Conceptualizing a National Data Infrastructure for Switzerland

Beat Estermann[1], Marianne Fraefel[1], Alessia Neuroni[1], Jürgen Vogel[1]

[1] Bern University of Applied Sciences, Switzerland;
corresponding author: beat.estermann@bfh.ch

Abstract

A national data infrastructure (NDI) provides data, data-related services and guidelines for the re-use of data as an easily accessible service to citizens as well as to public and private organizations. As such, it allows the efficient sharing of data between data providers and data consumers, supports new business models, and is thus a key enabler for the digital economy, societal collaboration and political processes. Several European countries have already taken steps in the direction of setting up such data infrastructures that reach across institutional silos. The approaches taken vary across the countries. In Switzerland, activities in this direction are at an early, conceptual stage. In this paper we present the findings of a preliminary study undertaken in order to identify key stakeholders, to foster a shared vision of what an NDI is about, to identify the major governance issues related to its inception and sustainable exploitation, and to sketch out a tentative roadmap in view of the establishment of a NDI in Switzerland. We carried out a series of stakeholder interviews with representatives of public administration, research, civil society, and the private sector, covering topics such as the collaboration across the data value-chain, the conceptualization of a NDI, as well as the governance issues and collective action problems that need to be resolved, such as the identification of key stakeholders and their expected contributions during the build-up phase.

Keywords: national data infrastructure, big data, open data, linked data, mydata, infrastructure resources

1   Introduction

As a recently published OECD study on data-driven innovation observes, data play the role of an infrastructure resource in that they generate value when used as inputs into a wide range of productive processes the outputs of which are often public and non-market goods (OECD 2015). Just as electricity, streets and water are core infrastructures that serve citizens, companies and governments alike, so too a **national data infrastructure (NDI)** is a community-wide need in the big data era, similar to education, healthcare, and public services (Shin 2007).

A NDI provides data, data-related services, and guidelines for data-reuse to public and private organizations and to citizens. As such, it allows the efficient sharing and effective use of data across organizational boundaries. It is beneficial for data providers and consumers alike, supports new business models, and is thus a key enabler for fostering the digital economy, enhancing societal collaboration and supporting political processes.

Fostering data-driven innovation requires the attention of policymakers. There is a **need for a coherent data governance** that balances openness and public benefit on the one hand, and privacy and data security on the other (OECD 2015). Good data governance not only needs to address the question of who gets access to what data for what purpose and under what conditions, but should also assign responsibilities and retribution mechanisms for data maintenance and enhancement in order to ensure the sustainability of the common data infrastructure. Furthermore, it is necessary to conceive a NDI in a way that ensures interoperability and alignment with international approaches. The European Commission (EC) perceives data policy to be one of the key issues to enhance the long-term competitiveness of Europe on a global scale. Data and the corresponding infrastructures for business, administration and society are thus an essential part of the European Digital Single Market (DSM) Strategy (cf. EC 2016a). Only recently, the Swiss government has adopted a **"Digital Switzerland" Strategy** that among other goals envisages the establishment of a NDI, the promotion of an equitable information ecosystem, the empowerment of Swiss citizens so that they have control over their own data, and the formulation of a coherent and future-oriented data policy for Switzerland (Federal Office of Communications, OFCOM 2016).

As the governance issues related to establishing and maintaining a NDI are still far from being fully understood by the key stakeholders in Switzerland, the Bern University of Applied Sciences, in cooperation with the opendata.ch association, carried out a preliminary study on the topic. The main goal of the study was to foster a **shared vision among**

key stakeholders regarding the data infrastructure that is to be established and to draw their attention to the **governance issues** surrounding it. To this end we conducted a first series of interviews with stakeholders from various sectors in order to analyze prevailing views on the notion of an NDI and to identify open questions in that respect. The study contributes to the current debate regarding the formulation of a future-oriented data policy. At the same time, it lays the foundations for a larger research project the goal of which is to address the governance issues related to the establishment of a NDI in more detail.

2   Relevant Research in the Field

Over the last couple of years, several authors have addressed governance issues related to the establishment and the maintenance of a NDI or similar data infrastructures (e.g. Nugroho et al. 2015, Jetzek 2016, Neuroni et al. 2016). It appears from these publications that a multi-disciplinary approach is needed in order to gather a holistic view of the challenges at hand, drawing on different strands of research in disciplines such as political and social sciences, law, economics, and computer science. This section summarizes present considerations in the different strands of research relevant to the discussion of governance issues surrounding a NDI.

In order to clarify the notion of a NDI we set out by drawing on literature related to (data) infrastructure development. Infrastructures for data sharing and re-usage are built for different areas of application. It is therefore important to view them in relation to context. From a socioeconomic viewpoint, recent developments in big data analytics are of particular importance. One of the characteristics of big data is the growing variety of data that is being integrated and tapped into for all sorts of analyses. In that respect, technological trends and challenges should be taken into consideration, such as the ones related to linked data. When it comes to analyzing personal data, privacy issues are a major concern (Jarchow / Estermann 2015), a topic that should be approached both from a technical and a legal perspective. And finally, we give a short overview of contributions regarding the notion of data governance at a societal level, which may serve as an umbrella to integrate the different research perspectives.

2.1   Existing notions of data infrastructure

Neuroni et al. (2016) explore the notion of a national data infrastructure and related concepts, observing that "the notion of a national data infrastructure is not straightforward, but bears connections and overlaps with different concepts dedicated to infrastructure development in a digital environment. This includes concepts on developing e-government infrastructures, national information infrastructures or open data infrastructures" (Neuroni et al. 2016). They also note that, irrespective of the given focus of interest, there is usually no common understanding of what a data infrastructure comprises. In fact, various authors stress different aspects and operate different distinctions in order to come to grips with the notion of a data infrastructure, such as Jetzek (2016) arguing that a distinction should be made between an **IT infrastructure** and a **regulatory infrastructure**.

There are at least three strands of research and practice that can be drawn upon when it comes to conceptualizing a NDI: e-government research, research on open government data (OGD), and the initiatives in various countries regarding the establishment of a "national information infrastructure" (NII) or a "national data infrastructure" (NDI).

In the context of e-government, the development of **shared infrastructures enabling interagency collaboration** is a core issue when it comes to improving public service provision (Bekkers 2009, Fraefel et al. 2013). As Neuroni et al. note, "the relevance of interoperable data and information sharing is also stressed in the European Interoperability Framework (EIF), which provides a conceptual model for public services and considers data-related services as a basic component for service provision. The focus is on base registers that are legally controlled and maintained by public administrations and provide authentic sources of information on items such as persons, companies, buildings, or roads." (Neuroni et al. 2016). According to the European Commission (EC) this information ought to be made available for re-use, and appropriate mechanisms for managing access and control should be put in place (EC 2010). This is consistent with the view that authentic data sources are a key enabler for enhanced service provision as they allow improving user experience and administrative efficiency (de Vries 2012, Capgemini et al. 2015).

