must be tightly controlled. Modern technology may be used not only to disseminate data quickly to those who need to use them for legitimate purposes, but also to strengthen the protection against illegitimate access to the same information. Hence, there does not need to be a conflict between integrity and open data in register based health research.

Data Protection and Health Data

Challenges in the age of Big Data

Manuel García Sánchez
Agencia Española de Protección de Datos
Ispra, 8th November 2014

41

Data are precious for research.


Data also have a growing economical value.

And a market, legal or dark




"They have created lists of victims of sexual assault, and lists of people with sexually transmitted diseases. Lists of people who have Alzheimer's, dementia and AIDS. Lists of the impotent and the depressed."

The Dark Market for Personal Data.
New York Times, Oct 16 2014




Balancing medical benefits and privacy (data protection) risks is inherently complex.



A proper balance between individual rights and community rights need to be achieved.

Should public good prevail at any cost?

Is privacy a barrier for quality research?




“Everyone has the right to respect for private life in relation to information about his or her health[...] everyone is entitled to know any information collected about his or her health.”

Article 10 Convention on Human Rights and Biomedicine

Is this hampering the right to equitable access to health care of appropriate quality (Article 3)?

44



Open data:

“Open data is publicly available data that can be universally and readily accessed, used, and redistributed free of charge. Open data is released in ways that protects private, personal or proprietary information. It is structured for usability and computability.”

The GovLab, UK

Open data:

“The interaction between open data and privacy and the potential for open data to have negative consequences on privacy was a recurring theme, though this did not translate into a general consensus that the potential risks outweigh the benefits of open data or that existing data protection measures are insufficient”

'Making Open Data Real' consultation, UK

Open data and privacy risks:

- Re-identification, whether spontaneous or deliberated, may equal to discrimination, loss of autonomy, stigmatization.
- The jigsaw effect – connection to Big Data.
- It may be an impact on the right to informational self determination.
- Data quality may be also affected.

Big Data:

“a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling the high velocity capture, discovery, and/or analysis.”

IDC(2011)

Big Data:

Personal data + cheap storage + increasing
computing power + powerful analytics =

?

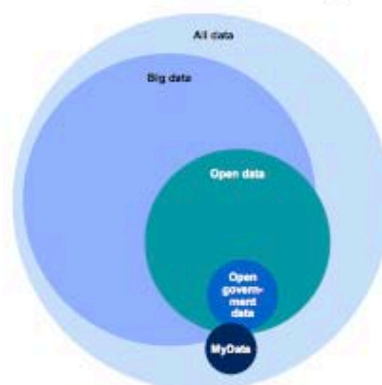
In other words:

Information on individuals is exposed to scrutiny, a condition that gives rise to concerns on profiling, stealing and loss of control.

Tene and Polonetsky(2012)

47

How open data relates to other types of data



SOURCE: McKinsey Global Institute analysis



Big Data:


“Health sector creates a big volume of data with an enormous potential. Release and reuse of this data could help to generate benefits for providers, payers, Government, citizens and researchers.”

PWC : Ten burning issues of the Spanish health care system for 2013




Tools for protecting privacy...

- **Privacy by policy**
- **Privacy by statistics**
- **Privacy by encryption**



De-identification has been proven insecure.




Nowadays, anonymization can only be assessed in terms of probability (likelihood to occur) of a successful re-identification.

Assessing the risks of re-identification:

As data linkage techniques and power computing develop, re-identification risks become unpredictable because there is a degree of uncertainty / knowledge on what data are already available or what data may be released in the future.


Assessing the risks of re-identification:

- What other data are available?
- The likelihood of re-identification being attempted – the motivated intruder.
- The likelihood of successful re-identification.




Assessing the risks of re-identification:

- The more linkable and individual-level the data are, the more limited the access to the data should be.
- The more aggregated and less linkable the data are, the more likely it is that they may be published and made available for re-use.




Assessing the risks of re-identification:

In general, release of individual-level datasets or other datasets posing a significant risk of re-identification will often not be appropriate.



Assessing the risks of re-identification:

A combination of technical legal and organisational measures can contribute to reduce the risk of re-identification: license terms, limited access, state of the art anonymisation techniques, recall of data, and, eventually, refrain from publication.



Data Protection Reform:

Directive 95/46 has been working well. But fragmentation is preventing free flow of data and giving rise to legal uncertainty. Moreover, there is a widespread perception of risk linked with on-line activity and new technological developments.

Data Protection Reform:

Different implementations of the Directive via national legislation have led to variability of interpretations in areas like definition of anonymisation, consent and research exemption, with an impact in access to patient data for research.

Working group of Confidentiality and Data Protection of the Network of competent Authorities of the Health Information and Knowledge strand of the EU Public Health Programme 2003-2008

Data Protection Reform:

The new Regulation aims to ensure an equivalent level of protection across the EU. However, there is a growing debate on the need to ensure MS the possibility to tailor the level of protection on some fields as well as to establish specific regimes of data protection. Research could be one of these fields.

Health data may be processed for:

- (a) the purposes of preventive or occupational medicine, medical diagnosis, the provision of care or treatment or the management of health-care services, [...]; or
- (b) reasons of public interest in the area of public health, [...]; or
- (c) other reasons of public interest in areas such as social protection, especially in order to ensure the quality and cost-effectiveness [...] in the health insurance system.

Art 81 Regulation – proposed text by COM

54

Health data processing:

Provides exception based on the needs of health-care as well as on reasons of public interest.

Appropriate safeguards for the rights of data subjects are requested, in particular protecting against incompatible secondary uses.

Public interest / commercial interest.

Summing up the discussion:

- > Processing of health data must be based on a robust legal framework supported by specific policies.
- > Risks at stake must be balanced with appropriate safeguards. Which and how?
- > Technical and procedural tools need to be developed. A PIA framework for Health Research ?
- > Fighting re-identification... unwinnable war? Alternatives?
- > The importance of the Ethical dimension. Public Interest boundaries? purpose compatibility? Commercial purposes? Need to know?
- > The need for transparency
- > The debate is not open data or privacy, it should be open data with privacy

Thank you!

Magnus Stenbeck
Karolinska Institute
Sweden

HEALTH RESEARCH, OPEN DATA AND PERSONAL INFORMATION

[HTTP://SIMSAM.NU/WP-CONTENT/UPLOADS/2013/04/SIMSAM_BOOKLET_ENG.PDF](http://simsam.nu/wp-content/uploads/2013/04/simsam_booklet_eng.pdf)





Register based research

- The population based Scandinavian registries contributed to these findings
- Other examples include the development of a vaccine against HPV infection, establishing breast cancer screening, improved hip replacement surgery, etc.
- Monitoring and comparisons of provider performance using registers has increased awareness and improved quality in the health care system
- These results are based on the existence of
 - many registers covering the entire population
 - a common personal identity number across all sectors of society

1328714 4

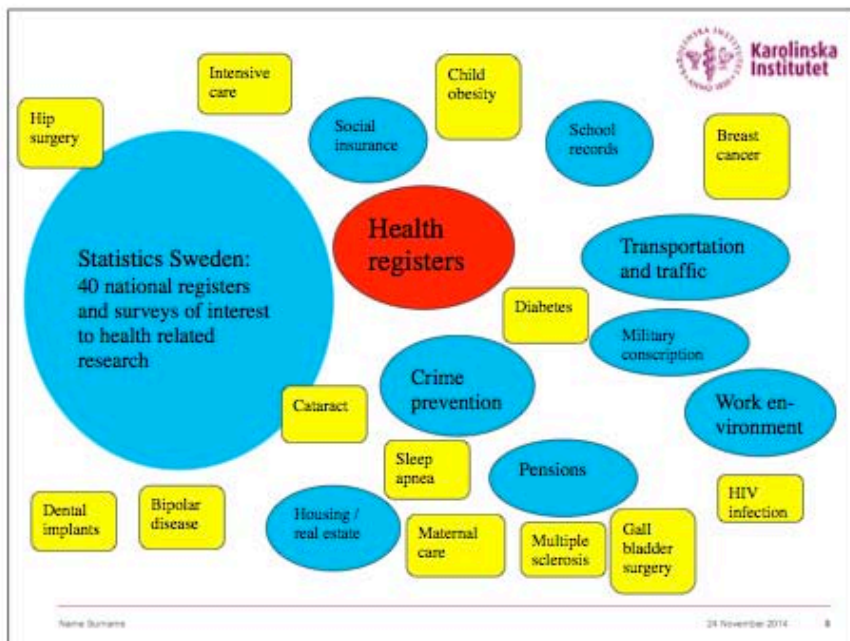
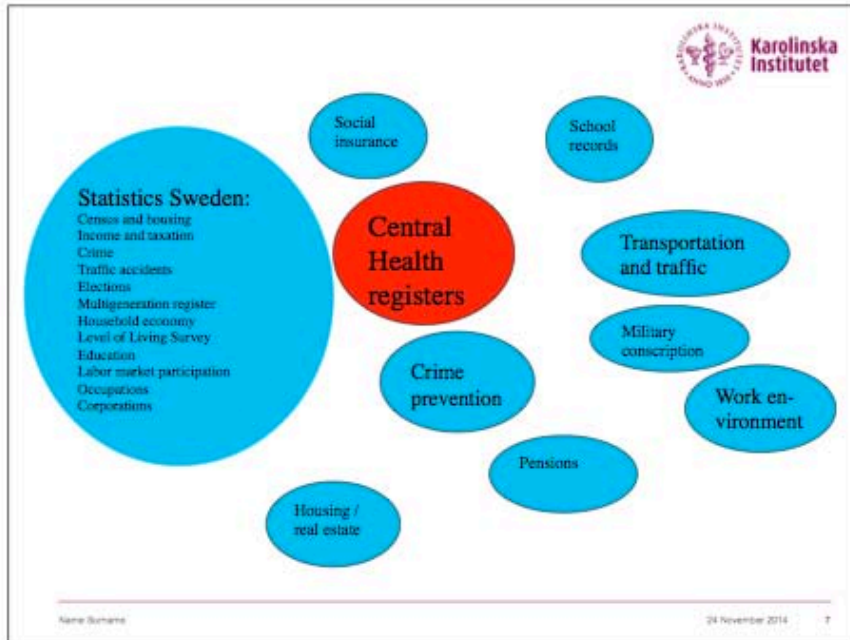
The slide features the Karolinska Institutet logo in the top right corner. The main title 'Register based research' is centered in a bold, dark red font. Below the title is a bulleted list of four items, each starting with a dark red square bullet. The fourth item includes two sub-points indicated by right-pointing arrows. At the bottom right of the slide, there is a small number '1328714' and a small square icon.

