Big data in the public sector
Uncertainties and readiness

Klievink, Bram; Romijn, Bart Jan; Cunningham, Scott; de Bruijn, Hans

DOI
10.1007/s10796-016-9686-2

Publication date
2016

Document Version
Final published version

Published in
Information Systems Frontiers: a journal of research and innovation

Citation (APA)
Klievink, B., Romijn, B. J., Cunningham, S., & de Bruijn, H. (Accepted/In press). Big data in the public sector: Uncertainties and readiness. *Information Systems Frontiers: a journal of research and innovation*, 1-17. https://doi.org/10.1007/s10796-016-9686-2

Inf Syst Front
DOI 10.1007/s10796-016-9686-2

Big data in the public sector: Uncertainties and readiness

Bram Klievink[1] · **Bart-Jan Romijn**[1] · **Scott Cunningham**[1] · **Hans de Bruijn**[1]

Abstract Big data is being implemented with success in the private sector and science. Yet the public sector seems to be falling behind, despite the potential value of big data for government. Government organizations do recognize the opportunities of big data but seem uncertain about whether they are ready for the introduction of big data, and if they are adequately equipped to use big data. This paper addresses those uncertainties. It presents an assessment framework for evaluating public organizations' big data readiness. Doing so demystifies the concept of big data, as it is expressed in terms of specific and measureable organizational characteristics. The framework was tested by applying it to organizations in the Dutch public sector. The results suggest that organizations may be technically capable of using big data, but they will not significantly gain from these activities if the applications do not fit their organizations and main statutory tasks. The framework proved helpful in pointing out areas where public sector organizations could improve, providing guidance on how government can become more big data ready in the future.

Keywords Big data · Use · E-government · Bold · Readiness · Assessment

✉ Bram Klievink
a.j.klievink@tudelft.nl

Bart-Jan Romijn
j.h.romijn@gmail.com

Scott Cunningham
S.Cunningham@tudelft.nl

Hans de Bruijn
J.A.deBruijn@tudelft.nl

[1] Faculty of Technology, Policy, and Management, Delft University of Technology, Delft, The Netherlands

1 Introduction and background

Technological and social advances have produced a flood of new digital applications and devices, and these are skyrocketing in value to users. Ever more mobile phones, websites, social media, smart household appliances, business software, industrial machines and smart cars are generating ever more digital data. Hence, the amount of digital data available is growing fast (Hota et al. 2015). These enormous volumes of digital data, combined with advances in data analysis, have attracted much interest from industry and research under the label of big data (Lohr 2012). Big data is not a technology itself. It refers to collections of data so large, varied and dynamic that they cannot be handled by conventional data processing technology. Processing big data is complex, due to its great variety, high velocity and extremely large volume (Kankanhalli et al. 2016). Yet, with advanced technologies this type of data can be combined and analyzed, revealing information that was hitherto virtually undiscoverable. Moreover, thanks to developments in computational, storage and analytical technologies, tools for handling and using this data are becoming ever more accessible (Bryant et al. 2008; Hota et al. 2015).

Firms, governments and academia are benefiting from this revolutionary means of knowledge discovery (Mayer-Schönberger and Cukier 2013). The private sector and science, particularly, are starting to use big data in their everyday activities, for example, for business intelligence (Chen et al. 2012). Retailers like Walmart (Bryant et al. 2008), Sears (Henschen 2012) and Amazon (Kelly 2013) use big data to better understand their customers and their buying decisions. Financial institutions, such as Morgan Stanley (Groenfeldt 2012), use big data to predict market behavior and investment performance. Companies like Google, eBay, Twitter and Facebook have created their entire business models around

Published online: 15 August 2016

huge volumes of digital data on individual behavior, information requests and preferences (Davenport and Dyché 2013; Simon 2013). Furthermore, big data is positioning scientific programs, like the Large Hadron Collider in Geneva and the Large Synoptic Survey Telescope under development in Chile, to make astonishing contributions to physics and astronomy (Mayer-Schönberger and Cukier 2013).

While the private sector and science are moving forward with big data use, the public sector seems to be falling behind (Mullich 2013). This is despite the substantial promise that big data holds for government organizations. Big data could help governments improve their efficiency, effectiveness and transparency (Milakovich 2012), which have long been among the main goals of public sector ICT use (Cordella and Bonina 2012; Heeks 1999; Weerakkody et al. 2011). For example, big data could enable better decision support information, more informed policymaking (Janssen and Kuk 2016), faster and richer images of evolving reality, and improved services based on better insight into citizen demands and needs (Chen and Hsieh 2014). Such benefits could be a potent tool for solving lingering social problems, such as transport congestion, healthcare provision and sustainable energy production (Scholl and Scholl 2014). While big data use could help transform government operations (Bertot and Choi 2013; Joseph and Johnson 2013; Yiu 2012), it could also potentially undermine public goals (Clarke 2016; Janssen and Van den Hoven 2015). Indeed, use of big data raises new challenges and poses new threats for government (Margetts and Sutcliffe 2013).

Perhaps in deference to these challenges and threats, actual use of big data in the public sector is still very limited in many countries, including the Netherlands (Van den Toorn 2014). Government organizations seem to still be in an orientation or contemplation phase regarding big data (Ministerie van Infrastructuur en Milieu 2014; TechAmerica Foundation 2012). We suspect that government organizations are postponing decisions on big data use because they are unsure whether they are ready for the introduction of big data and if their organizations are equipped to take advantage of the opportunities it offers (Malik 2013). In other words, public organizations are uncertain of whether and how to implement big data, and they lack the tools to determine if they are ready for big data use. Uncertainty about organizational readiness and inability to make an accurate judgment in this respect is problematic for public sector organizations. Not only does it slow the development of potentially valuable big data uses, it also increases the risk of premature big data implementations, which could undermine the success of future big data ventures in the public sector.

This paper examines public sector organizations' readiness for big data, constructing an assessment framework for establishing big data readiness. Section 2 describes the research approach. Big data is a fuzzy concept. We therefore conceptualize it in terms of how big data is *used* within organizations. This conceptualization is presented in section 3. Section 4 introduces our big data readiness assessment framework based on three components: organizational alignment, organizational maturity and organizational capabilities for big data use. The framework provides a methodology for public organizations to reflect on the uncertainties that may hinder their decision-making on big data projects. The idea is that big data use could work for the public sector if it is aligned with public organizations' goals and their ways of working. Section 5 describes our test application of the framework in the Dutch government. Section 6 summarizes our study findings, study limitations and suggestions for future research. Section 7 wraps up with conclusions.

2 Research approach

The concept of big data is used often and in various ways. Given the fuzziness of the concept (Cunningham and Thissen 2014), we first set out to define it to ensure clear communication with our study interlocutors. Instead of trying to gauge what big data itself precisely is, we based our conceptualization on how big data is *used* within organizations.

We first conducted a series of explorative interviews with officials in 11 organizations in the Dutch public sector. These interviews were held in 2014 and served two purposes. One was to verify the usefulness of our proposed way of looking at big data. The second was to learn what uncertainties they experienced that hindered decisions on big data use within their public sector organizations. A brief set of talking points guided the interviews. Structured analysis of the interview reports pointed out three main areas of uncertainty:

1. Uncertainty about what kinds of big data uses were suitable for the organization;
2. Uncertainty about whether the organization was sufficiently mature for big data use;
3. Uncertainty about the capabilities of and within the organization for big data use.

All three of these areas of uncertainty were taken to represent organizational readiness. We used these categories to direct our next step: construction of a framework for accurately assessing the big data readiness of organizations in the public sector. From the literature, we selected three established theoretical models, each corresponding to one of the three uncertainties. These models formed starting points for development of our framework. We selected literature on *organizational alignment* (specifically, the strategic alignment model of Henderson and Venkatraman 1993) to reflect the first uncertainty. For the second, we turned to the literature on *organizational maturity;* and for the third we used the literature on *organizational capabilities*.

