On State-Level Architecture of Digital Government Ecosystems: From ICT-Driven to Data-Centric

Preprint · February 2021

DOI: 10.13140/RG.2.2.16917.45283

3 authors, including:

Dirk Draheim
Tallinn University of Technology
222 PUBLICATIONS **1,022** CITATIONS

SEE PROFILE

On State-Level Architecture of Digital Government Ecosystems: From ICT-Driven to Data-Centric

Dirk Draheim[1], Robert Krimmer[2], and Tanel Tammet[3]

[1] Information Systems Group, Tallinn University of Technology, Estonia
dirk.draheim@taltech.ee
[2] Johan Skytte Institute of Political Studies, University of Tartu, Estonia
robert.krimmer@ut.ee
[3] Applied Artificial Intelligence Group, Tallinn University of Technology, Estonia
tanel.tammet@ttu.ee

Abstract. The "digital transformation" is perceived as the key enabler for increasing wealth and well-being by politics, media and the citizens alike. In the same vein, digital government steadily receives more and more attention. Digital government gives rise to complex, large-scale *state-level* system landscapes consisting of many players and technological systems – and we call such system landscapes *digital government ecosystems*. In this paper, we systematically approach the state-level architecture of digital government ecosystems. We will discover the primacy of the state's institutional design in the architecture of digital government ecosystems, where Williamson's institutional analysis framework supports our considerations as theoretical background. Based on that insight, we will establish the notion of data governance architecture, which links data assets with accountable organizations. Our investigation results into a digital government architecture framework that can help in large-scale digital government design efforts through (i) separation of concerns in terms of appropriate categories, and (ii) a better assessment of the feasibility of envisioned digital transformations. With its focus on data, the proposed framework perfectly fits the current discussion on moving from ICT-driven to data-centric digital government.

Keywords: digital government · e-government · e-governance · new institutional economics · digital transformation · data governance · consent management · data exchange layers · X-Road · GAIA-X

1 Introduction

The so-called "digital transformation" is currently perceived as a – or even *the* – key enabler for increasing wealth and well-being by many in politics, the media and among the citizens alike; and we find digital transformation initiatives as crucial building blocks in today's political agendas in all countries, recently, also under the keyword "smart city". In the same vein, digital government steadily

receives more and more attention by governments, actually, ever since it became mainstream in the 1990s under the name e-Government [1]. From all over the world, we hear about many great success stories in digital government. At the same time, when we look into concrete digital government projects, we often see tremendous project expenditures (with millions or even billions of USDs, even for single projects, are not an exception). Over and over again, we see project failures with tremendous cost overruns (millions, billions) and time overruns (years), with results far below initial expectations or even complete project abortions. Note, that we are not even talking about massive digital government initiatives here, i.e., often, these problems already show in single digital government projects that aim at realizing a single digital administrative process or delivering a single e-service, e.g., a concrete e-health information system, a single e-court system, a tax declaration service etc. How come? Our hypotheses is that essential aspects of the *state-level* ecosystem, in which the several single digital government solutions are realized are neglected (or: unknown, overlooked, not addressed properly etc.) And indeed, digital government gives rise to complex-adaptive systems [46], which are, actually, large-scale, *state-level* system landscapes consisting of many players (authorities, companies, citizens) and technological systems – we call such system landscapes *digital government ecosystems*.

Therefore, the purpose of this paper is to systematically examine the state-level architecture of digital government ecosystems. As a crucial step, we will discover the primacy of the state's institutional design, which, as we argue, provides the core reference point for all system design efforts in digital government. Based on that insight, we will establish the notion of data governance architecture. A data governance architecture links data assets with accountable organizations and represents the essence of co-designing institutions and technological systems of a digital government ecosystem. In our endeavours, Williamson's institutional analysis framework will support us as a valuable theoretical background. Our investigation results into a digital government architecture framework that can help in large-scale digital government design efforts through (i) separation of concerns in terms of appropriate categories, and (ii) a better assessment of the feasibility of envisioned digital transformations.