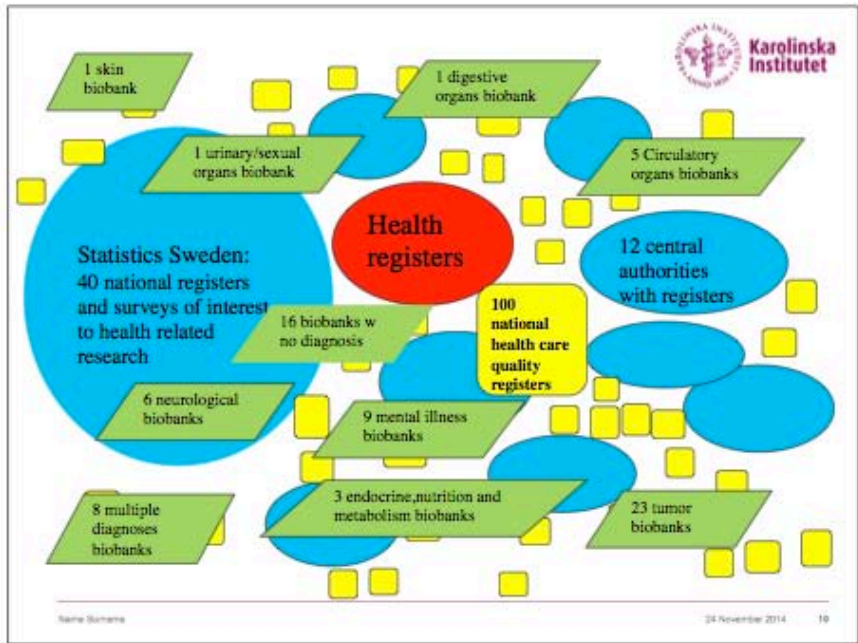
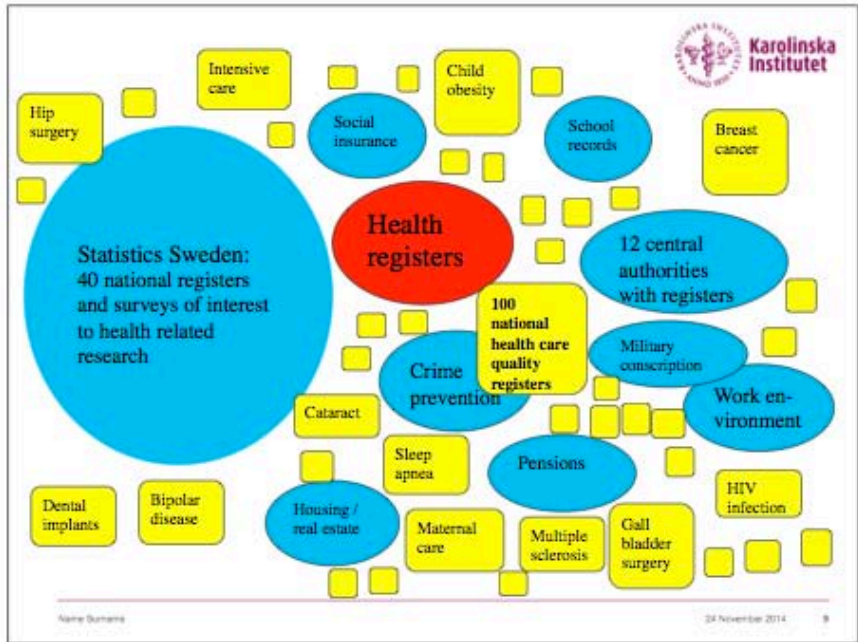
Why complete population registries with shared PIN?

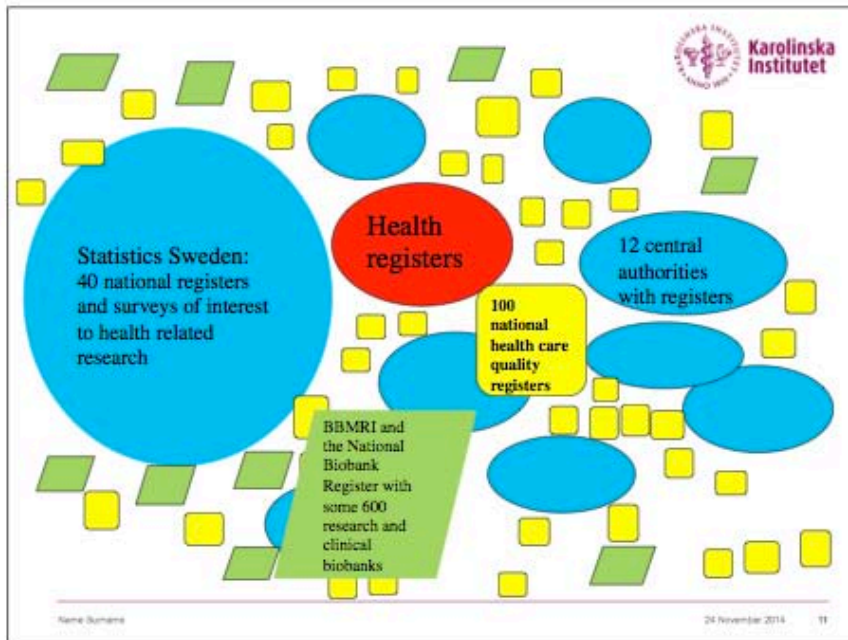
- You can combine your own sample data from the clinic or elsewhere, i.e. a sample survey with
 - background data taken from official records
 - outcome data in health registers
- You don't have to start a prospective study when the question comes up - you already have the prospective data
- With a common PIN for all sectors, you can combine information from many different sources
- You can study rare problems with a prospective design
- Hence, you can implement improvements in prevention and care much quicker and save lives earlier
- (Besides the benefits for individual patient care)

National Board of Health and Welfare

- Causes of death (1952)
- Inpatient hospital episodes (1964)
- Outpatient hospital visits (2000)
- Cancer (1958)
- Births (1972)
- Prescription drug sales (2005)
- DentalCare (2007)
- Community care (2010)







Great new potential

- A rapid increase in the number of datasources during the past 10 years
- Similar developments in the other Nordic countries
 - Norway: Helsedataprojektet, national health registries, quality registries, Biobank Norway
 - Denmark: centralization of data resources to SSI, Biobank Denmark
- Improved technical possibilities to share data
- Increased international collaboration

24 November 2014 12

Constitutional right of access to public information

- The Freedom of the Press Act (1942)
 - All documents kept by public authorities are openly available to the public
 - The right of media and the general public to examine the work of public authorities and is seen as a democratic cornerstone
 - Documents are open unless they are made secret by law
- The EU Data Protection Directive 1995 allows for this kind of openness
- The EU treaty prescribes free movement of services across borders, this should include data
- But in practice, there is no free movement of personal data

Other Swedish legislation

- The Law on Public Access and Secrecy (2009)
 - Personal data are protected
 - Sensitive personal data are more protected, but can be used in health research based on ethical vetting
- The Health Data Law (1998)
- The Patient Data Law (2008)
- Register specific legislations (some 200)
- The Law on Ethical Vetting (2003)
 - Six regional boards with high legal and scientific competence
 - The Ethics Board can waive informed consent
- The basic principle is open access but with limitations intended to protect national interests and personal integrity