We tested our framework by putting it into practice: assessing the readiness for big data use in the same 11 Dutch public sector agencies that were involved in the unstructured interviews. The Netherlands was selected as it has a stable public sector and is a leading country in global e-government indexes, such as the UN e-government survey 2014 (UNPAN 2014). These attributes suggest that the Dutch public sector is at the forefront of development and interest in big data technologies. Furthermore, within the public sector, organizations were selected that were considered likely to use large volumes of substantive data in their main activities. Such large data volumes are a main prerequisite for big data applications. The organizations and agencies chosen could therefore be assumed to be among the leading parties within the public sector when it came to big data use.

To assess the big data readiness of these organizations, we examined the three focal components of the framework (organizational alignment, organizational maturity and organizational capabilities) first separately and then as a single metric to develop an indicator of readiness per organization. Then, combining the readiness assessments for these 11 organizations, we produced a readiness assessment for the Dutch public sector. For details, see Romijn (2014).

The information required to implement the assessment framework was gathered via a questionnaire administered to experts within the 11 participating organizations. The questionnaire consisted of 41 items, with various subitems. Some questions were open-ended, while others were multiple choice or required answers on a 7-point scale. Organizational alignment, organizational maturity and organizational capabilities were assessed using the questionnaire results and an assessment scorecard. The appendix presents the scorecard rules (A.1) and an example of use of the scorecard to assess one of the organizations (A.2).

In brief, to assess *organizational alignment* we categorized the public organizations based on their main statutory tasks and current data activities, and considered the big data application types conforming with these (see section 3.3). Based on organizations' current IT activities and the requirements of each type of big data application and big data characteristics (see section 3.1), a degree of alignment between the organizations and potential big data application(s) could be established. For *organizational maturity*, we investigated the organizations' current activities and information sharing, their IT facilities for that purpose and the data systems currently in use (presented in section 4.2). Finally, we assessed the *organizational capabilities* of and within organizations in seven areas, selected for their relevance to big data use: IT governance, IT resources, internal attitude towards big data, external attitude towards big data, legal compliance, data governance and data science expertise (see section 4.3). These capabilities were assessed in three dimensions: importance for big data success in the organization, potential to develop a capability within the organization and current presence of the capability within the organization. Following Valdés et al. (2011), we derived our assessment on each capability by comparing given scores to a maximum score (see the appendix for details). This resulted in an assessment per capability for each organization, which yielded an overall organizational capability assessment.

3 Conceptualizing big data use

Big data is often defined as data so large, varied and dynamic that conventional hardware and software cannot process it (Laney 2001). However, due to fast advances in technology, exactly what can be considered big data is always changing, making it hard to express in specific and measurable terms. Furthermore, what hardware and software is conventional differs from industry to industry and from organization to organization. So, generalizing what constitutes big data is very hard to do. Fortunately, arriving at a set definition of big data is not necessary for us to understand big data. Indeed, in this paper we suggest a different approach. By describing how big data is *used* in organizations instead of big data itself, we sidestep the complexity introduced by its dynamic and specific nature. Instead, we focus on the way these dynamics are handled within organizations, exploring big data use in terms of its characteristics, processes and applications.

3.1 Big data use characteristics

Our first step in describing how big data is used was to clarify the difference between the use of conventional digital data and the use of big data. A review of the literature (Adrian 2011; Chen et al. 2014; Davenport et al. 2012; Gantz and Reinsel 2011; Hota et al. 2015; Janssen and Kuk 2016; Mayer-Schönberger and Cukier 2013; OpenTracker 2013; Simon 2013) suggested the following five differentiating characteristics of big data:

1. Use and combining of multiple, large datasets, from various sources, both *external and internal* to the organization;
2. Use and combining of *structured* (traditional) and *less structured or unstructured* (nontraditional) data in analysis activities;
3. Use of incoming data streams in *real time* or near real time;
4. Development and application of *advanced analytics and algorithms*, distributed computing and/or advanced technology to handle very large and complex computing tasks;
5. *Innovative use* of existing datasets and/or data sources for new and radically different applications than the data were gathered for or spring from.

Table 1 Big data activities

Main activity	Big data activities	Example literature sources
Collection		
	Collect, annotate	Bryant et al. 2008; Chen et al. 2014; Cumbley and Church 2013; Miller and Mork 2013
	Acquire, record	Agrawal et al. 2011; Chen et al. 2014; Dijcks 2012; TechAmerica Foundation 2012
	Generate	Chen et al. 2014; Gustafson and Fink 2013
	Choose, select	Tekiner and Keane 2013
	Sense	Bryant et al. 2008
Combination		
	Extract, clean, prepare, process	Agrawal et al. 2011; Chen et al. 2014; Miller and Mork 2013; TechAmerica Foundation 2012
	Combine	Cumbley and Church 2013
	Organize	Dijcks 2012; Miller and Mork 2013
	Store	Bryant et al. 2008; Chen et al. 2014; Gustafson and Fink 2013; TechAmerica Foundation 2012
	Integrate, represent	Agrawal et al. 2011; Miller and Mork 2013
Analytics		
	Analyze, model	Agrawal et al. 2011; Brohman et al. 2000; Bryant et al. 2008; Chen et al. 2014; Cumbley and Church 2013; Dijcks 2012; Gustafson and Fink 2013; Miller and Mork 2013; TechAmerica Foundation 2012; Tekiner and Keane 2013
	Visualize	Miller and Mork 2013
	Interpretation	Agrawal et al. 2011; Tekiner and Keane 2013
Use		
	Initiate	Brohman et al. 2000; Brohman 2006
	Deploy	Brohman et al. 2000; Brohman 2006
	Make decisions	Dijcks 2012; Miller and Mork 2013
	Apply, produce insight	Cumbley and Church 2013; Gustafson and Fink 2013; TechAmerica Foundation 2012
	Evaluation	Brohman et al. 2000; Brohman 2006

For our study, we qualified data use as big data use when it met three or more of these characteristics.

3.2 Using big data: A process

Our second step in describing big data use was to differentiate activities in the process of using big data. These are steps in the data value chain that take it from raw data in the environment to actionable knowledge for decision-makers. Describing these steps enabled us to relate the unique consequences and aspects of big data use to specific data-related activities. Many of the activities in the big data use process are likely to already be present within most organizations, though they are probably not geared towards big data use.

Based on contributions of a number of scholars, technology consultants and vendors on data value chains and knowledge creation processes from data, four big data use activities were formulated (Bryant et al. 2008; Chen et al. 2014; Cumbley and Church 2013; Miller and Mork 2013; TechAmerica Foundation 2012). These are collection, combination, analytics and use, with each consisting of various sub-activities. Table 1 presents these, along with relevant literature sources.

We argue that these activities, taken together, make up a value chain, or a *big data process* of sorts (Fig. 1). Note that communication and dissemination of use and results might be considered either an integrated part of the chain or a separate step. Through this process, data flows from collection, to combination and analytics, to use. The process is cyclical, reflecting the continuous nature of big data. A feedback loop extends from the use of the data back to the collection of new raw data.

3.3 Big data applications

Our third step in conceptualizing how big data is used in organizations was to formulate a typology of potential big data applications. One of big data's most attractive features is its potential for use in virtually any situation in which data is

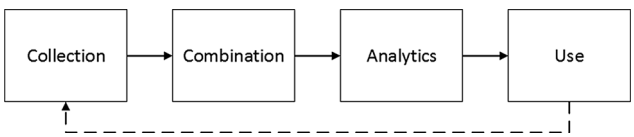

Fig. 1 Big data use process

Table 2 Big data application types

Application type	Object evaluation	Research	Continuous monitoring
Prominent big data use characteristics	- Internal & external datasets - Innovative use of existing data - Advanced analytics & algorithms	- Internal & external datasets - Structured & unstructured data - Advanced Analytics & Algorithms	- Real-time or near real-time - Advanced analytics & algorithms - Innovative use of existing data
Initial big data activity	Combination	Analytics	Collection
Added value	Decision support information	New insights	Enriched view of reality
Examples of relevant articles	Chen and Hsieh 2014; Ferro et al. 2013; Margetts and Sutcliffe 2013	Agarwal and Dhar 2014; Janssen and Kuk 2016	Bharosa et al. 2013; Chen and Hsieh 2014; Klievink and Zomer 2015

available. The possibilities for using big data are therefore manifold. However, some handles on this huge landscape of big data applications will enhance both insight and communication on the subject.