Following the UN e-Government Survey 2020, a dominating theme in digital government is to reach the ideal of a data-centric digital government. In a data-centric digital government, data would be used pervasively in decision-masking at all organizational levels and, beyond this, would enable the continuous optimization and innovation of people-centric services. With current Big Data [31, 47] and data science technologies [45], the necessary tools are available to realize such a data-centric digital government vision; however, yet, such data-centric digital government is far from becoming the standard. With its focus on data, the proposed architectural framework perfectly fits the current discussion on moving to data-centric digital government. In particular, it can help in identifying and understanding obstacles in the implementation of data-centric digital government.

The paper proceeds as follows. In Sect. 2, we briefly review the discussion of data in the current digital government discourse. We look into the UN e-Government Survey 2020 and what it tells us about the role of data in digital government. The section is meant to serve as background information and motivation. In Sect. 3 we explain, why institutions matter in the design of digital government architectures. We discuss Koppenjan and Groenewegen's systems design framework and briefly explain Oliver Williamson's institutional analysis framework. In Sect. 4 we establish the notion of data governance architecture and the notion of digital government solution architecture; we explain, how they relate to each other and to the state's institutional architecture. We arrange these components (data governance architecture, solution architecture, institutional architecture) into an architectural framework and aim at explaining the mutual dependent dynamics of changing these components. In Sect. 5, we review a series of digital government technologies against the background of the proposed architectural framework. Here, a digital government technology is a technology that have been explicitly designed for digital government or is otherwise relevant for building digital government systems. Among others, we look in the Estonian X-Road data exchange layer, the European federated data infrastructure initiative GAIA-X, and Tim Berner Lee's web-decentralization project Solid (Social Linked Data). The purpose of the section is, on the one hand, to reinforce the line of arguments embodied in the architectural framework and, on the other hand, to provide some confidence in the industrial-strength applicability of the framework. In Sect. 6 we further discuss the suggested digital government architecture framework. We discuss, in how far the framework can help in large-scale digital government design efforts. Also, we discuss how the framework is placed in the tension between e-democracy and e-administration. We finish the paper with a conclusion in Sect. 7.

2 From ICT-Driven to Data-Centric

In today's organizations, IT is about *data processing*, about the collection and manipulation of data in support of the business processes [15, 17]. But it is also about reporting on the basis of available data, in service of decision making [28] and knowledge management [53]. It is similar – at a higher level – in digital government. In digital government, we have a great deal of ICT being used to make administrative processes in the authorities as well as in between authorities more efficient and effective, however, the huge potential is now in exploiting data for better decision making and leveraging innovations. In that vein, the UN Agenda for Sustainable Development 2030 has stated:

> *"Quality, accessible, timely and reliable disaggregated data will be needed to help with the measurement of progress and to ensure that no one is left behind."* [69] (1)

In the note [68], the Committee of Experts on Public Administration of the UN Economic and Social Council identifies three main principles of effective

Approach	Description
ICT-driven	Where Governments are highly influenced by the use of new and existing information and communications technology (ICT).
Data-informed	Where Governments are guided by data; data play an inferential role in policymaking, with the understanding that data will inform rather than drive decision-making because there are rational, political and moral elements of decision-making and data are just one important aspect of the process [63].
Data-driven	Where Governments use analytics and algorithms in decision-making (elaborated in a recent OECD working paper on a data-driven public sector) [71].
Evidence-based	Where policy approaches reflect the practical application of the findings of the best and most current research available [...]
Data-centric	Where Governments place data and data science at the core of public administration; data are seen as a key asset and central to government functions and are leveraged for the provision, evaluation and modification of people-centric services [14].

Table 1. "Data as a key resource for Governments" [67]; literally compiled from the SOURCE: E-Government Survey 2020 [67], p. 150.

governance for sustainable development: *effectiveness*, *accountability* and *inclusiveness*. The UN e-Government Survey 2020 [67] systematically has screened the indicators and strategies that are connected to these three principles for those that are directly or indirectly related to data. And it finds many of them; here, we list a selection of them (for the full list see Table 6.2 in [67], p. 149): "investment in e-government", "monitoring and evaluation", "strategic planning and foresight", "results-based management", "performance management", "financial management and control", "risk management frameworks", "science-policy interface", "network-based governance", "open government data", "budget transparency", "independent audit", "participatory budgeting", "long-term territorial planning and spatial development"[67, 68]