But there are also problems

- Outdated database handling systems
- Outdated data sharing systems
- Outdated legislation
- No coordinated or consistent information
- Heterogeneous resources, knowledge and technology among data providers
- Lack of legal competence in the researcher community

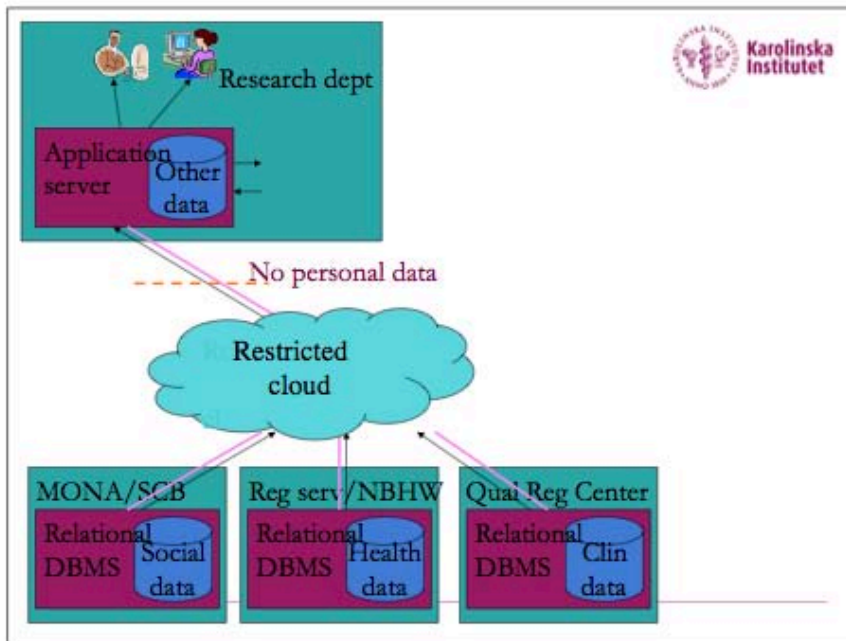
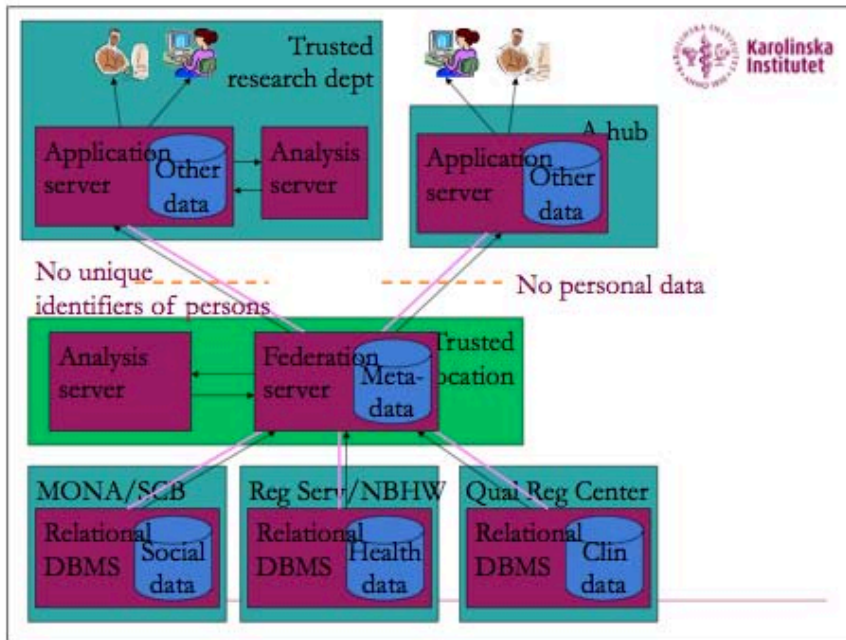
The Data Inspection in Sweden applied the outdated legislation

- The DI
 - stopped the collection of data to the national data collection effort "Life Gene" (similar to the UK biobank) which was building a general purpose cohort of 500 000 Swedes
 - prohibited the Swedish National Dataservice to service researchers with anonymized data if the PIN is still kept somewhere else
 - prohibited the research use of a national search system on biological samples in the 600 biobanks
- We needed a clearer legal basis for register based research

Government response

- A temporary law for Life Gene until 2015
- Government enquiries proposing new organisation and new rules for register based research
 - SOU 2012:36 Register data for research
 - improved information on register resources, improved collaboration on data sharing between data owners, legal support to researchers
 - SOU 2014:45 Unique knowledge through register based research
 - legal basis for building general purpose registers for research
 - strengthened data protection for research subjects
 - quicker ethical vetting for register based research
 - a national register with person identified information on where biological samples can be found
- SOU 2012:36 is being implemented now, SOU 2014:45 is reviewed by the government before putting a proposal to the parliament





Will the development be stopped by the data protection initiative in Europe?

- The parliament took a decision in March 2014. The parliament amendments to the commission proposal would end register based research.
- But the council amendments currently under way look much better for research
- A council decision may be expected within two months
- Then a trilogue between the commission, the parliament, and the council will commence.
- A new legislation requires agreement on a common text

Specific threats to register based research

- The exemption from the purpose limitation for archiving, historical, statistical and scientific use of the data was removed
- Informed consent required in every study
 - Impossible with several hundreds of thousands observations
 - Unethical in many situations when registers are used
- The right to be forgotten
 - We need complete data
 - We need to be able to reproduce the original data on which the findings were based

Specific threats to register based research

- Pseudonymization
 - Refers not only to direct but also to indirect identification
 - Removing information which permits indirect identification would render the data useless
- Time limits on the allowed period for keeping personal data
 - Threatens long term longitudinal studies
 - You do not know in advance which data will be valuable in the future

**Let us
continue
saving lives
through
research**



Session 3: Existing portals for Open Data, institutional duties, citizens' expectations: which roles for whom? Which role for the media?

Chair: Massimo Craglia, Institute for Environment and Sustainability, DG JRC, European Commission, Italy

massimo.craglia@jrc.ec.europa.eu

Since 2005 Max works at the Digital Earth and Reference Data Unit, European Commission- Joint Research Centre. The Unit is responsible for the technical coordination of the INSPIRE Directive, creating an infrastructure for Spatial Information in Europe to support environmental policy. Max is responsible for research on the socio-economic impact assessment of INSPIRE. He coordinates research projects related to the interoperability of multi-disciplinary e-infrastructures that contribute to the Global Earth Observation System of Systems. Recent activities have taken Digital Earth as a framework for evolving current e-infrastructures and integrate official data with real time data coming from citizens and sensors. This approach is being applied to develop novel indicators of Quality of Life in urban areas to include both quantitative data from sensors and qualitative information from the public

Antonia Rana, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy

antonia.rana@jrc.ec.europa.eu

Antonia Rana is a senior scientist at the Joint Research Centre of the European Commission. She holds a degree in Computer Science and a Master in Marketing and Communication Management and is a certified Intrusion Detection Analyst (CGIA 571). Her main scientific interests are related to network and systems security and interoperability, secure information exchange and communication in heterogeneous distributed systems, network and information architecture design, public-key infrastructures and information communication and marketing. During her career she has been studying different facets of communication, from the technical data communication aspect, including identity, security and privacy to the more usability/marketing oriented aspects. She has been responsible for the conformity testing and interoperability laboratory for electronic passports and communication officer responsible for external and internal communication for one of the institutes of the JRC.

Open Data websites and some reflections on raw data, usability and trust

Initiatives on open data have started flourishing recently involving administration and institution at every level and in many diverse fields. Open data is data that is made available publicly and that can be accessed, used and redistributed without any cost involved. It is released with the purpose to make institutions more transparent but at the same time data released with an open data type of license provide unique opportunities to re-use data and provide new insights into public issues, better information, new services, new solutions and also opportunities for new products. In particular in the

case of the health sector, open data has the potential to help citizens benefit from improved quality of care, lower costs and better informed choice.

In order to make open data widely available new portals are being set up which function also as aggregators or pointers to websites which publish open data. An example of these is the portal on EU initiatives on open data. The way data are made available differs considerably in terms of refinement of the data itself (i.e. pure raw data vs. information extracted from raw data by processing them in some form) and in terms of the format in which the data is published. These, in particular, can range from comma separated values (CSV), a format that can be easily imported into software that can be used to perform more or less complex processing or statistical analysis to more complex formats (such as RDF, for instance) which include, in addition to the raw data itself, metadata, i.e. data which describe the dataset and make it more interoperable to web services.

If we take the “trust” dimension, it is important to consider what elements are considered as generating trust in a website, information service or in a service provided via the web. Research in this area is often associated to research on reputation systems, a concept in which a more objective dimension can be introduced. Aspects related to the way information is presented, visualized and described have a role to play in this respect, as well as aspects related to the security mechanisms, if any, and how they are used. Not differently from information or service websites in general, recommendations related to usability and visual communication should be followed including possibly testing the functionalities in usability tests. This aspect would be particularly important if rather than raw data addressed to service providers, a website aims at publishing information resulting from new insights into and novel processing and aggregation of open data addressed to citizens.