A number of authors have created typologies of big data applications. However, none of these are particularly suited for our objective. For example, Chen et al. (2014) produced a typology based on the various data formats used for big data applications. Yet, as the greatest value of big data comes from combining different formats, this typology is not considered suited for assessing uses in public sector organizations. Other typologies (e.g., Milakovich 2012; Shiri 2014) are rather specific to an industry or sector, or lack sufficient detail to provide insights to practitioners facing decisions on big data use.

Based on the literature, we derived a categorization from the added value pursued with the big data application. Three application types were identified: object evaluation, research and continuous monitoring. Table 2 presents these, alongside the big data use characteristics associated with each, the initial activity in the big data use process and its primary added value. By using different information inputs and different analyses methods, the application types create different kinds of knowledge, for different purposes, adding value to organizations in different ways. In short, the three types of big data applications differ in the types of knowledge they deliver to users.

Our first application type, object evaluation (also called subject evaluation), involves use of big data to evaluate, rank or classify large sets of objects (or subjects). The evaluation is based on many different attributes of each of the objects for which data is available. By analyzing the values of each attribute for all the objects, the objects can be individually evaluated or ranked or classified in comparison with each other according to preset decision rules. The decision rules can be based on anything, from legislation or policy to the outcomes of predictive models or even expert opinions. Object evaluation applications identify very specific cases, or specific combinations of attributes, literally out of millions of objects. Big data comes into play when the set of records and attributes becomes extremely large, typically by combining data from multiple datasets, therefore requiring advanced algorithms and a large set of decision rules.

Our second application type is research, in other words, seeking new information and new insights. By combining huge amounts of data from various data sources and with very different formats and structures, and by using many different and advanced analytical methods, new relations and connections can be found, patterns identified and never before observed behavior recognized. These newly discovered connections and patterns provide clues for subsequent investigation. Thus, previously undetectable correlations may yield new insights pointing to research areas yet to be explored, or they may offer or represent powerful policy challenges or options.

Our final application type is continuous monitoring. Here, data is collected through extensive (sensor) networks and made available for analysis in real time or near real time. The resulting information is provided, for example, via 'dashboards' in easy-to-read visualizations or through exception reporting. These give human decision-makers information to which they can react, or which can be followed up by automated responses.

4 Assessing public sector readiness for big data use: Towards a framework

Reflecting our focus on organizational characteristics to evaluate big data readiness in the public sector, Fig. 2 presents our assessment framework's component parts: organizational alignment, organizational maturity and organizational capabilities. For each component, we drew on established models of the interaction between organizations and technology. However, none of these models were originally designed with big data in mind, so we had to adapt each. The overall assessment of an organization's big data readiness was understood to be

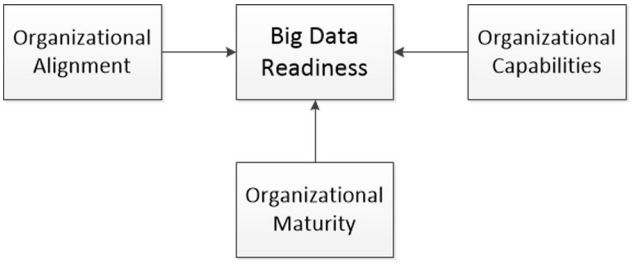

Fig. 2 Main components of the big data readiness assessment framework

the sum of its scores on the three components of the framework. This was expressed in a *big data readiness score* and associated description. This section examines in greater detail each of the components. As noted in section 2, our assessments of these were guided by a scorecard and scoring rules, as presented in the appendix.

4.1 Organizational alignment

The first uncertainty addressed was if big data use was suited for the organization in question and, if so, in what form. We termed this uncertainty organizational alignment. It concerns whether big data use could be reconciled with the organization's current structure, its main activities and its strategy.

Malik (2013) sensibly related big data readiness assessment to IT-business alignment assessment. The idea is that big data projects must be aligned with the organization they are to be executed within (Kiron 2013). We borrowed the IT-business alignment concept to assess big data readiness in the public sector. For this, we turned to a model that has often served as the basis of alignment assessments and variations of them: the strategic alignment model, as formulated by Henderson and Venkatraman (1993). According to this model, business strategy, IT strategy, organizational infrastructure and IT infrastructure need to be aligned, as they are interconnected within an organization.

To gear the strategic alignment model more to big data use and the public sector, we adapted the model's four components to big data and to public sector specific and relevant aspects. The *business strategy* component was taken to refer to organizational strategy, expressed in a typology of the main statutory tasks of public organizations. Indeed, the statutory tasks of public organizations are typically set by laws or regulations, and the organizations have a legal obligation to perform them. Public organizations are funded for these tasks, directly or indirectly. Furthermore, many public organizations are limited to *only* performing their statutory tasks and activities in direct support of these. They are prohibited from doing more, even if big data were to open up a potentially expanded portfolio. The statutory tasks thus largely determine a public organization's main activities and its data activities in support of these.

We translated *organizational infrastructure* as the intensity with which strategic (big) data use activities were being performed or could be executed by an organization. We took two of the big data process activities described earlier – collection and use – as representing strategic activities within organizations. All organizational data use processes must be aimed towards the choices made on these two data activities, regarding how and especially how intensive these activities are to be structurally performed.

We considered *the IT strategy* factor as indicative of the type of big data applications that the organization under assessment was most interested in, in terms of the three big data application types (i.e., object evaluation, research and continuous monitoring). The type of big data application chosen determines in a large part the IT strategy used by the organization to support its big data activities.

Lastly, *IT infrastructure* was taken to express the big data use characteristics (from section 3) needed for the chosen big data application types. By determining the extent that the five big data use characteristics were already among an organization's data activities, the alignment between the current IT infrastructure and the IT infrastructure required for the chosen big data application type could be determined. This was considered an indicator of alignment between the current situation in the organization and a prospective future situation that includes big data use.

Table 3 presents four organization types in which the main statutory task, data collection and data use intensity are ideally aligned with big data use characteristics and the chosen big data application type. This theoretical 'optimum' provided a basis for comparison when assessing the data readiness of real organizations in the Dutch public sector.

We began our big data readiness assessment of the Dutch public sector by evaluating whether the main statutory tasks (organizational strategy) and data activities (organizational infrastructure) were in alignment with the chosen big data application type (IT strategy) and supportive of big data use characteristics (IT infrastructure). We assessed the organizational aspects based on answers to our questionnaire (see the appendix for scoring criteria). For the IT aspects, we assessed four dimensions for each of the big data application types: interest in the application type, need for its specific benefits, its applicability in the organization and its feasibility.

4.2 Organizational maturity

The second uncertainty addressed was organizational maturity, operationalized as the maturity of e-government initiatives within the organization. The so-called 'e-

Table 3 Organizational alignment: The theoretical optimum

Organization type	1	2	3	4
Main statutory task	Coordination, project-based task, no data used	Research, evaluation	Registration, documentation	Administration, management
Data collection activity intensity	Low	Low	High	High
Data use activity intensity	Low	High	Low	High
Most present big data use characteristics	–	- Internal & external datasets - Structured & unstructured data - Advanced analytics & algorithms	- Internal & external datasets - Innovative use of existing data - Advanced analytics & algorithms	- Real time or near real time - Advanced analytics & algorithms - Innovative use of existing data
Best aligned big data application type	–	Research	Object evaluation	Continuous monitoring

government growth stage models' are among the established frameworks for assessing public sector organization maturity in e-government. These describe IT-enabled government reforms as following a typical growth path, which is presented in growth stage or maturity models with multiple distinct steps (Andersen and Henriksen 2006; Klievink and Janssen 2009; Layne and Lee 2001). Although a full assessment of e-government maturity using these frameworks falls outside the scope of this paper, the five stages described by Klievink and Janssen (2009) provided the basis for our assessment of the stage of e-government development in the public organizations currently under study. This model was chosen because it recognizes the inter-organizational aspects of e-government development, which are also very significant in big data development in the public sector.

Organizational maturity thus indicates how far organizations have developed towards a state in which they collaborate better with other public organizations (and their IT) and provide more citizen-oriented services and demand-driven policies. Big data can be an enabler of these developments, and vice versa, these developments can help make big data use more effective. More cooperation and more attention to citizen demands makes more data available for big data applications, which can be designed to help public organizations better provide and adapt their activities, services and decision-making to citizen demands. Hence, the maturity of the organization is not just related to big data implementation; it is also an indicator of how well an organization is able to use big data to its full advantage.