All of this strongly indicates the relevance of data for digital government. Different countries utilize data following different approaches, with different attitudes. The UN e-Government Survey 2020 [67] distinguishes between five such approaches: (i) *ICT-driven*, (ii) *data-informed*, (iii) *data-driven*, (iv) *evidence-based*, and (v) *data-centric*, see Table 1. Those are not merely qualitative characterizations of different possible approaches, but, clearly, the UN survey wants to express a "ranking" with the sequence (i)–(v), as it states the following in regards to Table 1 (Table 6.3 in [67]): "Table 6.3 shows the different approaches countries take and reflects a progression of sorts, illustrating how government data are increasingly leveraged for effective governance."[67] In that sense, the data-centric digital government seems to be the ideal to be reached. The question remains, what such *data-centric digital government* should be. It is clear, as it is put at the top of the ranking imposed by (i)–(v), that data are used here most pervasively (as compared with (i)–(iv), and most strategically, i.e., as

"key asset"). As eventual purpose, it is said that "data [...] are leveraged for the provision, evaluation and modification of *people-centric* services". Sure, *people-centric* sounds splendid – an eyecatcher. But what are *people-centric* services? Are they services to the citizens? Or to the government? Is people-centricity just a synonym for inclusiveness, which would be nice, or is it something else? Is it something that the citizens actually want? Without a definition, it is not possible to answer such and similar questions. The case study provided by the UN Survey is frightening in this regard, i.e., "The data-centric online-offline integration of digital government in Shanghai" (Box 6.1, p. 157 [67]). The digital government described in this success story incorporates the super-application [43] WeChat – a central building block of China's futuristic next-generation citizen surveillance programme, see the respective Human Rights Watch web page[4], compare also with, e.g., [49, 2]. The Gartner Group Report of Andrea Di Maio [14], which is given as reference for the data-centric digital government by the UN survey, does not help clarifying the concept of people centricity – it does not mention the notion of *people-centric* at all.

There are still vast, yet unused, opportunities to exploit data at state level to better the government's effectiveness, accountability and inclusiveness. At the same time, there are huge risks that citizens' data are exploited for citizens' monitoring and control. The challenge is in getting the data governance structure right. And this challenge needs to be understood early in all digital government design issues and, therefore, needs to be reflected in each approach to digital government architecture.

3 On Large-Scale ICT Systems and Institutions

Digital government ecosystems need to be analyzed and designed as socio-technical systems. In analysis of digital government ecosystems, we can receive guidance from Bruno Latour's actor-network theory [40], which "treats the social and the technical as inseparable" [72]. When it comes to design, a digital government ecosystem needs to be *co-designed* with respect to its *institutional architecture* and its *technological assets*.

In [37], Koppenjan and Groenewegen have provided a framework for the co-design of technological assets and institutions of complex, large-scale, technological systems – with foundations in Williamson's *new institutional economics* [74]. The class of systems that are addressed by Koppenjan and Groenewegen can be characterized simply as exactly those systems that have institutions as part of their solution, in particular, with resp. organizations that are not merely consumers of the solution, but make essential contributions to the solution, and that "have institutions" (i.e., have institutional setups) that matter. The class of these systems encompass systems such "energy networks, water management services (drinking water, sewage, protection, management), waste treatment, transport systems (rail, road, water, tube), industrial networks, information systems and

[4] https://www.hrw.org/tag/mass-surveillance-china

Level	Purpose	Frequency
L1 (social theory) *Embeddedness*: informal institutions, customs, traditions, norms, religion	Often noncalculative; spontaneous.	100–1000 years
L2 (economics of property rights) *Institutional Environment*: formal rules of the game – esp. property (polity, judiciary, bureaucracy)	Get the institutional environment right. 1st-order economizing.	10–100 years
L3 (transaction cost economics) *Governance*: play of the game – esp. contract (aligning governance structures with transactions)	Get the governance structure right. 2nd-order economizing.	1–10 years
L4 (Neo-classical economics / agency theory): resource allocation and employment (prices and quantities, incentive alignment)	Get the marginal conditions right. 3rd-order economizing.	continuous

Table 2. Economics of institutions; literally compiled from SOURCE: Williamson 1998 [74].

telecommunication networks, city service [...]."[37]. Rather obvious, many digital government solutions would fall into this category. No later than when it comes to whole digital government ecosystems at the level of states, the whole system can be conceived as belonging to this system class. For example, we have successfully used Koppenjan and Groenewegen's framework to compare the digital government ecosystems of the Netherlands and Estonia [7].