Rosy Battaglia, independent journalist and collaborator of Wired Italy, Nòva Il Sole 24 Ore, Italy

info@rosybattaglia.it

Guido Romeo, data&business editor Wired Italia, co-founder Diritto Di Sapere, Italy

guido.romeo@gmail.com

Rosy Battaglia is an online journalist and social media specialist, and blogs on Batblog. She specialises in environmental, cultural and social investigations, and is also a trainer and consultant on Social Media. She is the inventor and curator of Cittadini Reattivi, a civic journalism project and crowdmapping site on the environment, health and law that won the Fondazione ahref contest for multimedia investigations with a high civil and social impact, in 2013. She works with Nòva Il Sole 24 ore, Wired Italia, La Nuova Ecologia, and Terre di mezzo. She has worked with Lettera 43, Movimento Difesa del Cittadino, New Tabloid, Radiopopolare, and the press agency Redattore Sociale. She is a member of the office of the president of FIMA, Federazione Italiana Media Ambientali. First prize in the La Stampa Digital Information Quality Jury awards 2013.

Guido Romeo is data and business editor of the Italian edition of *Wired* (www.wired.it). He is co-founder of Diritto di Sapere (Your Right To Know) (www.dirittodisapere.it), a non-profit project advocating an Italian FOIA, and founder of Hacks/Hackers Italy in Milan. He has started and coordinated the iData project of the Fondazione Ahref (www.ahref.eu) for advanced training in data journalism. He

graduated from the University of Bologna and holds a journalism degree from the Ecole Supérieure de Journalisme in Lille, France and a masters in communications. In 2004 he was Armenise-Harvard science-writer fellow at the Harvard School of Medicine and winner of the Astra Zeneca award for science communication. In 2007 he won the Piero Piazzano science and environment reporting award and the Amundsen prize for coverage of climate change. In 2009 he was awarded the Voltolino, Italy's most prominent prize for science reporting. For *Nòva24*, the science and technology insert of *Il Sole 24 Ore*, he managed *Città illuminate* (enlightened cities), a series of reports and conferences on development and growth in urban centres investing in innovation and creativity. He was also producer and co-host on Radio24 of *NòvaLab24*, the daily programme on research, innovation and creativity.

The role of the media and the expectations of the civil society

The availability of data and cheap processing power are pushing data journalism techniques to a new level and have allowed to develop stories unprecedented in depth and detail, especially in the area of health and policy reporting. However, access to data and its quality vary enormously among countries. While the US and Northern European countries lead the way, Italy has made advances in the last few years opening data gathered by the National Agency for regional health services (Agenas), but a lot of progress still remains to be accomplished to have a real impact on society and businesses.



Publishing open data: formats, usability and trust aspects

Antonia Rana

Joint Research Centre (JRC)

The European Commission's in-house
science service
www.jrc.ec.europa.eu

Serving society - Stimulating innovation - Supporting legislation



Open data

- Evolution of the concepts:
 - Transparency
 - Participation
 - Collaboration
- Cultural change
- Universal participation:
 - Everybody must be able to use, re-use, re-distribute open data
- No discrimination towards single categories of individuals or groups
 - (including restrictions to commercialisation or allowing use only for educational purpose)
- Many initiatives: EU Data portal, national, regional, data catalogues

Guidelines for Public Administrations websites that publish
Open Data



2



Aspects of open data

- Focus on the instruments and technologies to make open data available and usable
- Focus on the process from identification through preparation (formats) and publishing of the data (including metadata)
- Focus on processing the data and producing information/stories (data-journalism)
- Consider need to address multiple targets and processing capabilities



3



GUIDELINES FOR PUBLISHING OPEN DATA



4



Guidelines for publishing open data

- Current websites of the Public Administration:
 - preferred place for publication of open data
- Criteria for rationalisation of existing websites
- Characteristics that open data should possess
- Adapted from guidelines published by the Sunlight Foundation, Access-Info Europe, OpenGovData.org, and others.



5



Guidelines for publishing open data

Accurate: Correctly representing an event. Accuracy has a cost and must be measured against the target audience (specialist or not)

Available: Accessible over time. URIs (unique resource identifiers) can help for long-term availability

Complete: All data comprising a public data set should be published, including non-digital archival data and data used to generate aggregate or derived figures.

Primary: Collected at the source, and published with the level of granularity with which it was collected

Transparency International Georgia, Ten Open Data Guidelines, <http://transparency.ge/en/ten-open-data-guidelines>



6

75



Guidelines for publishing open data

Timely: Data is made available as rapidly as possible in order to maximize its value to the public.

Accessible: Data is available to as many users as possible, for the widest range of purposes possible:

- Easy to share digitally, with a unique and easily obtained URI
- URIs should be in a human-friendly format
- Should never require registration or payment
- Bulk downloads should be available (e.g. ftp)
- Well-documented API for automated access



Guidelines for publishing open data

Machine-readable: Format and structure that allow automated processing (e.g. CSV, JSON, or XML). Meaning of fields documented and included with the data.

Non-proprietary: Open format, not subject to intellectual property controls, structure documented (e.g. HTML, XML, etc.)

Freely usable: Free for all types of use, including commercial use, without restriction



Guidelines for publishing open data

Reviewable: Point of contact to respond to questions and complaints about the data

Discoverable: Included in appropriate data catalogues, and accessible to search engines

Each public administration entity should adopt unified means for displaying data online, so that users can rapidly locate the data

Government data portal websites should be kept accurate and up-to-date

Permanent: As data ages, it should be archived in ways that satisfy the above criteria



Quality of data

Adapted from ISO/IEC 25012 – Data Quality Model

- **Syntactic accuracy:** higher when less data input errors, e.g. "thunderstorm" instead of "tunderstorm"
- **Semantic accuracy:** higher when there are less mis-typing errors where the change has a meaningful result, e.g. "Antonia" becomes "Antonio"
- **Completeness:** the extent to which data covers all possible values in the universe in which are defined (e.g. all addresses of all schools)
- **Internal consistency**
- **External consistency** (reference to external datasets)



Data Formats

- ★ Available on the web (whatever format) but with an open licence. Data can be printed, saved, but difficult to process automatically
- ★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table) – easier to process but require proprietary software
- ★★★ as (★★) plus non-proprietary format (e.g. CSV instead of excel) – can be processed automatically and without requiring proprietary software. This is the simplest and mostly used form of open data.
- ★★★★ All the above plus, use open standards from W3C (e.g. RDF and SPARQL) to identify data objects, data have a proper URI which can be easily accessed using automated means
- ★★★★★ All the above, plus links to other people's data to provide context – e.g. data from an administration about monuments linked to data from another administration to provide an overall view of monuments in a country (or similarly health data from different regions)

W3C: Linked Open Data



11



Data formats: a different viewpoint

Raw data (★): require pre-processing (OCR, extraction, conversions, manual processing, etc.)

Structured data (★★ and ★★★): open or proprietary formats but can be easily processed in a structured way (tables) through ad-hoc programs or programming environments (e.g. R)

Highly structured data (★★★★ and ★★★★★): automatically extracted from a database and exported in an open format based on the request and automatically imported into the requesters processing system



12



Data Formats: Metadata

- Facilitate discovery
- Enable refinement of a search
- Help organise electronic datasets
- Facilitate interoperability
- Allows use of standardised common across various systems and across disciplines
- Ensure common understanding of the meaning of the data.



13



Process to produce open data



EU open data - The basics for EU data providers - ISBN 978-92-78-41122-0



14



Process to produce open data

- Level of complexity and effort required to publish open data is proportional to the number of stars
- Costs associated to open data:
 - Personnel costs
 - Costs associated to compliance with open standards
 - Risk of cyber-theft



15



Data of interest for the community/civil society?

- No general answer. If data is available and is public, open it!
- Recent analysis (Italian Association for Open Government) identified the following priorities:
 - Budget of the public administrations
 - Activity of members of Parliament and local administrators
 - Environment pollution
 - Epidemiology and health
 - Public transport
 - Crime
 - School
 - Housing



16





Trust as reputation

- Assessment based on the history of interactions with an entity:
 - Directly (personal experience)
 - As reported by others (recommendations, rating systems)
- Trust and reputation in virtual communities
 - History of member's interactions
- *How can this concept applied usually to e-commerce websites be translated in the context of institutional websites and citizens?*



19



Web design and trust

- Appearance
 - Is the website professionally designed?
 - Does it focus on content or "advertisement"?
 - Is it easy to navigate and find information (e.g. how many clicks)?
 - Is it a closed site or does it link and is linked from other websites?
 - Is it regularly updated with date of update clearly indicated?
 - What about security certificates? Are they recognized by web browsers or do they raise alerts?



20





This Connection is Untrusted

You have asked Firefox to connect securely to [redacted] but we can't confirm that your connection is secure.

Normally, when you try to connect securely, sites will present trusted identification to prove that you are going to the right place. However, this site's identity can't be verified.

What Should I Do?