Table 4 presents e-government growth stages based on the Klievink and Janssen (2009) model. The big data readiness framework expresses these five growth stages as scores on three aspects: activities and information sharing, IT facilities and development path of data systems. The first two were taken from the original description of the five growth stages. They were chosen as clearly distinguishing between the different growth stages, while also reflecting the language, the expertise and the perspectives of the practitioners interviewed in the first step of this research project.

The third aspect – development path of data systems – did not come from the initial article. We added it here to adapt the growth stage model for our assessment of the readiness of public sector organizations for big data use. Hence, the stages in the 'data systems' column of the table can be seen as a model of data systems development, over time and towards increasing technical complexity. The explanation column also concerns these data systems.

To assess where an organization is on this maturity model, we used questionnaire items asking the government agency representatives which of the characteristics from Table 4 were present in their organization, and how strong that presence was. This provided insight into the organizations' maturity with respect to e-government initiatives and, in particular, big data enabled transformational government. See appendix for the assessment rules.

4.3 Organizational capabilities

The third uncertainty addressed was organizational capabilities, in particular, whether organizations possessed the requisite capacities to use big data, to create value from it for the organization and to ensure that no negative consequences arose from big data use.

Klievink and Janssen (2009) noted that advancement towards a further stage of e-government development depends on acquirement of the (dynamic) capabilities needed for that stage. Furthermore, progression to higher stages need not be sequential. Thus, merely assessing at which

Table 4 Organizational maturity: e-Government growth stages and characteristics

	e-Government growth stage	Activities and information sharing	IT facilities	Data systems	Explanation
1	Stove-pipe organizations	Information and activities are bounded within separate departments within the organization	All activities are, where possible, supported by IT	*Digitalization of processes*	Providing digital tools for operational processes in data entry and data search
2	Integrated organizations	Activities transcend the separate departments and information is shared throughout the organization	Organization-wide IT infrastructure from which information is accessible throughout the organization	*Business intelligence & data management*	Bundling of information storage across the organization and automated analysis tools that create information from organization's internal data
3	Nationwide portal	Activities go beyond the boundaries of the organization and information is accessible from outside the organization	IT infrastructure suited for external access by applications and for information access within the organization	*Business intelligence 2.0 & data management 2.0*	Business intelligence & data management with options to access and change data and information from outside the organization
4	Inter-organizational integration	Activities and information are extensively shared with other organizations	IT infrastructure suited for full access by applications and for information access by other organizations	*Open data*	Provision of as much anonymized data as possible from the organization as public data, in a standardized format
5	Demand driven, joined-up government	Activities and information are organized centrally and made available to all relevant organizations and stakeholders	Centrally built IT facilities supporting all information and applications and fully accessible to all involved stakeholders	*Big data*	Collection, combination and analysis of large, complex datasets with unconventional technologies to create new knowledge for the organization

stage a public organization currently is cannot provide an adequate view of readiness. This observation, combined with the notion that governments need to broaden and deepen their expertise, professionalism and capabilities (Janssen et al. 2015), suggests we seek an overview of the organizational capabilities needed for big data use, alongside a method for assessing them.

We derived the organizational capabilities considered vital for big data use in the public sector from the literature of relevant research fields. This included the literature on IT adoption (e.g. Ebrahim and Irani 2005; Jeyaraj et al. 2006; Kamal 2006), IT implementation (e.g. Finney and Corbett 2007; Premkumar 2003; Wixom and Watson 2001), innovation adoption (e.g. Robey et al. 2008), dynamic and core capabilities in IT (e.g. Daniel and Wilson 2003; Eisenhardt and Martin 2000; Klievink and Janssen 2009) and on big data specifically (e.g. Chen et al. 2014; McAfee and Brynjolfsson 2012; Milakovich 2012; Ross et al. 2013; Tambe 2014). We first identified all conditions considered important to IT and information systems implementation. The articles examined addressed different topics from various theoretical backgrounds, therefore providing a broad and diverse picture of conditions needed for IT implementation. Articles were selected based on the relevance of the core topic (IT implementation/adoption by organizations), number of citations and the unique contribution of mentioned conditions as applicable to big data use. Through content analysis of the articles and after selecting capabilities based on frequency and relevance, similar capabilities were combined, and reformulated in light of big data use. Based on this, we abstracted a list of seven overarching organizational capabilities for the use of big data: IT governance, IT resources, internal attitude, external attitude, legal compliance, data governance and data science expertise (Table 5).

Three dimensions of these organizational capabilities were assessed: importance for big data success in the organization, possibility to develop capability in the organization and current presence of the capability within the organization. Based on the maturity assessment methods used by Valdés et al. (2011), the level of each capability was derived by comparing given scores to the maximum score. This yielded a capability level per item for each organization, which could be used to calculate the overall capability levels per organization and organization type. See the appendix for the decision rules for the organizational capabilities assessment.

Before applying this big data readiness framework to organizations in the Dutch public sector, a remark on its theoretical foundation is in order. The components of the framework are here formulated without a proven connection to successful big data use. The capabilities from the literature have, in most

Table 5 Organizational capabilities for big data use

Capability	Explanation
IT governance	Capability to design and develop IT strategy, decision-making and responsibility structures, supporting the organization, including integration of new IT systems
IT resources	Capability to design, develop and maintain suitable IT infrastructure and expertise to facilitate current and new IT systems
Internal attitude	Capability to develop internal commitment and vision for new processes and systems, especially openness towards data-driven decision-making
External attitude	Capability to develop external commitment and support for new processes and systems with important stakeholders
Legal compliance	Capability to design and develop a compliance strategy including process design, monitoring and redesign of processes, especially regarding privacy protection, security and data ownership regulations
Data governance	Capability to design and develop a data strategy including collection, acquisition, quality control and data partnerships
Data science expertise	Capability to bundle/acquire, develop and retain data science knowledge in the organization, especially bundling knowledge on IT, business, statistics and mathematics

cases, been related to the (successful) implementation and use of other IT systems and innovations, but not to big data, as empirical evidence of successful big data use is sparse. The assessment framework should therefore be seen as a promising tool, but one in need of further specification, refinement and validation with more empirical evidence of successful big data use implementation in public organizations.

5 Findings: Assessing big data readiness in the Dutch public sector

5.1 Background on our test application

As this framework is largely theoretical, we put it to the test by assessing the big data readiness of 11 organizations in the Dutch public sector. Our goals in doing so were three. First, we sought to help decision-makers overcome their uncertainties regarding plans for or explorations of big data uses. Second, we wanted to know why the public sector lags behind other sectors and help prepare the ground for future big data applications. The overview of the readiness of Dutch public sector organizations for big data use and areas for improvement in this regard could help us to answer this question. Our third aim was to demonstrate how the framework could be used to provide that overview for practitioners and academics.

We started our assessment by translating the framework into a structured questionnaire, to ensure that respondents had similar conceptualizations of the terms and notions used. Furthermore, to understand how the respondents saw big data, the questionnaire explicitly asked what big data meant to them and what their organizations' perspectives on big data were. We also asked what interest, plans and expected value they and their organizations had for big data. Open-ended questions sought respondents' views on these aspects. Other questions required scores to be given on a 7-point scale. For example, for each type of big data application, we asked, "To what extent is this type of big data application relevant to the main tasks/activities of your organization?" There were also multiple-choice questions such as, "What type of organizational tasks best reflect those of your organization?" Because the questionnaire consisted of 41 items, 9 of which had up to 7 of subitems, it typically took more than an hour to complete. As sending such an extensive questionnaire to a large group of respondents was expected to result in a very low response rate, we directly asked the initial interviewees to fill in the questionnaire. In addition to being experts on the data activities of their organizations, these respondents had already been introduced to the project and were familiar with our conceptualization of big data in terms of its use within organizations. All of these respondents completed the questionnaire.