In the context of OGD, the focus is on infrastructures that allow third parties to **re-use government data**. As Neuroni et al. note, "OGD benchmarks assess data availability and accessibility, user support and functionalities for stakeholder. Availability of data mostly relates to coverage of sectoral data (education, health, etc.), while the provision of basic government data (key registers) is hardly a topic" (Neuroni et al. 2016). Also, the focus of OGD strategies generally is on public sector data that can be made available in an open access regime.

In contrast, NIIs or NDIs can comprise data from various types of data providers, such as government, businesses, or non-profit organizations. Also, they may be openly licensed, be made available to a group of specific stakeholders only, or be closed. As Neuroni et al. observe, "several countries have adopted the concept of a NDI/NII in order to effectively **share core government data sets within and outside government** and stimulate their use across boundaries in order to

create value. Thereby it can be observed that the adopted initiatives or policies conceptually strengthen connections between government data held in base registers and OGD. This is the case for instance in the UK, in Denmark, and in the Netherlands" (Neuroni et al. 2016).

No matter whether the notion of a data infrastructure is approached from the point of view of e-government research, OGD, or the creation of a NDI/NII, a data infrastructure can be understood as a community-wide need, respectively a good that should be provided as a public service (cf. Shin 2007). Data itself has become an essential driver of innovation, and it can be considered as **one of our society's central infrastructure resources** (OECD 2015). As Frischmann observes, from an economic point of view, infrastructure resources are fundamental resources that don't get consumed when being used and generate value when used as inputs into productive processes. As their outputs are often public goods that generate positive impacts for the society, it may be socially desirable to manage them in an openly accessible manner. Data meet the following characteristics of infrastructure resources: first, they may be consumed in a non-rivalrous fashion for some appreciable range of demand; second, the social demand for data is driven primarily by downstream productive activities that require them as an input and third, they may be used as an input into a wide range of (private, public or social) goods and services (Frischmann 2006 and 2012).

2.2   Potentials and policy implications of big data

The public sector is an important data producer and thus an important source of information for big data analytics. Public sector sources, such as classifications and statistics need to meet high quality requirements, are often produced on a regular basis, and are therefore an essential resource for providing stability to big data models (Washington 2014). In turn, big data technologies may change the way official statistics are produced (UNECE 2016). It is widely acknowledged that governmental data-provision may contribute to **socioeconomic development and growth** (cf. OECD 2013). In 2014, the European Commission has released its vision of a data-driven economy, addressing the relevance of big data and calling for the adoption of open data policies and the development of public data infrastructures (EC 2014).

A the same time, the public sector is also a data user and as such may benefit from the implementation of big data concepts. **Opportunities** comprise: increasing efficiency and saving costs, increasing transparency, innovating service provision, supporting and improving decision-making, detecting and preventing mistakes and frauds, reducing security threats, and conducting policy impact assessments (Eckert et al. 2014, BRZ 2015, EC 2016b). Typical domains include traffic planning and control, public security and intelligence services, social welfare and health care, disaster management or strategic management. Depending on the application domain and the type of data generated, big data analysis in the public sector is closely related to the concept of smart cities (Chen/Hsie 2014, Neirotti et al. 2014).

There are several accounts of practical cases of **(big) data analytics in the public sector** (Kim et al.2014, Eckert et al. 2014, Thapa 2016). As scholars point out however, many government data projects are by now rather focusing on structured and stored data, instead of using e.g. real-time, or in-motion data (cf. ibid., Desouza/Jacob 2014). Current gaps in exploiting the potential of big data in the public sector are also acknowledged by the European Commission, who in 2015 has launched an action on big data for public administrations. Its main objectives are to "*identify concrete big data and open knowledge opportunities and requirements in public administrations and in specific policy contexts*" and to "*promote cooperation (…) in order to accelerate the data-driven transformation*" (EC 2016b).

Big data analytics not only offers opportunities for the public sector, but also raises challenges (Chen/Hsie 2014, Höchtl et al. 2016). The first set of challenges relates to **big data governance**: establishing a data-driven decision-making culture (create public value through big data), developing agreements for integrating data sources across organizations to generate big data, and specifying roles and responsibilities for managing big data. The second set relates to **big data implementation**: recruiting specialized personnel (data scientists) and implementing adequate technological solutions. The third challenge relates to **risk management**: Ensuring privacy in view of data aggregation (individual profiling) and ensuring security (combination of disparate data sources).

Irrespective of the field of application (private vs. public), big data raises a range of **policy issues**. Both, the OECD and the European Commission stress the need for providing adequate framework conditions for big data, including policy frameworks "*for the generation, collection, transport and use of data*" (OECD 2013). Regulatory issues in particular relate to the areas of user empowerment and privacy protection (ibid., EC 2014).

2.3   Potentials and challenges of linked data

One of the corollaries of big data is the breaking-up of data silos, i.e. the growing integration of data across various data-sources. This has also been recognized by the open data movement. Thus, while the call to open up public sector information can be seen as a logical extension of the freedom of information regulations that have been adopted by many countries since the 1990ies, the open data movement is also driven by a technical and economical vision: a *semantic web* is to be created by linking many 'open' datasets from various sources. Thus, 'linked open data' will serve as an infrastructure resource for third parties to build value-added services on top of it, such as new combinations of

data, visualizations, or other data-driven services (Berners-Lee 2009, Bauer & Kaltenböck 2011).

When combining data from different sources, a difficult obstacle lies in the need to first understand the particular encoding (syntax) and meaning (semantics) of the individual data fields, which are often not well-documented and require substantial domain knowledge. Even worse, for the same type of data, different sources may use different syntax and/or semantics. As a result, larger projects with several data sources require labor-intensive and error-prone data integration. As case studies in different industries reveal, most integration projects suffer from syntax-related issues and encounter semantics-related problems (Schmidt et al. 2010). In order to overcome these challenges, experts and OGD researchers advocate publishing data as so-called **linked data** (Berners-Lee 2009, Gschwend et al. 2015).

A vital prerequisite for linked data is the availability of standardized metadata in the form of **ontologies** (Hitzler et al. 2010). While a substantial number of ontologies have been defined by standardization bodies and academic communities, many domains still lack appropriate ontologies. Also, serializing the data as linked data is difficult to automate, and induces considerable overheads in terms of the **storage and processing power** required to handle the data. On top of that a number of fundamental technical challenges remain: How to trace **provenance** on such platforms where pieces of data originating from various sources can be arbitrarily combined? – Even though RDF triples often contain a fourth piece of information to identify their source (or context), tracking such source information throughout the query processing pipeline is still an open research challenge (Wylot et al. 2014). Another fundamental problem that attracts the attention of researchers is the lack of **data integration** in data lakes, where multiple data types can coexist but cannot be interrelated or joined automatically (Demartini et al. 2012, Duggan et al. 2015). And finally, **user-friendly tools and software components** are required to provide the typical functionalities of the data lifecycle to model, create, store, secure, publish, search, retrieve, integrate, process, and visualize data originating from a NDI. Depending on the chosen data format and data provisioning platform, mature commercial solutions may exist (e.g. SQL), while for others mature, integrated solutions are still lacking (Auer et al. 2012).