Koppenjan and Groenewegen's (henceforth: KaG's) framework deals with the design process, and also with questions of the *design of the design process* (called 'process design' in [37]), and elaborates a four-level model for institutional analysis, which is ingrained in Williamson's institutional analysis framework (the details of differences between KaG's institutional model and Williamson's framework are not relevant to the discussions in this paper and we will not delve into them). KaG's framework fits scenarios, in which a solution is designed from scratch, as well as scenarios, in which some institutions already exist and need to become subject of re-design. The key insight (key takeway) from KaG's framework for our framework in Sect. 4 is that, whenever the shape of institutions is crucial for a solution, it needs to be incorporated into the design efforts of the solution. However, beyond that, we do not want our framework in Sect. 4 to be understood as a specialization KaG's framework, and also not as an extension of Kag's framework. We step from considering the design of large-scale solutions to the design of ecosystems of solutions. Such an ecosystem consists of many solutions (in our case, typically, a phletora of solutions; each owned by a different organization). At the same time, we specialize to digital government, which means, in particular, that our architectural considerations are always at *state-level*, i.e., address the state as a whole. Furthermore, we can assume that digital government ecosystems are never build from scratch, as a result of the existing institutional backbone. Building digital government ecosystem is, in major chunks, about adjusting and re-designing institutions.

We want to choose Williamson's institutional analysis framework as a theoretical underpinning. In Table 2, we have compiled the "four levels of social analysis" of the framework (from Fig. 1 in [74]). In [74, 75], the analysis framework is presented as part of wider discussions of *new institutional economics*. Institutions at the several levels, L1 trough L4, continuously evolve, at different pace and with different volatility; where they all influence each other (back and forth, even across several levels) in this evolvement. Level L1 is about culture at the societal level. Here, the level of analysis is about history and social science [29, 52]. Level L2 is about laws, regulations, government, i.e., formal rules. The investigation of this level dates back to Ronald Coases 'The Problem of Social Cost' [12]. Level L3 is about organizational governance, in particular, in so far it concerns inter-organizational transactions. It is the level of Coases 'The Nature of the Firm' [13]. Level L4 is the level of neoclassical economics (price/output, supply/demand etc.) as well as agency theory [32].

An institution is a compound of informal rules (social norms, customs, traditions, commitments etc.) and formal rules (legislation, regulations, contracts etc). An organization is an *organized* group of people. Organization adhere to institutions. Sometimes, *institution* is used to denote a group of people, for example, it might be that the "family" is called an institution. However, here we would usually not mean a particular group of people but rather the set of typical norms that shape families and that families adhere to, i.e., the notion of family. Similarly, in everyday language, some organizations are often called institutions, in particular, organization from the public sector: the police, a particular university etc. Douglass North defines institutions as "as humanly devised constraints that structure political, economic, and social interactions. They consist of both informal constraints (sanctions, taboos, customs, traditions, and codes of conduct), and formal rules (constitutions, laws, property rights!)" [54].

When we explain our digital government architecture framework in Sect. 4, we use the notions of

- state's institutions,
- (state's) institutional architecture,
- accountable organizations.

The *state's institutions* encompass all kinds of informal and formal rules existing in the society, in particular legislation, in so far they are relevant for government. We do not attempt a precise definition of "relevant for government" here. The notion of the state's institutional architecture is almost synonym to the state's institutions. It merely stresses the fact, that the state's institutions show mutually dependencies and interplay with each other. For the sake of the paper, the *accountable organizations* are deliberately formed, formal organizations and encompass all kinds of organizations from the public sector (agencies, authorities, offices, bureaus, commissions, chambers, chancelleries, public bodies, ministries, etc.) and organizations from the private sector (companies on behalf of public-private partnership, non-governmental organizations, associations etc.). Throughout the paper, we use also *state's organizations* for the organizations from the public sector for short.