The big data readiness of each organization was first determined for each of the three components of the framework and then combined to establish an overall big data readiness score. Table 6 presents these results. See the appendix (A.2) for an example of how we arrived at these outcomes. The assessments suggest that the planned big data applications were reasonably well aligned with the organizations, receiving an average assessment of 'medium'. Thus, organizations in the Dutch public sector appeared to have a fair understanding of big data applications and their implications for their organizations. The average assessment on organizational maturity, however, was 'low'. This indicates that few of the organizations were structurally collaborating with others on activities and sharing information. Further development will be needed in this regard before they are equipped to take full advantage of big data's potential. Organizational capabilities were generally assessed as 'medium'. While the organizations had developed many of the capabilities required for big data use, they did not yet seem to be at a level that would allow them to make effective use of it. Combining these three assessments established the overall big data readiness of organizations in the Dutch public sector. This was found to be 'medium', suggesting that these organizations were developing the requirements for using big data, but they were not ready for it yet. Further development will be needed before big data can (and should) be introduced, to ensure that big data applications are

Table 6 Big data readiness of 11 organizations in the Dutch public sector

#	Organizational alignment	Organizational maturity	Organizational capabilities	Overall big data readiness
1	Very high	Low	Medium	Medium
2	Medium	Medium	Medium	Medium
3	Low	Very low	Low	Very low
4	High	Low	High	Medium
5	Very low	Very high	Low	Medium
6	Medium	Low	Medium	Medium
7	High	Low	High	Medium
8	Medium	Very low	High	Medium
9	Very low	Very low	Low	Very low
10	Medium	Very low	High	Medium
11	Medium	Low	Medium	Medium
Overall	Medium	Low	Medium	Medium

approached in a way that adds value to organizations in the Dutch public sector. The remainder of this section discusses our main findings on each of the three components of the assessment framework.

5.2 Organizational alignment

Organization types in which data is used intensively (i.e., research/evaluation and administration/management) received much higher assessments on organizational alignment than the two types of organizations that did not use data very intensively (i.e., coordination/project-based and registration/documentation). A possible explanation for this large distinction is differences between the organizations' knowledge of and experience with data-intensive activities. Organizations less experienced with data-intensive activities might find it harder to judge the consequences and requirements of the different types of big data applications for their organization. The assessments of organizational alignment over the different statutory tasks of the organizations (not presented in the table) seemed to confirm this. Data-intensive organizations thus had a head start when considering big data use. They also scored much higher on big data readiness considering all categories of our assessment framework. In particular, coordination and project-based types of organizations scored low overall, indicating initially a need for orientation on potential big data uses, before moving forward with decisions and plans for big data use.

Unlike the intensity of data use, the intensity of data collection within the organizations did not appear to be positively associated with the alignment of the planned big data applications with the current situation of the organization. It therefore appears that experience with intensive *use* of data is a more important determinant of big data readiness than experience with intensive data *collection*.

5.3 Organizational maturity

We found no clear link between the organization types and assessment results on organizational maturity. Thus, there is no parallel on this component to the finding that organizations using data intensively scored higher on big data readiness. The 11 organizations under study were nowhere close to the fifth and final, transformational stage at which big data use becomes a key enabler. Most of the organizations had a maturity level two. This indicates that the organizations had mostly integrated activities, information sharing, IT facilities and data systems across departments within the organization, but were not structurally collaborating with other organizations on these aspects. This may be only logical, as the organizations were in different fields and had rather specialized statutory tasks, often without comparable organizations with similar goals to collaborate and integrate with. Introduction of big data use may result in wildly differing data from other fields and organizations becoming valuable to such organizations. Development of information and especially data sharing and cooperation could add significant value and quality to prospective big data initiatives.

In many organizations, substantial time and effort would have to be invested to approach an e-government growth stage that enables them to take advantage of the full potential of big data use. Organizations in the Dutch public sector did not seem to be currently performing activities that support extensive sharing and exchanging of data and information between organizations. This would make it harder for them to use big data successfully, as the available data is limited in amount and especially variety, which constrains the knowledge they can create from it.

5.4 Organizational capabilities

Most of the organizations examined appeared able to handle big data. Only a few were assessed as 'low' on organizational capabilities, with the remainder receiving 'medium' or 'high' average scores across all capabilities. That is, most received scores between 70 % and 81 % in our big data readiness framework. Just as with organizational alignment, organizations with intensive data use scored higher on organizational capabilities for big data use. They seemed to have acquired better developed capabilities for big data use and its introduction through their current intensive data use activities.

5.5 Overall big data readiness

Although some organizations scored quite high, the overall picture is that even those had substantial work to do in

orientation, planning and development, before big data could be successfully introduced within their organizations. None of the assessed organizations received consistently high assessment scores across all three of the big data readiness components. None, moreover, was close to the level of big data readiness required to use big data successfully.

From our overall big data readiness assessment another important observation can be made. That is, the organizational capabilities of Dutch government organizations were quite well developed, on average, which could lead organizations to believe they are ready to start using big data. At some level this may be true, as they do have the capabilities to implement and use big data and the associated technologies. However, the scope of a radical new concept like big data is much wider than just the organizational capabilities required for it. Big data applications are so comprehensive and potentially invasive for the organizations using them that organization alignment is as important as the technological nuts and bolts required for successful use.

A key finding is that organizations may be technically capable of using big data, but they will not significantly gain from these activities if the applications do not fit their organizations and main statutory tasks. Organizational alignment of big data applications is vital for their success. Without it, big data applications cannot deliver added value in line with an organization's main activities. Organizational alignment therefore should not be neglected, even if organizational capabilities are sufficiently available or developed.

6 Discussion, limitations and future work

6.1 Reflection on the framework and test application

Big data offers a great opportunity for the public sector to structurally improve and transform government organizations. It should not be wasted. The decision support information, new insights and richer images of reality offered by the three big data application types provide opportunities for public organizations to significantly improve their effectiveness, efficiency and transparency and to advance on the e-government maturity ladder. However, as our findings on big data readiness showed, much work remains to be done to unlock the full potential. Public organizations should not start using big data before they are ready, as that could unleash big data's darker side. Breaches of privacy and security of personal data, unfair treatment of citizens through overly extensive and unethical datafication of decision-making processes, wrong or suboptimal decisions because of incorrect data handling, analyses and interpretation, inadequate and faulty IT facilities and large IT investments that never pay off are just a few of the dangers (Clarke 2016; Janssen and Van den Hoven 2015; Margetts and Sutcliffe 2013). These threats are real and pose a substantial risk to the potential value of big data for society. They can be avoided only with careful planning and with development of adequate organizational alignment, maturity and capabilities for big data use. Only when organizations are able to fulfill all of these prerequisites should they consider embarking on big data use.

Against this background, the big data readiness framework served two purposes. First, it provided a valuable analysis tool for gaining more structured and detailed insight into public sector organizations' current readiness for big data use and possible areas of improvement in this regard. Second, by conceptualizing big data in terms of its use in organizations, and in terms of the types of organizations using it, the framework demystified the concept of big data, making it more understandable, communicable and practical for both practitioners and academics. Distinguishing the three main components in the readiness framework, and considering big data's characteristics, the big data use process and big data application types, allowed us to link the concept to established academic theory. Particularly useful were theories on e-government maturity, organizational change and IT implementation, and the dynamic capabilities of organizations. These links may serve as a pathway for academics to further refine, describe and explain the concept of big data and its implications and extend knowledge on it.

6.2 Implications and suggestions for practice

Given big data's high potential, organizations in the Dutch public sector are likely to continue developing activities towards big data use. In doing so, our research results suggest that they should learn and focus on what big data use will entail for their organization and what specific added value big data could bring. Public organizations should establish specific plans or roadmaps to guide their development. By ensuring that these plans hone in on the weak points identified by the big data readiness framework, organizations in the Dutch public sector can swiftly develop their big data alignment, organizational maturity and organizational capabilities. This could set the stage for future introductions of big data use, with added value to their organizations, thus advancing their development on the e-government maturity ladder and delivering value for society.

The current study identified four areas for improvement. First, organizational alignment could be improved by focusing only on big data applications with a good fit within the organization. Public organizations with research as a main statutory task could work towards improved alignment by strengthening data activities that support a combination of structured and unstructured data. Organizations with registration tasks could strengthen alignment by seeking ways to make more innovative use of their data.