2.4 Privacy protection and informational self-determination

The main legal issue in the context of big data is the **protection of personal data** in the face of possible reindividualization of previously anonymized data due to the recombination of various datasets (Weber 2016, Ohm 2009). With a steadily growing number of successful **de-anonymization** cases (Shapiro 2015, Sweeney 2000), the common practice to anonymize data sets before analysis is increasingly being called into question by public authorities and within research (Horvath 2013, Graham et al. 2014). One possible reaction is to qualify more data as personal data in the first place, thereby impeding re-usage and potential business cases. Another way of handling the problem is to develop new methods to ensure privacy beyond anonymization. IT research currently pursues **new approaches**, such as the combination of **decentralization** and **encryption**: The aim is to develop solutions that enable big data analysis without prior decryption of data, thus reducing the risk of (unintentional) data exposure by data re-users (Lindell/Pinkas 2009, Hayes 2012).

Recent developments in the area of big data have also led to calls to make data protection a basic requirement regarding the design of technical solutions. The **"*privacy by design and by default*"** principle has been included in the new data protection regulation of the EU which will enter into force in 2018. The European General Data Protection Regulation (OJEU, Regulation 2016/679) implements a new right to **data portability** according to which "*the data subject shall have the right to receive the personal data concerning him or her*" (Art. 20). The "democratization" of personal data is also subject of current initiatives under the label of "Midata" or "MyData" that aim at establishing platforms, which allow individuals to manage and share their personal data (cf. Shadbolt 2013, Hafen et al. 2014, Su et al. 2016).

2.5 Data governance at a societal level

From a cross-organizational perspective, governance can be perceived as "*the means by which to infuse order in a relation where potential conflict threatens to undo or upset opportunities to realize mutual gains*" (Williamson 1999). Thus, governance refers to the rules, roles and behavioral patterns which configure the way collective action occurs. The collective action problems identified in political science and institutional economics can roughly be attributed to four main types: free riding/opportunism, blockage, congestion, and depletion (Estermann et al. 2009). From an economic point of view, collective action problems tend to lead to underproduction or to excessive consumption of a good or a resource, putting at stake overall utility. Collective action problems vary in function of their context; one of the major influence factors are the characteristics of the goods or services concerned (cf. Ostrom 2003, Frischmann 2006). All these factors have an influence on the effectiveness of governance structures in a given setting and need to be taken into account when deciding about public or private provision and on the access regimes of the various elements of a NDI.

While company-wide data governance frameworks have been a topic in information science for decades (e.g. Weber et al. 2009), the big data era confronts us with the same task, but worldwide and in a setting where power structures are

less regulated. Data governance as overarching perspective therefore needs to focus on the key elements that are relevant for a cross-boundary common view of the reality, and the underlying architecture models need to support a vision that satisfies all relevant stakeholders (Ross et al. 2006).

Seizing the benefits from data driven innovation requires the willingness to collaborate. In order to create public and economic value, incentive systems for cooperation and collective action need to be developed, covering the entire data life-cycle (Bekkers 2009, Fraefel et al. 2011, Cordella/Bonina 2012, Harrison et al. 2012, Millard 2016). Coherent policies are needed to encourage investments, promote sharing and reuse, and reduce barriers to cross-border flows that could interrupt global data value chains (Jetzek et al. 2014, Jetzek 2016). The elements to be considered include data access and reuse, portability and interoperability, linkage and integration, quality and curation, "ownership" and control, motivation and responsibility, and data value and pricing (OECD 2015). Thereby it is important to strike the right balance between the social benefits of enhanced reuse and sharing of data, and individuals' and organizations' concerns about such openness (OECD 2015, Janssen/van den Hoven 2015).

Experiences show that non-technological design rationalities are as important as technological issues (cf. Bekkers 2009, Klievnik et al. 2016). The roadmap for a NDI in Denmark for example is built on fundamental choices regarding governance models (how to keep all stakeholders inside and outside government committed and on board), financing models (mitigating the fact that benefits of open data rarely accrue with the data holders making the effort) and active outreach to potential data (re-)users (recognizing them as important stakeholders), and lets those three aspects mutually reinforce each other (Jetzek 2016).

3   The Present Situation in Switzerland

In Switzerland, a series of steps have already been taken in view of the establishment of a NDI:

- Since 2003 a cooperative effort has been undertaken between the Confederation, the cantons and further public and private stakeholders in order to create a National Spatial Data Infrastructure[1].

- In 2013, the Confederation launched a pilot portal for open government data, and the year after, the Federal Council adopted an Open Government Data Strategy. The open government data strategy at the federal level has been complemented by similar strategies at cantonal and municipal levels – however still by far not by all the cantons or cities.

- In 2015, the Rectors' Conference of Swiss Universities launched a pilot Open Research Data Platform[2], while the State Secretariat for Economic Affairs (SECO) launched the LINDAS linked data service[3] permitting structured data from various sources to be integrated and networked using a cross-organisational approach, and commissioned an observatory for digital single market infrastructures.

- In 2016, the Confederation launched the official opendata.swiss portal[4] for government data and data of state-affiliated businesses throughout Switzerland (Weber 2016). It also published its "Digital Switzerland" Strategy (OFCOM 2016) which, among other things, calls for a "coherent and future oriented data policy in Switzerland", including "regulatory coordination on a national, and where necessary international basis, in relation to access to data and digital content as well as data processing and security", for the "construction of a national data infrastructure" to help the "open data principle permeate the entire administration, the research sector and parts of the private sector", and for the implementation of the "fundamental right of an individual to exercise self-determination with regard to information" by giving Swiss citizens control over their own data.

Besides that there is a range of data infrastructures where practices of data sharing among different administrations occur, but so far there is no overarching view or strategic vision in this regard.

---

[1] http://www.e-geo.ch

[2] http://openresearchdata.ch

[3] http://lindas-data.ch

[4] https://opendata.swiss

4   Research Questions and Methodological Approach

In spring 2016, the Bern University of Applied Sciences, in cooperation with the opendata.ch association, carried out a preliminary study in order to explore how key stakeholders in Switzerland conceive of a NDI and the governance issues surrounding it. The two research questions at the center of the study were:

1. How should a national data infrastructure in Switzerland look like?

2. How should inter-organizational collaboration and participation be organized in view of the establishment and the maintenance of a national data infrastructure?

In order to address these questions, 16 semi-structured interviews were carried out with key stakeholders from various backgrounds in Switzerland, representing the perspectives of public administration, state-affiliated enterprises, private enterprises, research, civil society, and politics. The interview guide comprised questions on five over-arching themes: data-based value creation, elements and stakeholders of a NDI, governance of a NDI, informational self-determination, and next steps in view of the establishment of a NDI in Switzerland. The interviews were exploratory in nature and aimed at understanding different stakeholder perspectives. The interview partners were selected from a list of potential interlocutors based on a stakeholder analysis according to the typology proposed by Mitchell et al. (1997), as this would at the same time allow us to identify the core stakeholders that are needed to further advance the establishment of a NDI in Switzerland. The 16 interviews were complemented by an international comparison of approaches towards national data/information infrastructures in various countries and by eight further interviews – six with stakeholders in Switzerland, and two with international experts – that allowed us to cover several issues in more depth. Note that neither the comparative study nor the eight additional interviews that also form part of the preliminary study are covered in this paper.