If you usually connect to this site without problems, this error could mean that someone is trying to impersonate the site, and you shouldn't continue.

[Get me out of here!](#)

▼ **Technical Details**

[redacted] uses an invalid security certificate.

The certificate expired on [redacted] 01 PM. The current time is [redacted] PM.
(Error code: sec_error_expired_certificate)

▶ **I Understand the Risks**


21


Trust by certification

- The Blue button initiative
- The Health on the Net Foundation: **HON Code of Conduct** (1996) - quality assessment for online medical and health information providers:
 - 3,600 participating websites in 72 countries
 - "The Use of the Internet and World-Wide Web for Telematics in Healthcare", September 7-8, 1995,


22



USABILITY AND TRUST



23



Website usability

- Quality of a user's experience when interacting with a product (websites, software, devices, applications, etc.):
 - **Intuitive design**: easy understanding of the architecture and navigation of the site
 - **Ease of learning**: how fast a new user can accomplish basic tasks
 - **Efficiency of use**: how fast an experienced user can accomplish tasks
 - **Error frequency and severity**: how often users make errors while using the system, how serious the errors are
 - **Subjective satisfaction**: if the user likes using the system
- Mistakes:
 - Non-scannable text
 - Anything that looks like advertising (selective attention)
 - Violating design conventions (inconsistency)



24



Open data can generate new revenue

- The GovLab: Open Data 500 study, U.S.-based companies that use open government data as a key business resource.
- Increase in open scientific data, particularly data about the human genome (to improve medical care)
- Growing number of start-ups
- Combinations of open data and consumers feedback to provide information about the quality of different healthcare options
- Examples:
 - Evidera: clinical trials, and other sources to develop models predicting how different treatment interventions will affect different kinds of patients.
 - Predilytics: machine learning to help health plans and providers deliver care more effectively



25



Community-generated value

- Hackathons
 - <http://opendataday.org/>
- Using open public data to show support for and encourage the adoption open data policies by the world's local, regional and national governments.
 - Developers
 - Designers
 - Librarians
 - Statisticians
 - Citizens to bring ideas and spread the word



26



Measuring the value of open data

- Production costs vs. benefits:
 - Investment in open data programs in government agencies at all levels
 - Little research on the effectiveness in terms of increasing transparency, collaboration, or participation.
 - Large variation in open data policies makes evaluations difficult
- Risks:
 - Cyber-theft?
 - Privacy?



27



Examples on the variability

What you can do with the datasets:

- download them
- process them (through an online GUI):
 - visualize them on a map
 - create histograms
 - create pie-charts
 - build a query through a graphical query system (<variable><operation><value>)



28



Examples – initiatives of policies on open data in health

- [NHS England](#) and The GovLab at New York University have jointly created a **blueprint** – The Open Data Era in Health and Social Care – to accelerate the use of open data in healthcare
- Define metrics to measure the impact of open data and to analyze the specific circumstances under which it is most (or least) effective:
 - improve the quality of care,
 - lower healthcare costs, and
 - facilitate patient choice.
- ***"The real value [of open data] comes from interpretation, analysis, linking-up and reflection – in short, from being used,"***



29



Examples – initiatives of policies on open data in health

Data should be:

- (1) published in a standard format;
- (2) published without proprietary conditions; and
- (3) available online in a downloadable format

Open data policies → different for different kinds of information:

Open Health Data: (e.g. organograms, statistical data about workforce at NHS, expenditure data, friends and family tests, healthy lifestyle behavior, etc.)

Restricted Health Data: Data that potentially carries personal or proprietary information, anonymized through the removal of personally identifying information. Researchers or charities can request to access this data.



30



Thank you

Joint Research Centre (JRC)
Web: www.jrc.ec.europa.eu

Open Data in Health: how knowledge may generate trust

The role of the media and the expectations of the civil society

Rosy Battaglia (Freelance Journalist, Contributor Wired Italy, Founder Cittadini Reattivi, Italy)
Guido Romeo (Wired Italy & co-founder Diritto Di Sapere, Italy)

Open data in Health @rosybattaglia
@guidoromeo

The screenshot shows the ProPublica website interface. At the top, the ProPublica logo is on the left, and a subscription button is on the right. Below the navigation bar, the main heading is "Patient Safety" with the subtitle "Exploring Quality of Care in the U.S.". A sidebar on the left lists "The Story So Far" and "Feature Stories", with the top story being "Citing 'Distraction,' Quality Forum CEO Resigns Board Seats". The main content area features a photograph of Dr. Christine Cassel speaking at a podium, with a text snippet to its right. At the bottom of the page, there is a filter section showing "All (62)" items and a "Contact Us" link. A footer at the very bottom of the screenshot contains the text: "Open data in Health @rosybattaglia @guidoromeo http://www.propublica.org/series/patient-safety".

PRO PUBLICA Journalism in the Public Interest

Home Our Investigations Tools & Data Slack/Reach Get Involved About Us

Receive our top stories daily

Search ProPublica

Patent Safety

Exploring Quality of Care in the U.S.

How You Can Help ProPublica Investigate Health Care Quality

by [Olga Pivov](#) and [Marshall Allen](#)
 ProPublica, Sep. 18, 2013, 2:38 p.m.

19 Comments | | |

Connect with Facebook to share articles you read on ProPublica.

75 10 70

Too many patients suffer harm instead of healing in U.S. medicine. That's why ProPublica's reporters have investigated everything from [deadly dialysis centers](#) and [dangerous hospitals](#) to the [failure of state boards](#) to discipline incompetent nurses.

As part of our ongoing reporting, we've created a [page](#) where readers and providers can be part of the patient safety conversation, get regular updates and share stories or views.

You'll find a mix of [Nursing Home](#)

This article is part of an ongoing investigation:
Patent Safety: Exploring Quality of Care in the U.S.
 More than 1 million patients suffer harm each year while being treated in the U.S. health care system. Even more receive substandard care or costly overtreatment.

The Story So Far
 Too many patients suffer harm instead of healing in U.S. medicine. That's why ProPublica's reporters have investigated everything from deadly dialysis centers and dangerous hospitals to the failure of state boards to discipline incompetent nurses.

This page allows patients, providers and readers to join the patient safety conversation. Our goal is to find out why so many patients are suffering harm and highlight the best ways to solve the problem. Here you'll find regular updates, and places to share your stories, views or opinions.

Open data in Health @rosybataglia @guidoromeo
<http://www.propublica.org/article/how-you-can-help-propublica-investigate-health-care-quality>

90

GHOST FACTORIES
 A USA TODAY Investigation

An invisible danger in our yards

In hundreds of neighborhoods across the United States, children are living and playing near sites where factories once spewed lead and other toxic metal particles into the air. The factories, which melted lead in a process called smelting, closed long ago but poisonous lead particles can still be found in the soil nearby. Families interviewed were unaware of the dangers posed by their yards — and the government has done little to warn them, a USA TODAY investigation has found.

VIDEO **PHOTOS** **EXPLORE**

DANGER
An invisible and ignored danger

FAILURE
A failure to protect the public

TESTING
USA TODAY's soil test findings

CASE STUDY
Tyner Metals, Cleveland, OH

TAKE ACTION
How you can make your home, yard safe

EXPLORE
More than 230 former factory sites


<http://www.ajl.usatoday.com/story/nation/lead-polluting>

← PREVIOUS RETURN TO LISTING NEXT →

#Checkyourhospital

Name of the organisation: Wired Italy
 Country: Italy
 Category: Data-driven investigative journalism small media
http://daily.wired.it/mappa_miglion_ospedale/

Vote



Description of the project: Italy has many troubles, but prouds itself for having the world's second best health system according to WHO estimates. This is surely true for access to care, but is it the