Second, organizational maturity could be improved by developing more structural collaboration between organizations. Sharing more information, activities and data will enable organizations to provide big data applications with increasingly diverse input, leading to more accurate and informative insights. Whereas in the past cooperation only with other organizations in the same field made sense, as relevant data and knowledge could usually only be found there, today's big data technologies can add value by combining data from very different sources.

Third, our big data readiness assessments suggest that organizational capabilities could be improved most of all by focusing on data science expertise, data governance and IT governance. Thus, public sector organizations should pay more attention to the recruitment and training of so-called 'big data talent' or data scientists. By employing more experts in the various professions concerning data, in-house expertise on data-intensive activities can be improved, to the benefit of current data activities as well as the big data readiness of the organization and successful utilization of big data applications in the future.

Finally, public sector organizations should consider starting small, with single-function big data applications suited to their current organizations and well known and understandable to them. Once these applications are successful, and only then, they might start to scale up towards larger and more diverse applications. This process would allow organizations to learn from their use of big data, while fostering positive attitudes among stakeholders and minimizing associated risks.

6.3 Limitations of the study

Given the novelty of the topic, our study had to draw on several fields of research, and its resulting assessment of the Dutch public sector organizations cannot be viewed as definitive. Moreover, several drawbacks of our approach must be noted. First, both our identification of the uncertainties and our final assessment of the public sector organizations were based on a limited number of interviews (11) and questionnaires (also 11). Although the prevalence of the same uncertainties across respondents and corroboration in the literature suggest that these represent key uncertainties in big data use decisions, we do not know the extent that this holds for all public sector organizations in the Netherlands, let alone those in other countries. Despite the structured approach of the assessment, our basing the framework on the literature and our selection of experts, we still cannot generalize our findings from the questionnaire. Nonetheless, we believe that these limitations do not undermine the value of this study, as its objective was to paint a rich picture of the uncertainties that public sector organizations face in making decisions on big data use and how readiness in this field could be assessed.

Second, limitations regarding the assessment framework remain and should be taken into account when applying it in other situations and contexts. First, the assessment method is primarily suited to public organizations with clear and intensive data activities. Public organizations working much less intensely with data in their day-to-day operations may not be able to express their main statutory tasks in terms of data activities. It may not, therefore, be possible to assess these organizations using the proposed framework. In our test of the assessment framework, these organizations received a low organizational alignment score (though this may not be inaccurate). Second, the method's assessment of organizational alignment is partly based on a comparison between the characteristics of current data use and the planned use of big data. Although an organization will be more ready for big data if its current data use is more similar to a future big data use, our method somewhat neglects the possibility that organizations might make a larger and more discrete step in their data use when they implement big data, instead of more incremental development. This omission in the assessment method results in a lower assessment on organizational alignment for public organizations aiming for a larger transformation, even when the planned big data applications are fully suited to the main data activities and statutory tasks of the organizations concerned. This was not the case in the organizations we assessed, but it should be taken into account if the framework is used elsewhere.

Finally, the e-government growth stages model used here focused primarily on service provision functions within public organizations and the corresponding organizational structures and requirements. Yet, the public organizations we assessed did not all have an extensive and dominant service provision function. The five growth stages used may therefore not fully apply to all. Those organizations that were not aimed at service provision may have had no need to develop an organizational structure that would support e-government development. This might result in a lower maturity level assessment than would be given based on an assessment not designed with service provision in mind.

6.4 Future work

Before wrapping up with conclusions, a couple of suggestions can be offered for future academic endeavors. Working intensively with big data requires a particular approach to the way data and knowledge are created and managed, the way decisions are made, and the way governance and control activities are organized. The aim must be to ensure that data activities are in line with the goals and objectives of the organization and with stakeholders' needs. Organizations must realize that any radical change will have profound effects on all of these area, but academics too will need to update their models on collaboration, decision-making and governance, to include big data. One example from this study is the stage models,

which are prominent in the field of e-government. These should be updated to guide comprehensive ICT development in government, including but not solely focusing on, big data use. Furthermore, work on adoption and diffusion of innovations should be expanded to seek a better understanding of the antecedents and impediments to adoption of big data. Specific focus should be paid to interorganizational dependencies, many of which will be new, related to big and open data, new collaboration models for them, and their related transformation of government operations (Klievink et al. 2016). Another prospective focal point is the requirement for data management and governance that transcends organizational boundaries.

7 Conclusion

This paper set out to establish the readiness of public sector organizations for big data use. This quest was driven by the gap between the potential that big data seems to hold for the public sector (Chen and Hsieh 2014) and actual use of big data in government organizations (Mullich 2013). We suggested that the scarcity of actual big data use by the public sector may be related to uncertainties among big data champions and decision-makers about organizations' readiness for big data use. We combined literature from the domains of information systems and e-government to construct a readiness assessment tool, specifically attuned to big data. A side benefit of this approach was that, in the process, we operationalized big data in terms of the *use* of big data within organizations. This also served to demystify the big data concept, expressing it in terms of alignment with organizational goals, organizational maturity and the capabilities of an organization. This constitutes an advancement on definitions of big data solely in terms of its great promise or in terms of specific applications that already exist in some form.

This approach helped us to understand big data use opportunities and challenges in terms of existing organizations and the roles big data may have within organizations. Our big data readiness assessment framework helped us to understand why public sector organizations are lagging behind in big data use. It furthermore demonstrated that big data use can work for the public sector, if aligned with the public sector's organizational goals and its ways of working. Our test of the assessment framework in the Dutch public sector showed that overall, the Dutch public sector was not entirely ready for the large-scale introduction of big data and should further develop its readiness for big data use. In general, the organizations we assessed did not seem to fully understand that big data applications would add value to their organizations only if they supported and were supported by all the main organizational activities. Unlike organizations in the private sector, where big data can unlock new possibilities and enable new goals, the mere fact that big data and the tools to analyze it are available does not in itself constitute a value proposition for the public sector. These organizations' goals are often given, and they must guide big data use. Organizations may be technically capable of using big data, but they will not significantly gain from these activities if the applications do not fit their organizations and main statutory tasks. Organizations that were more experienced in using data seemed to better understand the organizational implications of big data, compared to organizations that used data less intensively, regardless of how set they were on the path towards actual big data use. Areas of improvement for the public sector were identified. These provide helpful pointers for practitioners seeking to improve the big data readiness of their organizations. They could also point national government towards areas where common organizational constraints can be overcome collectively.

Acknowledgments This open access publication was supported by the "Governing public–private information infrastructures" research project, which is financed by the Netherlands Organisation for Scientific Research (NWO) as Veni grant 451-13-020.

One of the authors wishes to thank the European Commission for its support for this research as part of the BYTE project (grant number 619551).

Appendix

A.1 Scorecard and scoring rules

A.1.1 Alignment readiness assessment

A. Big data application's fit with type of organizational statutory task and data intensity.

Highest scoring (>50 %) application fits statutory task: 3 points

Multiple high scoring (>50 %) applications, among which best fitting application: 2 points

Best fitting applications scores low (<50 %): 1 point

B. Big data characteristics currently present in data use, fitting with needed characteristics:

All three characteristics highly present (>70 %): 3 points

One characteristic not highly present (>70 %): 2 points

Two or more characteristics not highly present (>70 %): 1 point

Score (= A + B: range 2–6 points)	Alignment assessment
2	Very low
3	Low
4	Medium
5	High
6	Very high

A.1.2 Maturity readiness assessment

Maturity Level = highest level in which average score > 75 % and all three aspects are above 60 % and previous level also qualifies.

Maturity level	Alignment assessment
1	Very low
2	Low
3	Medium
4	High
5	Very high

A.1.3 Capabilities readiness assessment

Capabilities score is the average score of all separate capabilities. Separate capabilities scores are actual scores divided by the maximum score expressed as a percentage.

Average score (%)	Alignment assessment
0 %–60 %	Very low
61 %–70 %	Low
71 %–80 %	Medium
81 %–90 %	High
91 %–100 %	Very high

A.1.4 Overall big data readiness assessment

The overall score is the sum of the assessment scores on the three aspects of the framework.

Assessments	Score
Very low	1
Low	2
Medium	3
High	4
Very high	5

Readiness score range is therefore 3–15.