The following step in our research design consists in playing the results of the preliminary study back to the stakeholders of a NDI in Switzerland in order to gather reactions to them and to formulate questions for further inquiry. The preliminary study thus constitutes the first iteration in a longer research process to be carried out according to the design science methodology (with elements of action research, comprehensive prototyping and co-production).

5   Findings

5.1   Assessment of the status quo

In order to set the grounds for discussing issues related to the establishment of a NDI in Switzerland, the interviewees were asked to position their organization / themselves along the data value-chain and to assess the current situation regarding secondary data use. As was to be expected due to the selection process, many interviewees indicated that their organizations fulfill multiple roles as data providers, data users and as intermediaries. For some of them, one of these roles clearly prevails, while for others the roles are context-dependent. And finally, the sample included three interviewees who are neither data providers nor data users, nor typical intermediaries, but are mainly involved in setting the framework conditions (e.g. through lobbying or as members of parliament).

When asked to assess the current interplay between stakeholders of the data value chain, interviewees put **positive** emphasis on the following aspects: Governmental data provision in selected areas (geospatial data, statistical data, and environmental data) is appreciated, in particular if made available through central portals in a coordinated manner (international, federal). Several stakeholders, especially from the private sector, have established processes in place for re-using data directly from third parties (government, others) that generally work well. Interviewees also pointed to the favorable conditions in Switzerland regarding the basic communication infrastructure and legal certainty as prerequisite for exploiting data value chains. The ongoing political dialogue on data-related issues is considered as important and the role of the open data community as catalyst for raising political awareness is appreciated in that respect.

Stakeholders also pointed out some **negative** aspects of the current situation: Governmental data provision shows deficiencies with regard to data accessibility with respect to heterogeneous data formats and pricing models as well as with regard to data interoperability across organizational, federal and national boundaries, resulting in unnecessary effort and expenditures for data users. From the point of view of governmental data providers, it is not always clear who is to fund the activities necessary to change this situation. Interviewees also pointed to a lack of knowledge about existing data value chains (who uses which data for what purpose?). Related to this, there are still observable reservations towards opening up data on the one hand and uncertainties concerning data access on the other.

Accordingly, interviewees also identified **room for improvement**. There is a broader understanding that the full potential of the secondary usage of data hasn't been tapped yet and that associated opportunities and risks haven't been understood well enough yet, neither in the public nor in the private sector. The adoption of an ecosystem perspective is considered essential to enhance value generation through data re-usage (e.g. through data combination). Also, several interviewees see a need for a broader social discussion and for raising public awareness, especially on data privacy issues

and data ethics in general (e.g. informational self-determination). According to them, politics should develop a coherent position on the provisioning of open government data (including the clarification of the financial aspects), as clear conditions and rules would be favorable both to data provision and usage.

5.2   Understanding of a national data infrastructure

In order to reach a better understanding of the way various stakeholders conceive of a "national data infrastructure, interviewees were asked to state their opinion on the idea of considering data as an "infrastructure resource" (cf. section 2.1). While it was undisputed that data are a highly valuable resource that was prone to be re-used by third parties, identifying them as an "infrastructure resource", similar to the electricity grid, the water system, roads, railroads, etc. was not straightforward for all interviewees. This explicitly or implicitly raised a discussion on the lines along which types of data should be differentiated when conceptualizing a NDI. When it came to deciding which data should belong to a NDI, various interviewees applied different criteria (cf. table 1), which obviously also had an impact on their view of the necessary governance mechanisms and the question to what extent data provision is to be considered as a public responsibility.

Differentiating characteristics of data				
Provenance	State	State-affiliated enterprises	Private enterprises	Private individuals
Production context	Public	Enterprise		Personal
Accessibility	Open	Shared		Closed
Data protection requirements	Particularly sensitive		Non-sensitive	
Purpose of use	Functioning of the state		Specific purposes	
Scope	Possibly all data		Demand-oriented data	

Table 1: Criteria used to differentiate the data potentially within the scope of a NDI

Participants were also invited to point out relevant elements of a NDI. Again, the notion of a NDI was not equally familiar to all interviewees. While some had a rather clear vision of what a NDI is, others would have preferred to be presented with a pre-defined concept. Some interviewees adopted an overarching perspective; others rather put a focus on specific areas of interest. In sum, the interviews provided broad input with regard to a common understanding of a future NDI that appears to converge on the following definition:

*A **national data infrastructure** is a nationwide (distributed) **technical infrastructure** (comprising portals, platforms, services, etc.) that allows the access to and exchange of data on the basis of **predefined rules**.*

The interviews also made apparent however that speaking of "a" data infrastructure or of a "national" data infrastructure raises certain concerns and bears some potential for misunderstanding. The following differentiation provides clarification: A national data infrastructure is **not a monolithic block** (but has a distributed architecture) and it **does not exist in isolation** (but is inter-connected at the international level). On an abstract level, a distinction was also made between conceiving a NDI as some form of tangible result in terms of technical solutions and data provided (**technical artifact**) and its ideational value as a concept that helps to consciously shape developments related to the digitalization of society with the participation of various stakeholders (**boundary object**).

The above definition puts the focus on the first understanding. In that respect, participants also provided concrete input on what a NDI should or could comprise and what issues its implementation would entail. As for the different types of data, there is a common understanding that government data are at the core. Correspondingly, provision is at least in part considered to be of public responsibility. However, the question of which other types of data should be made available via a NDI and what responsibilities arise from this requires further clarification. Further potential elements of a NDI that need to be assessed in a coordinated way are presented below (cf. section 5.4).

Most interviewees would agree that the **purpose** of a NDI is to support data-driven value creation and in particular to help realize the potential value of existing data. The interviewees were asked to state if and what kind of **benefits** they could gain from a future NDI. The following potential benefits were most salient:

- efficiency gains (access to more data, standardized data, inter-organizational exchange of data and shared infrastructures);
- better services for users (new or improved services thanks to new insights or better integration of information across organizational boundaries);
- an improved image of government agencies and state-affiliated enterprises;
- reduced costs related to data acquisition.

For several interviewees however the concrete purpose and use of a national data infrastructure still needs clarification, particularly from the point of view of private enterprises. Some suggested to develop the business case for a NDI and to

analyze its economic and societal benefits in more depth, e.g. on the basis of use cases.

5.3 Key stakeholders and their roles

In order to explore the perceptions different stakeholders have regarding the key agents and their roles in setting up a NDI, the interviewees were asked what contribution their organization could make to establishing a national data infrastructure in Switzerland and what they would expect from other stakeholders in that respect. Again, answers ranged from clearly formulated expectations to rather reluctant statements. The results underline that the implementation of a NDI requires the participation of various stakeholders with well-defined roles, some of which are already rather clear, while others remain to be clarified. There were no notable gaps with regard to self-perception and external perception of the various stakeholders in this regard.