<http://www.wired.com/voting/2013/11/01/11/04/1524/>



DJA 2013 shortlist


18 maggio 2013
 WIREDDAILY
 Daily
 A WIREDDAILY CHANNEL ON TWITTER

HOME ABOUT BLOG NEWS MORE

DOVE TI CURI

SCOPRI IL NUMERO DI OSPEDALI IN TUTTA ITALIA

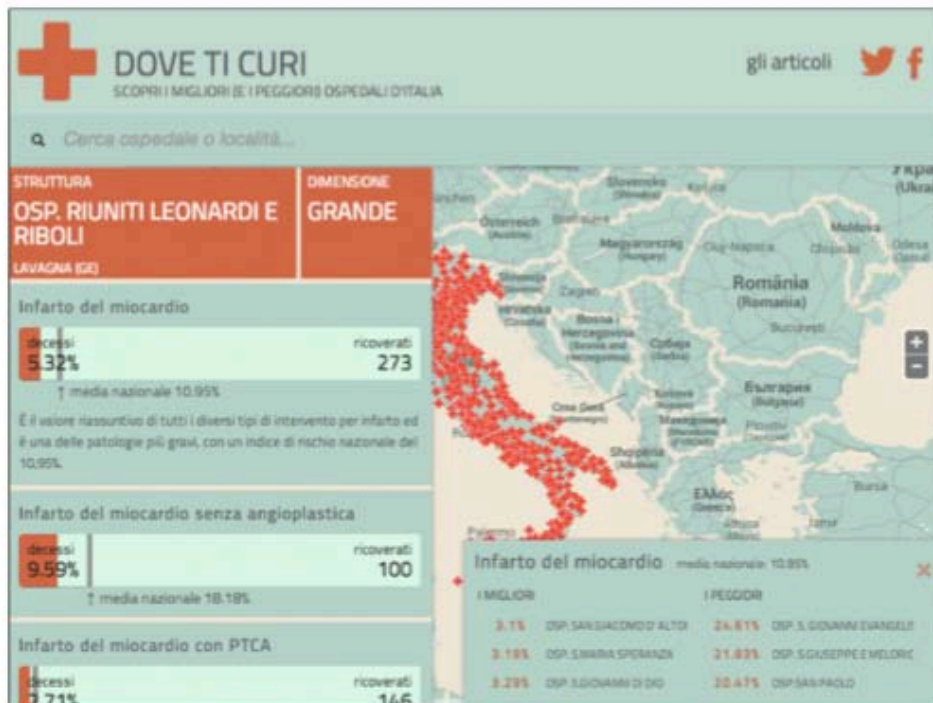
già online  



mitecube WIREDDAILY

http://daily.wired.it/mappa_miglion_ospedale/

DJA 2013 shortlist



92



Case of Impact of the COMPARATIVE EVALUATION RESULTS in ITALY



"We must not, in fact, confuse the effect that a communication option that produces the behavior of various actors in the health care system with its impact on performance.

If it is confirmed a positive relationship between public disclosure and change in a hospital, the literature is not, however, agree to establish a unique link between disclosure of results and improving the standard care.

At the root of this failure, the negative impact of adverse selection (which would otherwise offset the positive effect of deterrence) or the weak response of the citizen-patients (deterrence is not enough and you need the levers are actually activated)?"

Open data in Health @rosybataglia
@guidoromeo

93

And the Expectations of the Civil Society?



Open data in Health @rosybataglia
@guidoromeo

Credits Ernesto Bellisario <http://www.slideshare.net/ernestobellisario/datacademy-pzsmart>

Act. React. Impact.



94

Stop biocide: 100 000 citizens Naples, 16 november 2013



Open data in Health @rosybattaglia
@guidoromeo

La Spezia: Stop Veleni, 8 mars 2014



Open data in Health @rosybataglia
@guidoromeo

95

Stop biocidi: Brescia, 10 may 2014



“Information is the first form of health protection”

Pietro Comba ISS , Interview for Cittadini Reattivi 2013



Open data in Health @rosybataglia @guidoromeo

Collaborazioni e Partner Newsletter Disclaimer

cittadini reattivi
Siamo, così, attivi tutti per tutti

La bonifiche ambientali Ambiente, salute, legalità Blog Sei un cittadino reattivo? Mappati Italia

INQUINAMENTO E SALUTE

L'IMPATTO SANITARIO DELLE ATTIVITA' A RISCHIO E DEI SITI INQUINATI

ART. 41 COSTITUZIONE
L'iniziativa economica è libera.
Non può svolgersi in contrasto con l'utilità sociale o in modo da recare danno alla sicurezza, alla libertà, alla dignità umana.

Inquinamento e salute nei siti contaminati e a rischio: a Martova dal 19 al 20 settembre per capire di più

ULTIME DAL BLOG

NOI SOSTENIAMO

Open data in Health @rosybataglia @guidoromeo <http://www.cittadinireattivi.it/>

Cittadini reattivi > Project of Civic Journalism

**Health, Environment, Legality,
Open data
Citizen's Best Practices Crowdmapping**

**Stakeholders> Reactive Citizens, Committees,
Associations, Local Government Officials, Civic
Networks, Scientists**

**Speakers> Public Administration, Public Health and
Environmental Control, Companies**

**The methods and Channels> Social Network, Civic and
Citizen Media, platform, newspapers**

**Tools> Open Data Monitoring from the bottom, right of
information, Data journalism, Social Storytelling**

Open data in Health @rosybattaglia
@guidoromeo

97

Crowdmapping and Community Science Citizens



Open data in Health @rosybattaglia
@guidoromeo

Open data extrapoleted from Italian Citizens = Civic Haking



<http://www.cittadinireattivi.it/2013/09/10/taranto-la-spezia-e-brindisi-quando-sono-i-cittadini-a-contare-i-malati-di-tumore/>

Open data in Health @rosybattaglia
@guidoromeo

Collaboration between Scientists, Civic Journalists, Teachers and Young Citizens



Citizens involvement led us to many others Inquiries



Open data in Health @rosybataglia
@guidoromeo

What is missing in Italy?

Open data in Health @rosybataglia
@guidoromeo

Civil Society, Journalists and Citizens ask Freedom of Information Act



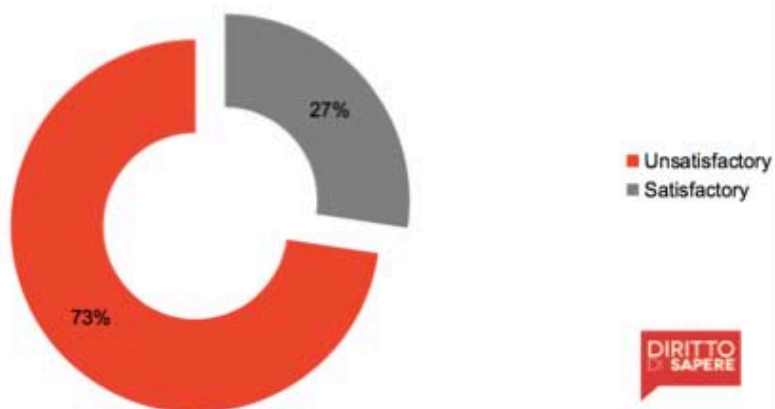
Open data in Health @rosybattaglia
@guidoromeo

100



The State give NO answers

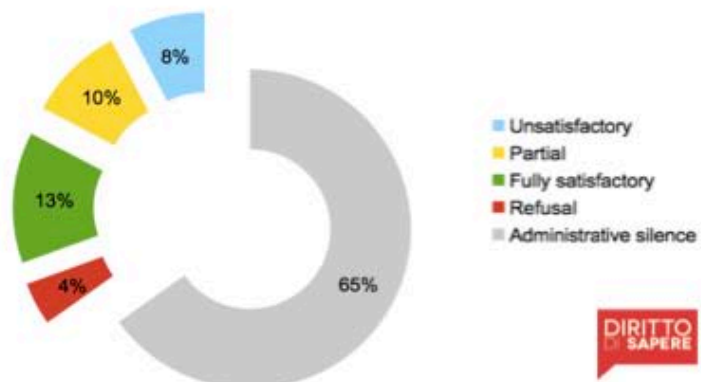
How Italian administrations answer information requests



101

Answers in detail

How did they answer your request?



manuale per cittadini e giornalisti

LegalLeaks

media e democrazia
come usare il diritto di accesso all'informazione

seconda edizione - aprile 2013



www.dirittodisapere.it

Open data in Health @rosybattaglia
@guidoromeo

Thank you for kind attention

@rosybattaglia

www.cittadinreattivi.it

@guidoromeo

www.dirittodisapere.it

www.wireditalia.it

www.foia4italy.it

Open data in Health @rosybattaglia
@guidoromeo

Participants

Estefanía Aguilar Moreno, (ex) Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy

eaguilarmo@gmail.com

Estefanía Aguilar Moreno did a stagiare at Institution for the Protection and Security of the Citizen (IPSC) at the European Commission –Joint Research Center, until November 2014. She was working on the TRUDI project about building trust in digital interactions: citizens, institutional and corporate ethics, being mainly focused on digital memories governance and its ethical implications. Graduated in Information Science and Master in Information Knowledge and Society by Universitat Oberta de Catalunya (Spain), she has Librarianship and Competitive Intelligence as professional background, having some publications about these topics.

103

Laurent Beslay, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy

laurent.beslay@jrc.ec.europa.eu

Laurent Beslay is an European Commission Official at Joint Research Centre, the Institute for Protection and Security of the Citizen (IPSC) since September 2011 and works as Action Leader of the Surveillance Technologies and the Citizen action. He manages research activities on fight against cybercrime, biometric systems and privacy safeguards. From 2004 until September 2011, he worked as Coordinator on Security and Technology for the European Data Protection Supervisor (EDPS) in Brussels. He also previously worked, for six years, for the JRC - Institute for Prospective Technological Studies (IPTS) as a PhD candidate (Electronic surveillance: benefits and risks for the EU) and as a project officer in the field of cyber-security. He holds a Post-master's degree in Global Management of Technological Risks and Crisis (University of Paris, la Sorbonne) and a Master's degree in International Relations.