Readiness score	Big data readiness assessment
3–5	Ready for orientation on big data use
6–7	Ready to conduct research into big data use possibilities and requirements
8–9	Ready for planning and early design of big data use
10–11	Ready for further development of big data use
12–13	Ready for detailed development and testing of big data use
14–15	Ready for big data implementation

Big data readiness assessment based on the phases typically formulated in design and implementation of projects and used in typical IT roadmaps.

A.2 Example

The table below presents an example of the assessment for one of the organizations in our study.

Organization	1
Alignment	
Main statutory task type	Administration, management
Data collection intensity	High
Data use intensity	High
Best aligned application	All, but especially continuous monitoring applications
Corresponding big data characteristics	- Real time or near real time - Advanced analytics & algorithms - Innovative use of existing data
Application types	
Object/subject evaluation	82 %
Research	82 %
Continuous monitoring	82 %
Big data characteristics	
Internal & external	71 %
Structured & unstructured	57 %
Real time	71 %
Advanced analytics	71 %
Innovative data use	71 %
Alignment assessment	Very high
Alignment area for improvement	–
Maturity	
Stove-pipe organizations	*86 %*
Activities and information sharing	86 %
IT facilities	71 %
Data systems/concepts	100 %
Integrated organizations	*86 %*
Activities and information sharing	86 %
IT facilities	71 %
Data systems/concepts	100 %
Nationwide portal	*68 %*
Activities and information sharing	43 %
IT facilities	71 %
Data systems/concepts	90 %
Interorganizational integration	*71 %*
Activities and information sharing	86 %
IT facilities	29 %
Data systems/concepts	100 %
Demand-driven, joined-up government	*73 %*

(continued)

Activities and information sharing	71 %
IT facilities	71 %
Data systems/concepts	76 %
Maturity level	2
Maturity assessment	Low
Maturity areas for improvement	Sharing more activities and information with other organizations and creating IT facilities that make all data in the organization accessible by other organizations
Capability	
Data governance	81 %
Internal attitude	71 %
Data science expertise	62 %
Legal compliance	86 %
IT governance	81 %
IT facilities	81 %
External attitude	71 %
Capability average score	76 %
Capability assessment	Medium
Capability areas for improvement	Increasing data science expertise, while improving internal and external attitudes towards big data
Alignment assessment	Very high
Maturity assessment	Low
Capability assessment	Medium
Big data readiness score	10
Big data readiness assessment	Ready for further development of big data use

References

Adrian, M. (2011). *Information management goes' Extreme': The biggest challenges for 21st century CIOs*. SAS: Business Analytics and Intelligence Software.

Agarwal, R., & Dhar, V. (2014). Editorial —big data, data science, and analytics: the opportunity and challenge for IS research. *Information Systems Research, 25*(3), 443–448. doi:10.1287/isre.2014.0546.

Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U., Franklin, M., et al. (2011). *Challenges and opportunities with big data 2011–1*. Cyber Center Technical Reports. Paper 1. http://docs.lib.purdue.edu/cctech/1.

Andersen, K. V., & Henriksen, H. Z. (2006). E-government maturity models: extension of the Layne and Lee model. *Government Information Quarterly, 23*(2), 236–248.

Bertot, J. C., & Choi, H. (2013). Big data and e-government: issues, policies, and recommendations. In *Proceedings of the 14th Annual International Conference on Digital Government Research* (pp. 1–10). ACM.

Bharosa, N., Janssen, M., van Wijk, R., de Winne, N., van der Voort, H., Hulstijn, J., & Tan, Y. (2013). Tapping into existing information flows: the transformation to compliance by design in business-to-government information exchange. *Government Information Quarterly, 30*, S9–S18. doi:10.1016/j.giq.2012.08.006.

Brohman, M. K. (2006). Knowledge creation opportunities in the data mining process. In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*. IEEE.

Brohman, M. K., Parent, M., Pearce, M. R., & Wade, M. (2000). The business intelligence value chain: Data-driven decision support in a data warehouse environment: An exploratory study. In *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on* (p. 10–pp). IEEE.

Bryant, R., Katz, R. H., & Lazowska, E. D. (2008). Big-data computing: creating revolutionary breakthroughs in commerce, science and society. Computer Community Consortium.

Chen, Y.-C., & Hsieh, T.-C. (2014). Big data for digital government. *International Journal of Public Administration in the Digital Age, 1*(1), 1–14. doi:10.4018/ijpada.2014010101.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly, 36*(4), 1165–1188. doi:10.1145/2463676.2463712.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: a survey. *Mobile Networks and Applications, 19*(2), 171–209.

Clarke, R. (2016). Big data, big risks. *Information Systems Journal, 26*(1), 77–90. doi:10.1111/isj.12088.

Cordella, A., & Bonina, C. M. (2012). A public value perspective for ICT enabled public sector reforms: a theoretical reflection. *Government Information Quarterly, 29*(4), 512–520. doi:10.1016/j.giq.2012.03.004.

Cumbley, R., & Church, P. (2013). Is "Big Data" creepy? *Computer Law & Security Review, 29*(5), 601–609. doi:10.1016/j.clsr.2013.07.007.

Cunningham, S. W., & Thissen, W. A. H. (2014). Three business and societal cases for big data: which of the three is true? *IEEE Engineering Management Review, 42*(3), 7–9. doi:10.1109/EMR.2014.2341476.

Daniel, E. M., & Wilson, H. N. (2003). The role of dynamic capabilities in e-business transformation. *European Journal of Information Systems, 12*(4), 282–296.

Davenport, T. H., & Dyché, J. (2013). Big data in big companies. *International Institute for Analytics*. Retrieved from https://www.sas.com/resources/asset/Big-Data-in-Big-Companies.pdf.

Davenport, T. H., Barth, P., & Bean, R. (2012). How big data is different. *MIT Sloan Management Review, 54*(1), 43–46.

Dijcks, J.-P. (2012). Oracle: Big data for the enterprise. Oracle White Paper. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Oracle+:+Big+Data+for+the+Enterprise#0.

Ebrahim, Z., & Irani, Z. (2005). E-government adoption: architecture and barriers. *Business Process Management Journal, 11*(5), 589–611. doi:10.1108/14637150510619902.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic Management Journal, 21*(10–11), 1105–1121.

Ferro, E., Loukis, E. N., Charalabidis, Y., & Osella, M. (2013). Policy making 2.0: from theory to practice. *Government Information Quarterly, 30*(4), 359–368. doi:10.1016/j.giq.2013.05.018.

Finney, S., & Corbett, M. (2007). ERP implementation: a compilation and analysis of critical success factors. *Business Process Management Journal, 13*(3), 329–347.

Gantz, J., & Reinsel, D. (2011). Extracting value from chaos. *IDC Iview, 1142*, 1–12.

Groenfeldt, T. (2012). Morgan Stanley Takes On Big Data With Hadoop. Forbes. Retrieved from http://www.forbes.com/sites/tomgroenfeldt/2012/05/30/morgan-stanley-takes-on-big-data-with-hadoop/.

Gustafson, T., & Fink, D. (2013). Winning within the data value chain. *Strategy & Innovation (Innosight Newsletter)*, *11*(2). Retrieved from Winning within the Data Value Chain

Heeks, R. (1999). *Reinventing government in the information age: International practice in IT-enabled public sector reform*. London: Routledge.

Henderson, J. C., & Venkatraman, N. (1993). Strategic alignment: leveraging information technology for transforming organizations. *IBM Systems Journal, 32*(1), 472–484.

Henschen, D. (2012). Why Sears is going all-in on Hadoop. Information week. Retrieved July 1, 2014, from http://www.informationweek.com/it-leadership/why-sears-is-going-all-in-on-hadoop/d/d-id/1107038?.

Hota, C., Upadhyaya, S., & Al-Karaki, J. N. (2015). Advances in secure knowledge management in the big data era. *Information Systems Frontiers, 17*(5), 983–986. doi:10.1007/s10796-015-9593-y.

Janssen, M., & Kuk, G. (2016). Big and open linked data (BOLD) in research, policy, and practice. *Journal of Organizational Computing and Electronic Commerce, 26*(1–2), 3–13. doi:10.1080/10919392.2015.1124005.

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*(4), 363–368. doi:10.1016/j.giq.2015.11.007.