Overall, the following key roles emerged: the interviewees expect from **politics** to provide the necessary framework conditions and to issue mandates to public administration. **Public administration** in turn shall provide and publish data, foster the debate, play a coordinating role, and contribute to the setup of the technical infrastructure. **Civil society** is expected to foster the networking and the dialogue between stakeholders, while **academia** is expected to provide data and infrastructure components, whereas the exact role of **state-affiliated** and **private enterprises** still needs clarification.

5.4 Fields of action requiring coordination

As mentioned in section 3, Switzerland doesn't start from zero when it comes to establishing a NDI, and as the new "Digital" Switzerland Strategy as well as ongoing debates on issues related to data policy illustrate, time is ripe to advance its conceptualization and implementation. As the interviews indicated, the creation of a NDI should be seen as a step-by-step endeavor that combines a bottom-up and a top-down approach: While publishing first datasets, involving citizens and private companies, and developing use cases can be done by many government agencies in a **bottom-up** approach, the strategic framework should be provided in a **top-down** manner, ensuring that it is backed by political decisions providing a clear mission to the government agencies involved in the establishment of a NDI. In order to be effective, **the two approaches need to be synchronized**; several interviewees pointed to the fact that civil society ought to play the role of a mediator between the two.

Based on the stakeholder interviews, seven fields of action could be identified that need to be tackled in a coordinated manner. The most important issues raised in the interviews are presented in table 2; they have been categorized according to an extended version of the European Interoperability Framework's interoperability levels (EIF) (EC 2010).

Dimensions	Fields of action requiring coordination
Political	Legal foundations / legal mandate (clarification of government agencies' responsibilities)
	Coordination (across sectors and federal levels)
Ethical	Ethical norms regarding the use of personal data
Legal	Regulation of data protection and security
	Regulation of the usage of government data (e.g. fees, conditions limiting its usage)
	Regulation of data provision (attribution of tasks, regulation of liability issues)
	Regulation of the use of personal data
Economical	Funding for data provision / basic investment
	Business models for data refinement
	Competitive situation of state-owned enterprises
	Data monetization (as an incentive for data provision)
Organizational	Access regimes (open vs. club model, private data)
	Data-lifecycle management
	Processes (data provision, data usage, etc.)
	Knowledge management, sharing of know-how, support to data providers and users
Semantical	Shared ontologies
	Metadata
Technical	Basic infrastructure (for the storage, transport, and processing of data)
	Infrastructure components (data portals, platforms and interfaces for service provision)
	Identity and access management
	Knowledge management functionalities

Table 2: Fields of action that require coordination

Many interviewees also expressed their views on how this coordination should take place. Thus it has been observed that the dialogue and the establishment of a 'code of conduct' among the different stakeholder groups are at least as

important as legal regulations. There is no agreement as to the right mix between incentives and legal obligations when it comes to getting government agencies to publish their data in form of open data. Several interviewees expressed the view that public agencies need to go ahead with regard to opening up their data, setting an example in order to encourage other players (e.g. private enterprises) to follow suit. Others stressed the importance of successful use cases, as only demonstrable usefulness of providing data via a national infrastructure (as open or shared data) could create the momentum necessary to trigger a self-carrying dynamism. And finally, in view of ensuring governance in a collaborative context, it was suggested to elaborate several possible variants regarding the overall objective before deciding on a common roadmap.

5.5 Role of research / requested knowledge base

The interviews were also used as an opportunity to get better insights on the knowledge base needed by key stakeholders to engage in conceptualizing a NDI and in order to identify potential research gaps. Needs of information touch upon various fields of interest (cf. table 3). In particular, participants stated that further considerations on establishing a NDI should be based on an ecosystem perspective, and that the cost-benefit ratio of a NDI should be assessed in more detail, e.g. based on use cases. Also, there was a strong interest in issues related to data ethics and promising developments related to strengthening privacy.

Focus	Fields of interest
Ecosystem	Multiple perspectives on the ecosystem (governance, legal, organizational and technical issues) International best practices and implications for Switzerland Approaches to defining responsibilities in a multi-stakeholder context Assessment of existing cooperation models (public / private) Assessment of existing models for data brokerage Ways of ensuring the public value of a NDI (e.g. alignment between providers and users of data)
Economic aspects	Analysis / provision of evidence on the benefits of a NDI Assessment and prioritization of potentials for value creation Overview and assessment of business models (along the range of types of data and access regimes: from personal to public, and from closed to open) Approaches to balancing investments and benefits in a multi-stakeholder context (public / private) Assessment of adequate pricing models for data re-use Assessment of models for remunerating re-use of personal data Assessment of suitable (cooperation) models for operating data infrastructures (central / decentral etc.)
Ethical aspects	Permissibility / legitimacy of linking disparate data (especially with regard to personal data) (Non-)legitimacy of analysis / validity of conclusions (Limits of) acceptance within society (big data, privacy) Definition of adequate basic conditions for the analysis of personal data and ways to enforce them
Privacy	Overview of the range of personal data traces and (technical) possibilities for generating insights Overview of recent developments in the field of privacy at an international level Overview and assessment of mechanisms aimed at securing control over personal data Overview and assessment of technological possibilities required to install trusted third parties as intermediaries regarding the provision of / access to personal data
Technology	Overview of the technological possibilities and challenges of a NDI Clarification of differences between network infrastructures and data infrastructures Assessment of options for organizing data management, e.g. the (dis-)advantages of centralization Assessment of different design options and their (dis-)advantages Overview and assessment of approaches to increase data quality in complex domains (sectorial data) Requirements towards the technical infrastructure regarding the management of big volumes of data Assessment of the potential of blockchains for managing / regulating data access (from a technical and legal perspective)

Table 3: Fields of interest related to NDIs

6 Discussion

6.1 Four perspectives on a national data infrastructure

As the review of the existing literature and the interview results suggest, the notion of a NDI should be approached from four distinct perspectives (see figure 1):
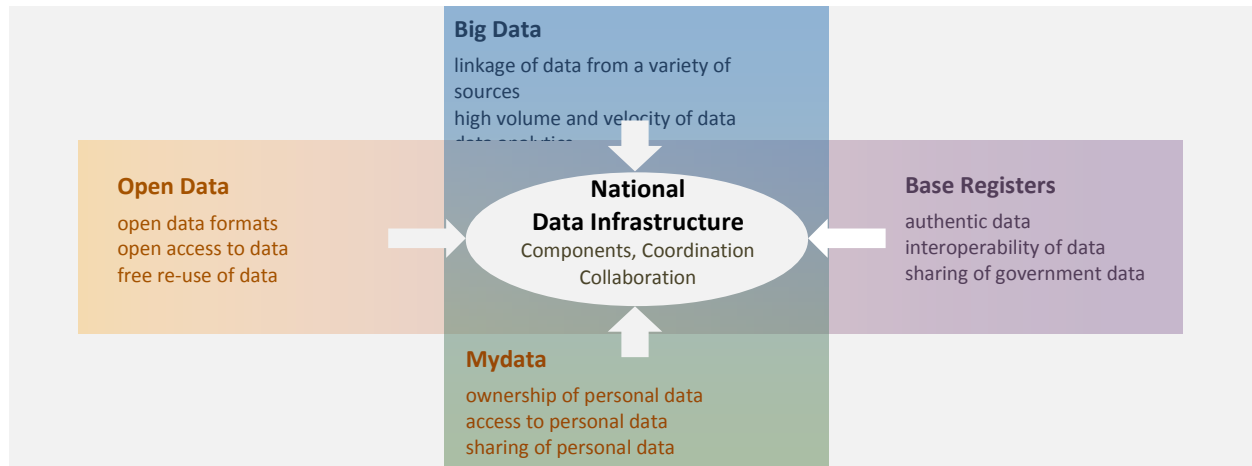

Figure 1: Four perspectives on a "national data infrastructure"

- The **big data perspective** stresses the role of novel analytical methods, spurred by an ever increasing quantity of data collected and facilitated by ever larger data storage capacities and a trend towards breaking up data silos. Data analysis requires access to all sorts of data (authentic, open, personal).