4 Digital Government Architecture

This section aims at elaborating a digital government architecture framework, depicted in Fig. 1, that is, essentially, based on the following line of hypotheses:

- The state's institutions are formed following the state's functions. The entirety of the state's institutions, how they are shaped and the way how they interplay makes the state's *institutional architecture*. The institutional architecture usually changes slowly. More precise: substantial changes to the institutional architecture, i.e., those that are the result of societal change, usually occur non-disruptively and take significant time.
- The state's institutional architecture determines the state's *data governance architecture*. The data governance architecture links data assets with accountable organizations along two dimensions: the *interoperability* dimension and the *provisioning* dimension.
- The data governance architecture limits the design space of the *digital government solution architecture*, which consists of all *digital administrative processes* and delivered *e-services*, i.e., those assets that are eventually perceived as digital government by end-users and citizens. The digital government solution architecture can show small, ad-hoc and fast changes.
- Changes in the institutional architecture are so severe, that they can trigger immediate changes in the digital government solution architecture, whereas changes in digital government solution architecture (usually) can only have a long-term influence on changes in the institutional architecture.

In our framework, we say that the data governance architecture and the digital government solutions architecture together form the *digital government architecture*. The data governance architecture forms the backbone, or let us say, the core of the digital government architecture that deals with the necessary fulfilment of data governance; whereas the solutions architecture addresses all kinds of quality aspects of the offered solutions, i.e., usefulness, adherence to good service-design principles, maturity of processes etc.

It is important to note that the discussed architectural framework is not limited to transforming the classic services of public administration or what we would call *e-administration*. Our concept of e-service delivery in Fig. 1 definitely encompasses all e-services, also from the realm of what we would call *e-democracy* including initiatives such as open government data, e-participation, or i-voting. In the current section, we rather not delve into a discussion of the different kinds of state functions. This is deferred to Sect. 6, where we look into digital government architecture in the tension between e-democracy and e-administration.

We claim that a key to understand architecture of digital government ecosystems is in understanding *data governance*. In the context of digital government, data governance is an ultra large-scale, cross-organizational challenge. As a next step, we need to discuss the most important data governance principles in Sect. 4.1, before we can continue with a definition of data governance architecture in Sect. 4.2.

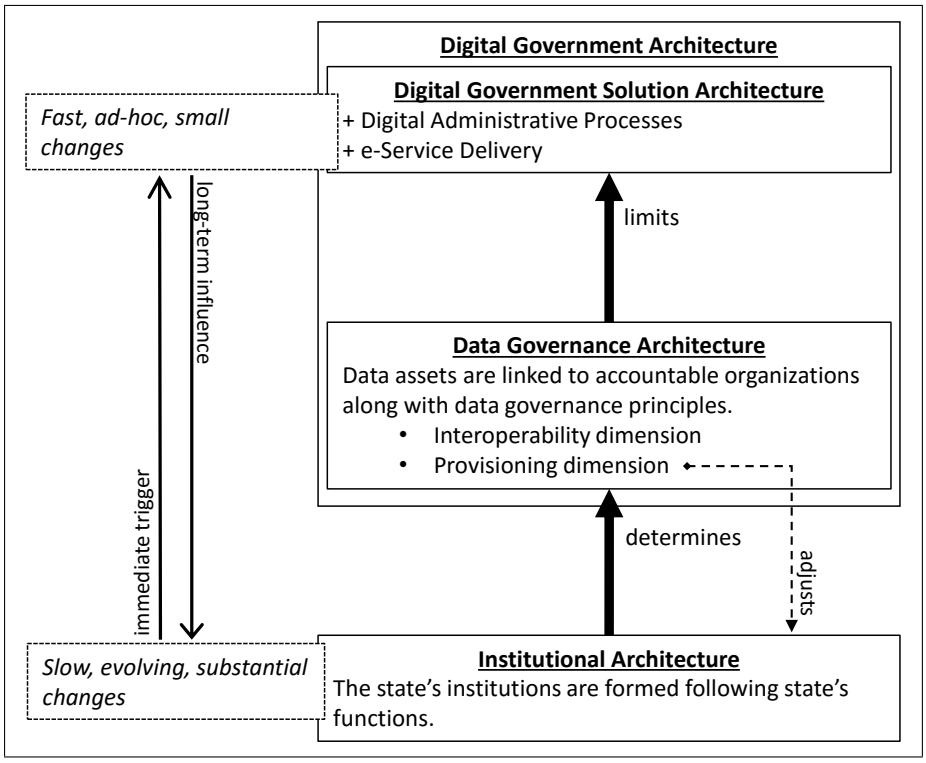

Fig. 1. A digital government architecture framework.