Annibale Biggeri, University of Florence, Italy

abiggeri@disia.unifi.it

Annibale Biggeri is Professor of Medical Statistics at the University of Florence, and Head of the Biostatistics Unit at the Institute for Cancer Prevention and Research, Florence (IT). His research interests are Surveillance and Environmental Epidemiology. He has been member of the National Toxicological Committee and of the Council of the International Biometric Society. He also has been was

President of the Italian Epidemiological Association, and he is involved in several projects with impact on local communities – e.g. The Sarroch Bioteca Foundation; HIA of ILVA Steel Plant (Taranto IT). As President of the Epidemiologia & Prevenzione non-profit social enterprise, he is a co-promoter of Participant-led research.

Sergio Cima, Spazio A, Italy

sergiocima@gmail.com

Sergio Cima was born in 1975. He graduated in Philosophy at the University of Milan and later he achieved a master's in science communication. He is a freelance journalist and deals with health, research and science for several national newspapers and website (Corriere della Sera, Espresso, Pertecipasalute). He is member of editorial board of medical and scientific journals (Occhio Clinico, Scienza in rete). Currently he is primarily concerned with precision journalism, data analysis, data visualization and open data. As web manager in sever european project (Ecrane, Tellme) he has developed expertise in social network analysis and web analytics. It's a founding member of Spazio A, a company engaged in corporate communication , social media management and organization of cultural events. In 2014 he was one of the organizers of Sports Hackdays , hackathon on sports-related open data.

Maria Luisa Clementi, Inferenze scarl, Italy

clementi@inferenze.it

Maria Luisa is a scientific journalist, editor in chief of Epidemiologia & Prevenzione, the journal of the Italian Association of Epidemiology, working to build communication bridges between researcher-epidemiologists and communities.

Marco Crespi, Inferenze scarl, Italy

crespi@inferenze.it

Physics degree and master in science communication. I worked for ten years as editor in math magazines and I have been working since 2006 as editor and webmaster in the journal Epidemiologia & Prevenzione on the field of epidemiology and public health. In the last two years I attended the course "Digital education for enhanced editorial products" organized by Lifelong Learning Programme - Education and Culture DG - European Union, and "Data driven journalism MOOC: Doing Journalism with data: first steps, skills and tools" by European Journal.ism Centre.

Giuseppe d'Acquisto, Italian Data Protection Authority, Italy

G.DAcquisto@gpdp.it

Giuseppe D'Acquisto graduated in Electronic Engineering in 1995 (110/110 cum laude) and received the PhD in Telecommunications Engineering in 1999. After graduation, he pursued an industrial career, working for consulting companies in the area of ICT strategies. In 2008 he joined the Italian Data

Protection Authority, where he currently works as technology policy advisor. He is the Italian delegate at the Technology Subgroup of the Article 29 Working Party, and member of the International Working Group on Data Protection in Telecommunications (IWGDPT) "Berlin Group. He is the author of several publications on data protection related issues, such as data breach and security investments, the right to be forgotten, and on other regulatory topics, such as the economics of data, the net and search neutrality and the evaluation credit risk.

Silvia Deandrea, Institute for Health and Consumer Protection, DG JRC, European Commission, Italy
Silvia.DEANDREA@ec.europa.eu

Silvia obtained her degree of Medical Doctor and specialisation in Public Health at University of Pavia, and obtained her Biostatistics PhD at University of Milano in 2011. Before joining the Joint Research Centre in March 2012, she worked in healthcare quality consultancy for the Joint Commission International, in cancer epidemiology research at Mario Negri Institute of Pharmacological Research (Milano, Italy) and in population-based cancer screening programmes organisation and evaluation at Cancer Prevention Unit of Milano Local Health Authority. For the Breast and Colorectal cancer programmes she covered the role of Quality Manager and she coordinated local activities in the context of multicentre research projects. Her current research interests include quality assurance and standardisation in breast and colorectal cancer screening, cancer pain epidemiology and Bayesian methods for evidence synthesis. She is author of more than 20 articles published in peer-reviewed international journals.

105

Anders Friis-Christensen, Institute for Environment and Sustainability, DG JRC, European Commission, Italy
anders.friis@jrc.ec.europa.eu

Anders Friis-Christensen is currently employed as Scientific / Technical Project Officer at the European Commission, Joint Research Centre where he among others is working on a JRC data policy facilitating open access to data. He has a Ph.D. in computer science and a master degree in geography and computer science, and has a background in conceptual modelling of geospatial information and design of IT architectures. Until June 2013 he was employed at the Danish Geodata Agency and involved in the implementation of the European Directive INSPIRE. Furthermore, he contributed to the basic (open) data initiative under the Danish eGovernment program, primarily on analysis and design of a common data distribution platform.

Monica Gemo, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy
monica.gemo@jrc.ec.europa.eu

Monica Gemo is working as a scientific technician at the JRC since 2011 at Digital Citizen Security Unit, where is investigating participatory mobile surveillance. She holds an MSc degree in electric engineering from Politecnico di Milano, Italy. From 2002 to 2007, she was working as a research engineer at the Telecommunications Laboratory of the Catholic University of Louvain, Belgium, specializing in e-health

applications, namely standardised annotation in breast cancer screening and digital health records for emergency health care units. In 2008 she became member of the JRC to develop tools and services for real-time alerts and world press reviews from Internet sources (collation, multilingual information extraction, semi-automatic edition, trend/impact follow-up), including MediSys European Monitoring system.

Alessia Ghezzi DDG01 Econometric and applied statistics, DG JRC, European Commission, Italy
alessia.ghezzi@jrc.ec.europa.eu

Alessia Ghezzi works as assistant researcher at EC-JRC. Graduated in Humanities at the Sapienza University in Rome. Degree in archival science and master on semantic cataloguing and indexing. She has worked at the National Archives of Rome, at the Vatican Library and at the Constitutional Court as librarian and archivist and as video and audio documentalist for the RAI. Currently works on Digital Memories and ethical implications.

Ângela Guimarães Pereira DDG01 Econometric and applied statistics, DG JRC, European Commission, Italy
angela.pereira@jrc.ec.europa.eu

Ângela G. Pereira is a scientific officer of the European Commission. She has a PhD in Environmental and Social systems from the New Univ. of Lisbon (Portugal). For the past two decades she has been involved in several projects that have as key ingredients public engagement in science and technology, science communication and science governance. She is currently leading projects that study ethical issues of emerging information and communication technologies.

Hal Levin, Building Ecology Research Group, Santa Cruz, USA
hlevin6@gmail.com

Hal Levin is a Research Architect with Building Ecology Research Group, Santa Cruz, California. He is President of the Indoor Air Institute and Administrator of the International Society of Indoor Air Quality and Climate (ISIAQ). Mr. Levin has conducted research and consulted on building's impacts on occupant health and comfort as well as on the larger environment since 1978. His work has focused on the integration of knowledge about indoor and outdoor air pollution as well as other risk factors into the design, construction, and operation of buildings and communities. Mr. Levin was educated at Cornell University and the University of California, Berkeley, receiving a Bachelor of Arts degree (major in English) and professional degree, Bachelor of Architecture. He taught at two University California campuses, at the College of Environmental Design at UC Berkeley from 1978 to 1989, and in Environmental Studies at the Santa Cruz campus, 1978-1983. He was Scientist at Lawrence Berkeley National Laboratory from 2000-2005. He also held appointments at UC San Francisco, Riverside and Davis campuses UC and at the School of Public Health and at the Graduate School of Design at Harvard University. He currently works under a grant from the Alfred P. Sloan Foundation to study the microbial ecology of the indoor air environment. He is a Fellow of ASHRAE, ASTM, and the International Academy of Indoor Air Sciences. He serves on editorial boards of the journals Indoor Air, Building Research and Information and formerly

that of Building and Environment. He coined the term “Building Ecology” in the late 1970s (first published an article by that title in 1981), focusing on the dynamic and interdependent relationships between buildings, their occupants, and the larger environment. He has published extensively on various aspects of Building Ecology and has been invited to lecture on four continents.

Jan Löschner, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy jan.loeschner@jrc.ec.europa.eu

Jan Löschner has a degree in Electrical Engineering from the Technical University of Ilmenau in Germany. He works as a Scientific Support Officer for Research in the European Commission, in the Institute of the Protection and Security of the Citizen. He has more than 15 years of experience in assessing and testing of security components. He has contributed in advisory groups for international bodies like in IAEA and ISO. He is administrator of the European Root Certification Authority for the digital Tachograph.

Davide Lunardi, Bocconi University Law student, Italy
lunardi.d@gmail.com

Davide Lunardi has completed in September 2014 his traineeship at the European Commission – DG Joint Research Center, Institute for the Protection and Security of the Citizen (IPSC). During this period he was working on the institutional project TRUDI about Building Trust in Digital Interactions: Citizens, Institutional and Corporate Ethics. He is a fifth-year student in the Combined Bachelor and Master of Science (LL.B./M.Sc.) in Law at Bocconi University in Milan, focusing on International and Public Law. His academic interests focus on International and European Law, ICT Law and Human Rights Law. At the moment he is completing research for his Master Thesis about the state-of-the-art of ICT Law in Europe and in particular the legal implications of Privacy by Design.