- The **base register perspective** focuses on data that is collected on the basis of a specific legal mandate and that is trusted, authentic, and under access control (cf. EC 2010). Infrastructures for basic government data have long been in place, are shared among authorized parties, but are in most cases not open to third parties.

- The **open (government) data perspective** highlights the quality of "openness" of the data on a technical and legal scale. In recent years, national, local, and sectoral infrastructures (e.g. spatial data) have been built to provide open government data. Evidence on the impact is still limited.

- The **mydata perspective** stresses the re-use of personal data in combination with an increased control of individuals over their own personal data (cf. Poikola et al. 2015). This implies a paradigm shift in the area of data protection, requires a special focus on societal acceptability, and calls for new technological solutions.

On an abstract level, each perspective shows a number of peculiarities with regard to types of data, existing infrastructure components as well as prevailing and/or emerging technologies. Also, in accordance with its target stakeholders and its focus, each perspective will bear specific policy implications. Reaching agreement over the right mix of the four perspectives and the pace at which each of them shall be developed will be crucial when it comes to advancing the cause of a NDI. For this purpose it will also be useful to reflect on the **interactions and inter-relations between the different perspectives**. Thus, adopting the base register perspective may re-focus the efforts in the field of open government data on areas that are particularly prone to lead to efficiency gains both within public administration and among private enterprises. As the Danish example shows (cf. Jetzek 2016), this may indeed unlock the discussion about the actual (economic) use of open data. Furthermore, it may spark reflections about the pooling of data as "shared data" among various public authorities, possibly even with the participation of private enterprises. At the same time, it can be expected to spur the discussion about the re-use of personal data, which is also at the center of the mydata perspective. Here, however, the preoccupation is not primarily with efficiency gains, but with ethical questions and with ensuring informational self-determination. In fact, the informational self-determination and the empowerment of users endorsed by the mydata perspective is two-pronged: On one hand it is about civil liberties, about the right to know and to regain control over who is doing what with one's personal data. On the other hand, it is about economic rights, namely the right to put one's own personal data to profit – be it by releasing it for re-use for a social cause or by trying to monetize its use by profitable enterprises in the context of big data applications. And finally, this two-pronged nature of the mydata perspective shares similarities with the open data perspective's two-fold objective of ensuring transparency and good governance and of unleashing the slumbering (economic) potential related to the re-use of existing data within public administration and beyond thanks to new developments in the area of data analytics.

6.2   Informational self-determination as a salient aspect of a NDI

As noted above, the mydata perspective, i.e. the issue of informational self-determination, has emerged as a salient aspect of a NDI, which in this form hadn't been mentioned in the existing literature on national data infrastructures. The issue is highly topical and the object of an unfolding political debate, where many stakeholders haven't adopted a clear position yet. As a recent study on big data in Switzerland shows (Jarchow/Estermann 2015), there is broad support for a **new paradigm in the area of personal data**, which consists in giving individuals greater control over the data that pertains to them. Giving individuals a '**right to a copy**' of the personal data held by companies and organizations would be a first step towards this new paradigm. The issue has been debated in Swiss parliament and is currently subject of a study mandated by the federal government. A similar development can be observed within the EU, where the European General Data Protection Regulation (EU 2016) implements a new right to data portability according to which "*the data subject shall have the right to receive the personal data concerning him or her*". A further step promoted by some would consist in giving individuals some form of **property right over the personal data concerning them**, empowering them to exercise greater control over what use third parties are allowed to make of their data. In view of the implementation of a NDI the following questions should be asked:

- What role do personal data play in the context of a NDI?

- What approach should be taken towards data anonymization in the face of a growing risk of de-anonymization?

- To what extent could a NDI serve as a catalyzer for user empowerment through the provision of platforms for the management of personal data and the creation of trustworthy intermediaries between holders and potential re-users of personal data?

- What has a NDI in store for private individuals (e.g. in terms of a 'right to a copy' or in terms of monetization of one's personal data)?

- How to balance the different interests of data-collecting organizations, the public (open data), and individuals (mydata)?

6.3   Implications for further research activities

As has been pointed out by previous research and has again been underlined by the present study, the establishment of a NDI clearly must be tackled from an **interdisciplinary perspective**. The research in the diverse fields of interest from the point of view of key stakeholders is not equally developed. While some areas are fairly well studied (e.g. pricing models) other research fields are only emerging (e.g. systematic analysis of the impact of open data, new privacy mechanisms). Overall, the interviews suggest that there is a need to strengthen **knowledge transfer** from research to practice and vice versa. Furthermore, there are opportunities to learn from the experiences in other countries; in the case of Switzerland this is particularly the case with regard to some countries' focus on base registers that hasn't received much attention so far by Swiss authorities. Table 4 gives an overview of the **questions to be addressed by further research**, organized along the dimensions of the extended version of the European Interoperability Framework's interoperability levels.

7   Conclusions and Outlook

7.1   How should a national data infrastructure in Switzerland look like?

The findings regarding our first research question can be summarized as follows:

A **national data infrastructure** can be conceived of as a nationwide (distributed) **technical infrastructure** (comprising portals, platforms, services, etc.) that allows the access to and exchange of data on the basis of **predefined rules**. It is **not a monolithic block**, but has a distributed architecture, and it is **not confined to Switzerland**, but is inter-connected at the international level. Also, the notion of a NDI should be approached from **four distinct perspectives**: a big data perspective, a base register perspective, an open data perspective, and a mydata perspective. Each of the four perspectives shows a number of peculiarities with regard to types of data, existing infrastructure components as well as prevailing and/or emerging technologies. Also, in accordance with its target stakeholders and its focus, each perspective will bear specific policy implications.

With regard to the data a NDI shall comprise, there is general agreement that non-sensitive government data and research data shall be part of it. With regard to other types of data, such as data from private enterprises or private individuals, further clarification is needed.