4.1 Data Governance Principles

When it comes to the single authority or organization, data governance is always about *responsibility* for adherence with *data governance principles*; often, it is about *accountability* for adherence with data governance principles. Accountability goes beyond responsibility, i.e., it is given whenever the relevant data governance principles are subject to laws and regulations. If data governance principles are merely recommendations or best practices, we call them *soft* data governance principles; if they are subject to laws or regulations, we call them *hard* or *strict* data governance principles.

In today's digital government initiatives, the following (partially overlapping and mutually dependent) categories of data governance principles can be identified:

– *Data Protection Principles.*
  - *Minimality Principles.* Citizens' data are collected, stored and processed only for *defined* purposes and for *defined* time periods. Data are deleted, if the purpose of its storage becomes obsolete. Depending on the *data category*, the citizen has the right to enforce the deletion of his or her data.

  - *Transparency Principles.* The citizen has the right to know, for which purposes and time periods his or her data are collected, stored and processed. As an advanced principle, for certain *data categories*, he or she has the right to know the data processing history, i.e., *who* has accessed his or her data *when* and for *which* purpose.
  - *Consent Principles.* Depending on the *data category*, the citizen has the right to determine whether his or her data are stored. Consent can be granted resp. withdrawn as *opt-in* or *opt-out*.
- *Data Quality Principles.* Correctness and consistency (resp. non-redundancy), including referential integrity (which is particularly challenging in cross-organizational settings such as digital government [64]). Or, in terms of ISO 25012 [30]: accuracy; completeness; consistency; creditability; currentness. Etc.
- *The Once-Only Principle.* The once-only principle (OOP) [38, 36] is about ensuring that "citizens and businesses supply the same information *only once* to a public administration"[23].

Observe that actually following these principles may come with a cost, and higher costs in combination with little to no direct gain for the state agencies will make it less likely that the principles are actually followed. Assessing the costs and gains of a data governance principle in a concrete scenario is a difficult, complex endeavor. The cheapest among these appears to be the minimality principle: collecting less data is typically cheaper than collecting more. On the other hand, deleting data after a prescribed time period is already associated with a cost. Similarly, the data quality principles are important for the functioning of the e-administration and are thus expected to be followed. On the other hand, transparency, consent and the once-only principles provide rather few – if any – immediate benefits to the e-administration; and their implementation is associated with significant costs. Hence, these principles are always less likely to be followed.

On Data Consent The relationship between the citizen and the state authorities is different from the relationship between a company and its customers, as it is again different between a company and their employees. This matters in so far, as the citizen has no consent right for all kind of data. In accordance with the minimality principle, the state can consider (and regulate) certain citizen data as critical; only beyond this, the consent principle applies. Following the minimality principle, we would expect that critical data are usually master data; whenever it comes, e.g., to log data, trajectory data, or any kind of aggregated personal data, we would expect that a consent principle is granted. Super-application such as Tencent's WeChat in China, as we have described in [43], are the counter-example. We would not count such an application as digital government. In the ideal world, each digital government initiative is expected to be in service of strengthening our democracies, independent of whether the concrete initiative follows a democratic or rather a technocratic narrative [19].

Data consent comes with several synonyms, each with different flavors and different, yet overlapping communities, such as MyData, Self Data, Internet of Me, or PIMS (Personal Information Management Services), compare with the MyData Declaration[5].

Consent Management The first question of consent management is whether and how citizens can actually block or enable data collection and processing. We argue that building such mechanisms is complex, expensive and may lower the efficiency of e-administration, despite the positive aspects such as increased trust. For example, in the digital government system practiced in Estonia there are almost no consent management mechanisms for citizens. The first pilot project for managing consent (in the context of providing health data to insurers) will be launched in 2021. Instead, a specialized state authority acts as an overseer of digital government systems with the goal to block the unnecessary collection of citizen data, i.e., enforcing the minimality principle. As an additional mechanism, several digital government systems participate in a monitoring system enabling citizens to see when and why their personal data has been transferred from one organization to another. However, participation in this monitoring system is not obligatory and most state organizations do not participate in it, apparently in order to avoid related development costs.