Janssen, M., Van Der Voort, H., & van Veenstra, A. F. (2015). Failure of large transformation projects from the viewpoint of complex adaptive systems: management principles for dealing with project dynamics. *Information Systems Frontiers, 17*(1), 15–29.

Jeyaraj, A., Rottman, J. W., & Lacity, M. C. (2006). A review of the predictors, linkages, and biases in IT innovation adoption research. *Journal of Information Technology, 21*(1), 1–23.

Joseph, R. C., & Johnson, N. A. (2013). Big data and transformational government. *IT Professional, 15*(6), 43–48.

Kamal, M. M. (2006). IT innovation adoption in the government sector: identifying the critical success factors. *Journal of Enterprise Information Management, 19*(2), 192–222.

Kankanhalli, A., Hahn, J., Tan, S., & Gao, G. (2016). Big data and analytics in healthcare: introduction to the special section. *Information Systems Frontiers, 18*(2), 233–235. doi:10.1007/s10796-016-9641-2.

Kelly, J. (2013). Amazon Web Services: 1 million Hadoop clusters and counting. Services angle. Retrieved July 1, 2014, from http://servicesangle.com/blog/2012/04/27/amazon-web-services-1-million-hadoop-clusters-and-counting/.

Kiron, D. (2013). Organizational alignment is key to big data success (interview with Randy Bean). Retrieved from: http://sloanreview.mit.edu/article/organizational-alignment-is-key-to-big-datasuccess/.

Klievink, B., & Janssen, M. (2009). Realizing joined-up government - dynamic capabilities and stage models for transformation. *Government Information Quarterly, 26*(2), 275–284. doi:10.1016/j.giq.2008.12.007.

Klievink, B., & Zomer, G. (2015). IT-enabled resilient, seamless, and secure global supply chains: Introduction, overview and research topics. In M. Janssen, M. Mäntymäki, J. Hidders, B. Klievink, W. Lamersdorf, B. van Loenen, & A. Zuiderwijk (Eds.), *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 9373). Cham: Springer International Publishing. doi:10.1007/978-3-319-25013-7.

Klievink, B., Bharosa, N., & Tan, Y.-H. (2016). The collaborative realization of public values and business goals: governance and infrastructure of public–private information platforms. *Government Information Quarterly, 33*(1), 67–79. doi:10.1016/j.giq.2015.12.002.

Laney, D. (2001). 3D data management: controlling data volume, velocity, and variety. *META Group - Application Delivery Strategies, 949*.

Layne, K., & Lee, J. (2001). Developing fully functional E-government: a four stage model. *Government Information Quarterly, 18*(2), 122–136.

Lohr, S. (2012). The age of big data. *The New York Times*. Retrieved from http://www.nytimes.com.

Malik, P. (2013). Governing big data: principles and practices. *IBM Journal of Research and Development, 57*(3/4), 1.

Margetts, H., & Sutcliffe, D. (2013). Addressing the policy challenges and opportunities of "Big data". *Policy & Internet, 5*(2), 139–146. doi:10.1002/1944-2866.POI326.

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data. A revolution that will transform how We live, work and think*. London: John Murray Publishers.

McAfee, A., & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard Business Review*. doi:10.1007/s12599-013-0249-5.

Milakovich, M. E. (2012). Anticipatory Government: Integrating Big Data for Smaller Government. In *Internet, Politics, Policy 2012: Big Data, Big Challenges*. Retrieved from http://ipp.oii.ox.ac.uk/sites/ipp/files/documents/Anticipatory%20Government%20Oxford.pdf.

Miller, H. G., & Mork, P. (2013). From data to decisions: a value chain for big data. *IT Professional, 15*(1), 57–59.

Ministerie van Infrastructuur en Milieu. (2014). *Voorbereiding thematische kenniskamer Omgaan met open data en big data*. The Hague: Workshop.

Mullich, J. (2013). Closing the big data gap in public sector. Bloomberg Businessweek. Retrieved from https://www.sap.com/bin/sapcom/fr_be/downloadasset.2013-09-sep-23-16.closing-the-big-data-gapin-public-sector-pdf.html.

OpenTracker. (2013). Definitions of big data. *OpenTracker*. Retrieved June 6, 2014, from http://www.opentracker.net/article/definitions-big-data.

Premkumar, G. (2003). A meta-analysis of research on information technology implementation in small business. *Journal of Organizational Computing and Electronic Commerce, 13*(2), 91–121.

Robey, D., Im, G., & Wareham, J. D. (2008). Theoretical foundations of empirical research on interorganizational systems: assessing past contributions and guiding future directions. *Journal of the Association for Information Systems, 9*(9), 497–518.

Romijn, J.H. (2014). Using big data in the public sector. Uncertainties and Readiness in the Dutch Public Executive Sector. Delft University of Technology.

Ross, J. W., Beath, C. M., & Quaadgras, A. (2013). You may not even need big data after all. *Harvard Business Review, 91*(12), 90–98.

Scholl, H. J., & Scholl, M. C. (2014). Smart governance: A roadmap for research and practice. In *iConference 2014 Proceedings* (pp. 163–176). iSchools.

Shiri, A. (2014). Linked data meets big data: A knowledge organization systems perspective. *24th ASIS SIG/CR Classification Research Workshop*, 16–20.

Simon, P. (2013). *Too big to ignore: The business case for big data*. Hoboken: Wiley.

Tambe, P. (2014). Big data investment, skills, and firm value. *Management Science, 60*(6), 1452–1469.

TechAmerica Foundation. (2012). Demystifying big Data*: A Practical Guide to Transforming the Business of Government. TechAmerica Foundation's Federal Big Data Commission*. doi:10.1109/MCSE.2011.99.

Tekiner, F., & Keane, J. A. (2013). Big data framework. In *Systems, Man, and Cybernetics (SMC), 2013 I.E. International Conference on* (pp. 1494–1499). IEEE.

UNPAN. (2014). *UN e-Government Survey 2014. E-Government for the Future We Want*. New York.

Valdés, G., Solar, M., Astudillo, H., Iribarren, M., Concha, G., & Visconti, M. (2011). Conception, development and implementation of an e-government maturity model in public agencies. *Government*

Information Quarterly, 28(2), 176–187. doi:10.1016/j.giq.2010.04.007.

Van den Toorn, M. (2014). Big data Kent restricties. PM Publiek Denken. Retrieved July 10, 2014, from http://www.pm.nl/artikel/3303/big-data-kent-restricties.

Weerakkody, V., Janssen, M., & Dwivedi, Y. K. (2011). Transformational change and business process reengineering (BPR): lessons from the British and Dutch public sector. *Government Information Quarterly, 28*(3), 320–328.

Wixom, B. H., & Watson, H. J. (2001). An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly, 25*(1), 17–41.

Yiu, C. (2012). *The big data opportunity: Making government faster, smarter and more personal*. London: Policy Exchange.

Dr. Ing. Bram Klievink is associate professor and head of the section Policy, Organisation, Law and Gaming at the Faculty of Technology, Policy and Management at Delft University of Technology. His work is on collaborative digital governance, public-private information infrastructures, smart logistics, living labs, and social and collaborative innovation. He was awarded a prestigious VENI grant in 2013 by the Netherlands Organisation for Scientific Research (NWO). Bram holds a PhD degree in technology and public management, a MSc degree in political science, and a degree in business information systems.

Bart-Jan Romijn obtained his MSc degree from the Technology, Policy & Management faculty of Delft University of Technology in 2014 and worked on this research as a graduate intern. He currently works as a consultant for KPMG Management Consulting in the Netherlands and supports financial services companies, primarily banks and insurance companies, with the development of Big Data initiatives that improve regulatory compliance, risk assessment and customer service.

Scott Cunningham is an associate professor of policy analysis at the Delft University of Technology. He researches the social and political impacts of big data. Prior to joining TU Delft he worked as a data scientist in California, where he helped companies in manufacturing, retail and e-commerce make the best use of their data in supporting decision-making.

Hans de Bruijn is Professor of Public Administration, Organisation and Management at Delft University. His research is on networked, multi-actor governance, both between and within organizations. Areas of application are the various utility sectors, large infrastructural projects and environmental issues. Much of his research is also on management within networked organizations – e.g. on performance management, decision making, knowledge management, coordination.