Dimension	Research questions
Political	What shall be the main purpose of the NDI? What are the key aspects of a coherent data policy for Switzerland?
	What type of data shall a NDI in Switzerland provide and which infrastructure elements are needed to this end?
	Which elements are to be provided as a public service?
	What are the potentials and requirements with regard to data sharing and using big data analysis in the Swiss public sector?
Ethical	In which areas does big data analytics tread on new grounds that require new norms regarding the handling of (personal) data?
	How should new norms regarding the handling of (personal) data look like? What are key elements of a code of conduct that respects basic rights and personal freedoms and is in the interest of the public good?
Legal	What legal provisions are needed in view of the establishment and the maintenance of a NDI?
	How can requirements related to user empowerment and privacy protection be addressed from a legal point of view?
Economical	What are appropriate financing models and access regimes for the different elements of a NDI?
	To what extent are the setup and the maintenance of a NDI of public responsibility? To what extent do its economic characteristics justify its conceptualization as a public good?
	How can collective action problems in inter-organizational, collaborative settings be overcome?
	How can the chances that open data has a greater impact be improved, within public administration and outside?
Organizational	Which are the different roles in the NDI ecosystem and particularly, what is the concrete role of the state when it comes to establishing and maintaining a NDI?
	How can a viable NDI architecture be set up, taking into account existing infrastructure elements and organizational requirements?
	How can good data governance over the entire data life-cycle be ensured?
	How can trustworthy intermediaries best be set-up in view of giving users control over their personal data (mydata)?
Semantical	How can data integration be improved on a semantic level in a sustainable manner?
Technical	What is the best way to ensure the tracking of data provenance in a distributed environment?
	What technical measures should be taken to enhance user empowerment and privacy protection?

Table 4: Questions to be addressed by further research

7.2 How should inter-organizational collaboration and participation be organized in view of the establishment and the maintenance of a national data infrastructure?

There is general agreement that a step-by-step implementation process combining bottom-up and top-down elements is most likely to lead to success. There is also general agreement as to the role of several key stakeholders: the parliament is expected to provide the necessary framework conditions and to issue mandates to public administration. Public administration in turn shall provide and publish data, foster the debate, play a coordinating role, and contribute to the setup of the technical infrastructure. Civil society is expected to foster the networking and the dialogue between stakeholders, while academia is expected to provide data and infrastructure components and to feed back current research developments into practice. At the same time, it also appeared from the interviews that the role of state-affiliated and private enterprises still needs to be clarified.

The study has also allowed to identify a set of concrete issues on which coordinating action is needed. It has been shown that such coordinating action is needed along seven dimensions: political, ethical, legal, economical, organizational, semantical, and technical. Given the prominence of the issues related to open data, personal data, and informational self-determination, the central role of data governance shouldn't been underestimated when it comes to balancing the different interests of data-collecting organizations, the public (open data), and individuals (mydata). Reaching a good balance in this respect is one of the keys to the success of a NDI. Another one is the demonstration of its usefulness through the example of concrete use cases.

7.3   Next steps

Our research project is contributing to the discussion of a highly topical issue in Switzerland: the blueprint of a national data policy. The focus is on key stakeholders within the big data ecosystem in their diverse roles as data providers, data users and analyzers, legislators and regulators, as well as providers of public goods and guarantors of security. With view to the economic expectations and societal concerns associated with big data, developing a concept of a NDI is first of all a means to define the frame conditions within which development can be driven forward. This includes the clarification of roles of the different stakeholders, the definition of rules and decision rights, as well as the implementation of appropriate accountability mechanisms.

In close collaboration with key stakeholders we will further develop a common vision of a NDI and identify the need of actions from a governance perspective. The intended result of the envisioned follow-up project is a governance framework and a draft roadmap of prioritized actions and policy recommendations for the future Swiss NDI. As the preliminary study has confirmed, a multidisciplinary approach is needed with regard to tackling these issues. While there is plenty of existing research to be drawn upon in the various domains, the real challenge lies in painting a holistic picture, taking into account the interactions and the interdependencies between the various dimensions and perspectives. The preliminary project has allowed us to stake out the ground and to develop an analytical framework that can be used as a boundary object in further interactions with and among stakeholders.

Acknowledgements

References

- Auer, S. et al. (2012): "Managing the life-cycle of linked data with the LOD2 stack." In: Cudré-Mauroux P. et al. (eds.) (2012): The Semantic Web – ISWC 2012.
- Bauer F., Kaltenböck, M. (2001): Linked Open Data: The Essentials: A Quick Start Guide for Decision Makers. SEMANTIC-WEB COMPANY.
- Bekkers V. (2009): Flexible Information Infrastructures in Dutch E-Government Collaboration Arrangements: Experiences and Policy Implications. Government Information Quarterly, 26, pp. 60-68.
- Berners-Lee T. (2009): Putting Government Data online, https://www.w3.org/DesignIssues/GovData.html.
- BRZ (2015): Big Data in der öffentlichen Verwaltung. White Paper, Bundesrechenzentrum GmbH. https://www.brz.gv.at/downloads/studien/2015-03-30_Big_Data_in_der_oeffentlichen_Verwaltung_v1.2_pub.pdf?55m9h7
- Capgemini et al. (2015): Future-proofing eGovernment for a Digital Single Market. DOI: 10.2759/32843.
- Chen Y.C., Hsieh T.C. (2014): Big Data for Digital Government: Opportunities, Challenges, and Strategies. International Journal of Public Administration in the Digital Age, 1, pp. 1-14.
- Cordella A., Bonina C.M. (2012): A public Value Perspective for ICT Enabled Public Sector reforms: a Theoretical Reflection. Government Information Quarterly, 29(4). pp. 512-520.
- Demartini G. et al. (2012). ZenCrowd: Leveraging Probabilistic Reasoning and Crowdsourcing Techniques for Large-scale Entity Linking. Proceedings of the 21st World Wide Web Conference, pp. 469-478.
- de Vries M. (2012): Funding of a System of Key Registers in a PSI-conomics and Contemporary Perspective. The Dutch Experience in a Danish Context, http://gst.dk/media/2915637/funding-of-a-system-of-key-registers-2012.pdf.
- Duggan J. et al. (2015): The BigDAWG Polystore System. SIGMOD Record, 44(2), pp. 11-16.
- Eckert K.P. et al. (2014): Big Data - Ungehobene Schätze oder digitaler Alptraum, https://www.oeffentliche-it.de/documents/10181/14412/Big+Data+ungehobene+Sch%C3%A4tze+oder+digitaler+Albtraum.
- Estermann B. et al. (2009): Integrated and Transcendent e-government: Keys for Analyzing Organizational Structure and Governance. In: Proceedings of the 10th Annual International Conference on Digital Government Research, dg.o, pp. 162-171.
- European Commission (EC) (2010): European Interoperability Framework (EIF) for European public services. Annex 2, Bruxelles, http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf.
- European Commission (EC) (2014): Towards a Thriving Data-Driven Economy (COM(2014) 442 final, http://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1404888011738&uri=CELEX:52014DC0442.
- https://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf.
- European Commission (EC) (2016a): Digital Single Market, https://ec.europa.eu/digital-single-