Andrea Montanari, Università degli Studi di Roma Tre, Italy
andrea.montanari@uniroma3.it

Andrea Montanari obtained the Degree in Law at the University of Rome, Rome 3, and he carried out a PhD in Private Law with a scholarship provided by the University of Palermo, where he discussed a thesis on damages for termination of contract. During the PhD he also carried out a period of research in the UK at the University of Warwick. In the UK, after his PhD, Montanari also obtained a contract of research at the London School of Economics & Political Science where he took part in a research project on "Preserving Historic Buildings". Furthermore, he had the occasion to teach a number of courses, such as a course on Commercial Law at the University of Malta (based in Rome) and a course on Information Technology Law and Data Protection at the University of Palermo. More recently, he was offered a traineeship at the European Commission, DG for Informatics (DIGIT), where he gained expertise on EU public procurement law and practice. Ultimately, he obtained a postdoctoral fellowship from the University of Rome, Rome 3, in order to broaden his PhD research on damages for termination of contract. The outcome was the publication of a book which has been published in the Rome 3 University book series: “Il danno da risoluzione”, Jovene, 2013. Montanari is also author of various journal articles:

“Programma di clemenza e azione risarcitoria nella direttiva europea sul risarcimento del danno: convivenza possibile?” (Leniency programme and action for damages in the EU Directive on antitrust damages actions: is cohabitation possible?), in *Conc. merc.*, Giuffrè, 2014; «Lex Google»: Copyright Law and Internet Providers, future enemies or allies?, in *EIPR*, 8/2013; Questions and answers on AdWords’ cases, in *Dir. comm. int.*, 2012; “Il risarcimento in forma specifica e la rilevanza giuridica dell’attività di compensazione del danno” (Compensation in specific form and the legal relevance of activity to compensate damage), in *Europa dir. priv.*, 2013; only to name a few. Montanari is also lawyer, qualified with Rome Bar.

Susana Nascimento, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy

susana.nascimento@jrc.ec.europa.eu

Susana Nascimento is currently working at the Institute for the Protection and Security of the Citizen at the Joint Research Centre – European Commission. She is also Associate Researcher at CETCOPRA / Centre d’Etude des Techniques des Connaissances et des Pratiques in Université Paris 1 Panthéon-Sorbonne, and at CIES-IUL / Center for Research and Studies in Sociology in ISCTE-IUL / University Institute of Lisbon. She holds a PhD in Philosophy from Université Paris 1 and a PhD in Sociology from ISCTE-IUL, under joint tutorship. Her research interests are in science and technology studies, with present emphasis on digital fabrication and internet of things, environmental and social sustainability, open science and technology, transdisciplinarity and interdisciplinarity, participatory and community-based research, and social methods for technology development.

108

Giorgia Randi, Institute for Health and Consumer Protection, DG JRC, European Commission, Italy

Giorgia.RANDI@ec.europa.eu

Giorgia is Italian and has a degree in Demographic and Social Statistics obtained at the University of Padua and a Ph.D. in Medical Statistics at the University of Milan. Before starting at the Joint Research Centre in November 2013, she gained experience for more than ten years in the epidemiology field, especially on cancer epidemiology. She has mainly worked on case-control studies and meta-analyses at Mario Negri Institute in Milan and at WHO International Agency for Research on Cancer in Lyon, and on cancer registry data at the Cancer Registry of Milan. She has also worked on the acute air pollution health effects in population studies at the University of Milan. At the Cancer Information Group of the JRC, she is collaborating on the harmonization of European cancer data. She is author of more than 40 articles published in peer-reviewed international journals.

Lorenzo Richiardi, University of Turin, Italy

lorenzo.richiardi@unito.it

Lorenzo Richiardi is Associate Professor in Biostatistics and Epidemiology at the University of Turin, Italy. He graduated in medicine in 1999, specialized in biostatistics at the Milan University and received his PhD in epidemiology in 2004 at the Karolinska Institutet in Sweden. Since 2004 he has been working at

the University of Turin. His main research interests include cancer epidemiology, with specific interests in prostate and testicular cancer, occupational cancers and head and neck cancers, birth cohort research and causal inference methods. In 2005, he established the NINFEA birth cohort, a nationwide web-based cohort study. He has been teaching epidemiology and biostatistics in national and international courses, including different editions of the IEA international course on epidemiological methods and the EEPE summer school. He has more than 130 peer-reviewed publications and is a member of the editorial board of the International Journal of Epidemiology

Ignacio Sánchez Martín, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy

Ignacio.SANCHEZ@ec.europa.eu

Ignacio Sanchez is an ICT Engineer with over 12 years of experience in the field of Information Security. He is CISSP, PRINCE2 and ISO 27001 Lead Auditor. Ignacio is currently working as a Commission Official in the Digital Citizen Security of the Institute for the Protection and Security of the Citizen (IPSC) at the Joint Research Centre (JRC) of the European Commission. He is project leader for the Privacy and Data Protection Measures project and his current research activities focus on cyber security, data protection and privacy.

David Shaw, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy

david.shaw@jrc.ec.europa.eu

David has been experimenting with communication/internet tools for over 40 years. He spent a few years at CERN making antimatter, then joined the European Commission where he now is trying to make privacy matter.

Mariachiara Tallacchini, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy

mariachiara.tallacchini@jrc.ec.europa.eu

Mariachiara Tallacchini is professor of Science, Technology, and Law at the Faculty of Economics and Law of the Università Cattolica S.C. of Milan (Italy), and teaches Bioethics at the Faculty of Biotechnology of the State University of Milan (Italy). After graduating in law, she earned a PhD in Legal Philosophy at the University of Padua and has been a postdoctoral fellow at the Kennedy School of Government (Harvard University) under a NSF grant. She is currently working (2013-2015) for the European Commission at the JRC/IPSC (Ispra, Italy) in the Digital Citizen Security Unit on topics at the interface between ICT and health. Her interests concern the legal regulation of science and technology, particularly in the field of life sciences, the patentability of biotechnological inventions, the issues of scientific uncertainty and the law, the relationships between science and democracy.

Benedetto Terracini, University of Turin
benedetto.terracini@fastwebnet.it

Benedetto Terracini is Professor (retired) of cancer epidemiology and of Medical Statistics at the Università degli Studi di Torino. Since the 1960s he has pioneered cancer epidemiology research in Italy. In 2003 he was awarded the John Goldsmith prize of the International Society of Environmental Epidemiology, and in 2014 he was given the annual International prize of the Collegium Ramazzini. Between 2000 and 2010 he has been the Editor-in-chief of the peer-reviewed journal "Epidemiologia e Prevenzione." He is still active in the fields of asbestos- related diseases and pediatric tumours, but his current interests also concern risk communication and the relationships between experts and non-experts in research with human subjects.

Lucia Vesnic- Alujevic, Institute for the Protection and Security of the Citizen, DG JRC, European Commission, Italy
lucia.vesnic@jrc.ec.europa.eu

Lucia Vesnić-Alujević is a Postdoctoral Researcher at the European Commission's Joint Research Centre in the IPCS (Ispra, Italy) in the Digital Citizen Security Unit. She obtained her PhD in Communication Science from Ghent University based on the research on political communication on social media in Europe. Her research interests include EU policies, online communication and social media. She has authored several peer reviewed research papers

Europe Direct is a service to help you find answers to your questions about the European Union
Freephone number (*): 00 800 6 7 8 9 10 11
(*): Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

It can be accessed through the Europa server <http://europa.eu/>.

How to obtain EU publications

Our priced publications are available from EU Bookshop (<http://bookshop.europa.eu/>),
where you can place an order with the sales agent of your choice.

You can obtain their contact details by sending a fax to (352) 29 29-42758.

European Commission
EUR 27091 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: OPEN DATA IN HEALTH: how knowledge may generate trust

Authors: Estefania Aguilar Moreno, Monica Gemo, Nicholas Nicholson, Antonia Rana, Mariachiara Tallacchini,

Luxembourg: Publications Office of the European Union

2014 – 110 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series –ISSN 1831-9424

ISBN 978-92-79-45348-9

doi: 10.2788/421226

Abstract

Developed within the context of the JRC project on Trust in Digital Interactions (TRUDI), the Workshop on “Open Data in Health: how knowledge may generate trust” (Ispra, 18 November 2014) aimed to investigate some general issues surrounding Open Data in the EU normative perspective, reflect on institutional and civic imaginaries about Open Data, and identify how trust between institutions and citizens can be improved in health matters. The workshop encompassed three sessions, each asking a different set of questions: a) What is the current state-of-the-art on Open Data, and what does Open Data in health mean? b) Are European policies on Privacy and Open Data conflicting or complementary? c) Do existing portals for Open Data meet citizens’ expectation and which role can be envisaged for the media?
This report provides a summary of the topics and arguments presented at the Workshop and offers some recommendations in the still unfolding field of Open Data in Health.

JRC Mission

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new methods, tools and standards, and sharing its know-how with the Member States, the scientific community and international partners.

*Serving society
Stimulating innovation
Supporting legislation*