market/en/digital-single-market.
- European Commission (EC) (2016b): Providing Big Data Opportunities for Public Administrations, http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-22action_en.htm.
- European Union (EU) (2016): Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, L 119.
- Fraefel M. et al. (2011): Inter-organizational Cooperation in Swiss e-Government. In: M. Janssen et al. (ed.): Electronic Government. Proceedings of the 10th International Conference, EGOV 2011, pp. 259-272.
- Frischmann B. M. (2006): An Economic Theory of Infrastructure and Commons Management. Minnesota Law Review, 89, pp. 917-1030.
- Frischmann B. M. (2012): Infrastructure: The Social Value of Shared Resources. Oxford Scholarship Online.
- Graham S.L. et al. (2014): Big Data and Privacy: A Technological Perspective. President's Council of Advisors on Science and Technology, https://www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf.
- Gschwend A. et al. (2015): Publication and Reuse of Linked Data: The Fusepool Publish-Process-Perform Platform for Linked Data. In: Tambouris E. et al. (eds.): Electronic Government and Electronic Participation. Amsterdam: IOS Press, pp. 116-123.
- Hafen E. et al. (2014). Health data cooperatives - citizen empowerment. Methods of Information. Medicine, 53(2), pp.82-86.
- Harrison T. et al. (2012): Open Government and E-Government: Democratic Challenges from a Public Value Perspective. Information Polity, 17(2), pp. 83-97.
- Hayes B. (2012): "Alice and bob in cipherspace," American Scientist, 100(5), pp. 362–367.
- Hitzler P. et al. (2010): Foundations of semantic Web technologies, Boca Raton: CRC Press.
- Höchtl J. et al. (2016): Big data in the policy cycle: Policy decision making in the digital era. Journal of Organizational Computing and Electronic Commerce, 26(1-2), pp. 147-169.
- Horvath, S. (2013): Aktueller Begriff Big Data. Wissenschaftliche Dienste des Bundestages. https://www.bundestag.de/blob/194790/c44371b1c740987a7f6fa74c06f518c8/big_data-data.pdf.
- Janssen M., van den Hoven J. (2015): Big and Open Linked Data (BOLD) in Government: A Challenge to Transparency and Privacy? Government Information Quarterly, 32, pp. 363-368.
- Jarchow T., Estermann B. (2015): Big Data: Chancen, Risiken und Handlungsbedarf des Bundes. Ergebnisse einer Studie im Auftrag des Bundesamts für Kommunikation. Bern: OFCOM.
- Jetzek T. (2016): Managing Complexity Across Multiple Dimensions of Liquid Open Data: The case of the Danish Basic Data Program. Government Information Quarterly, 33(1), pp. 89-104.
- Jetzek T. et al. (2014): Generating Sustainable Value from Open Government Data in a Sharing Society. In: Bergvall-Kåreborn B., Nielsen A.P. (eds.), Creating Value for All Through IT, Berlin, Heidelberg: Springer, pp. 62-82.
- Kim G.H. et al. (2014): Big-data Applications in the Government Sector. Communications of the ACM, 57(3), pp. 78-85.
- Klievink B. et al. (2016): The Collaborative Realization of Public Values and Business Goals: Governance and Infrastructure of Public–Private Information Platforms. Government Information Quarterly, 33(1), 67-97.
- Lindell Y., Pinkas B. (2009): Secure Multiparty Computation for Privacy-preserving Data Mining. Journal of Privacy and Confidentiality, 1(1), pp. 59-98.
- Millard J. (2016): Open Governance Systems: Doing More with More. Government Information Quarterly (forthcoming).
- Mitchell, R. K. et al. (1997): "Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What really Counts." in: Academy of Management Review 22(4): 853 - 888.
- Neirotti P. et al. (2014): Current Trends in Smart City Initiatives: Some Stylised facts, Cities, 38, pp. 25-36.
- Neuroni A. et al. (2016): Exploring the Notion of a National Data Infrastructure and the Governance Issues Surrounding It. Joint Proceedings of Ongoing Research of IFIP EGOV and ePart 2016 (forthcoming).
- Nugroho, R.P. et al. (2015): "A comparison of national open data policies: Lessons learned." Transforming Government: People, Process and Policy 9.3: 286-308.
- OECD (2015): Data-Driven Innovation: Big Data for Growth and Well-Being, Paris: OECD Publishing.
- OECD (2013): Exploring Data-Driven Innovation as a new Source of Growth. Mapping the policy issues raised by "Big Data", DOI: 10.1787/20716826.
- OFCOM (2016): "Digital Switzerland" Strategy. Bern: Federal Office of Communications.
- Ohm P. (2009): Broken Promise of Privacy: The Surprising Failure of Anonymization. UCLA Law Review, 57, pp. 1701-1777.
- Ostrom E. (2003): How Types of Goods and Property Rights Jointly Affect Collective Action. Journal of Theoretical Politics, 15(3), pp. 239-270.

- Poikola A. et al. (2015). MyData – A Nordic Model for Human-Centered Personal Data Management and Processing, http://urn.fi/URN:ISBN:978-952-243-455-5.
- Ross J. et al. (2006): Enterprise Architecture as Strategy: Creating a Foundation for Business Execution. Boston: Harvard Business School Press.
- Shadbolt, N. (2013): Midata: towards a personal information revolution. *Digital Enlightenment Yearbook*, 202-224.
- Shapiro Z. (2015): Big Data, Genetics, and Re-Identification, Bill of Health, http://blogs.harvard.edu/billofhealth/2015/09/ 24/big-data-genetics-and-re-identification/#_ftn1.
- Shin D.-H. (2007): A Critique of Korean National Information Strategy: Case of National Information Infrastructures. Government Information Quarterly, 24, pp. 624–645.
- Su, X. et al. (2016): Privacy as a Service in Digital Health. *arXiv preprint arXiv:1605.00833*.
- Sweeney L. (2000): Simple Demographics often Identify People Uniquely, Carnegie Mellon University. Carnegie Mellon University, Data Privacy Working Paper, 3, pp. 1-34.
- Thapa B. (2016): Big Data in der Verwaltung: Chance mit Fallstricken, http://www.sgvw.ch/2016/01/14/bigdata_thapa/.
- UNECE (2016): Big Data in Official Statistics. Hg. v. United Nations Economic Commission for Europe, http://www1.unece.org/stat/platform/display/bigdata/Big+Data+in+Official+Statistics.
- W3C (2013): SPARQL 1.1 Overview, https://www.w3.org/TR/sparql11-overview.
- Washington A. L. (2014): Government Information Policy in the Era of Big Data. Review of Policy Research 31, pp. 319-325.
- Weber K. et al. (2009): One Size Does Not Fit All – A Contingency Approach to Data Governance. ACM Journal of Data and Information Quality, 1 (1), pp. 1-27.
- Weber R.H. (2016): Transparenz und Open Data. Jusletter IT 25. Mai 2016.
- Williamson O.E. (1999): Strategy Research: Governance and Competence Perspectives. Strategic Manage J, 20:1087- 1108.
- Wylot M. et al. (2014): TripleProv: Efficient Processing of Lineage Queries in a Native RDF store. Proceedings of the 23rd World Wide Web Conference, pp. 455-466.
- Zech H. (2015): Information as Property, Journal of Intellectual Property, Information Technology and Electronic Commerce Law, 6, pp. 192-197.