Need for systematic consent management arises also as companies enter the scene. The players in a digital government ecosystem are not restricted to government authorities. Companies and other organization, that are no government authorities can be involved on behalf of public-private partnership (PPP) [55].

On the Once-Only Principle and Data Consistency At a first sight, the once-only principle (OOP), looks rather like a *service design principle* than a data governance principle, as we have said that it is about not asking the same data from the citizen more than once. It becomes a data governance principle as it can only be resolved by joint coordination efforts of all authorities of the digital government ecosystem together. Sometimes, you can hear that the once-only principle is about not storing a data item in more than one location. But it is not. As an example, a simple act of transferring data from one system to another immediately yields a replication of data in different locations. It is also usual – and architecturally sensible – to cache some of the transferred data for longer time in order to avoid frequent requests for the same data items from external sources. Similarly, adherence to the once-only principle cannot help with *data consistency*, which is a data quality principle. Similarly, it can not help resolving lack of consistency in cross-organizational transactions (*long transactions*) [20]. All in all, the OOP appears to be beneficial to the citizens interacting with the e-services, but not so much to the efficiency of e-administration.

[5] https://mydata.org/declaration/

4.2 Data Governance Architecture Defined

Each digital government ecosystem has a data governance architecture. A data governance architecture links data assets to accountable organizations along with data governance principles.

First, there is a primary institutional design of the state authorities. This design follows the branches of the state's government with all its entities from the executive, judiciary, and legislature – embodying the entire public administration. The authorities are designed following the functions to be fulfilled by the state's government, following the principle of separation of powers, implementing checks and balances and targeting good governance principles. This primary institutional design is hierarchical and cannot, in general, be arbitrarily changed. Of course, we sometimes see that ministries are re-shaped, e.g., a super ministry might be formed by merging; a new ministry might appear in a legislation period and again disappear in the next etc. However, at the lower level, changing the primary institutional design amounts to major efforts or even reforms.

The state authorities need to collect citizen data to fulfill there functions. By collecting and processing citizen data they become accountable for the fulfillment of data governance principles. This is how data governance architecture is determined by the institutional design. Actually, we can see now that the data governance architecture is the architecture of the digital government ecosystem *per se*. Indeed, data governance may be determined by legislation. For example, in the case of Estonia, both the obligation and right to collect and store specific kinds of data are given to organizations by lower-level legislative acts. These acts then become a primary enforcer for the creation or modification of corresponding IT systems.

A data governance architecture achieves the following. It creates a correspondence between data assets and accountable organizations together with lists of specified data governance principles. More precisely:

> A *data governance architecture* specifies for each data asset α, each accountable organization ω, and each data governance principle γ, in how far exactly ω is accountable for α in regards of γ. (2)

The complete description of a data governance architecture as defined in (2) can quickly become quite complex, because there can be overlaps. Several different organizations might be accountable with respect to the same data asset and data governance principle. Then, it needs to be clarified, what there specific roles are with respect to this data asset in regards to the resp. data governance principle and how they interplay. Actually, we have used accountability as a rather broad term in (2). An accountability in (2) can come in various forms and need to be specified in each single case. In particular, accountability comes with different levels of strictness; a typical approach is, e.g., to distinguish accountabilities in a range from *accountable* (in a more narrow sense then) as the most strict notion (as hard legal accountability) over *responsible* to *consulted* [58]. We will make no attempt here to elaborate a concrete data governance specification approach.

We have seen that the data governance architecture *essentially* follows the primacy of the institutional design; but, at the same time there is some degree of freedom in how the concrete data governance architecture materializes. This degree of freedom shows in two dimensions:

- IT system interoperability
- IT service provision

the first dimension is about *IT system interoperability* (or just *interoperability* for short) and the second dimension is about *IT service provision* (or just *provisioning* for short). The distinction of these two dimension is crucial in design efforts for digital government solutions; we delve into the IT system interoperability dimension in Sect. 4.3 and the IT service provision dimension in Sect. 4.4.

4.3 IT System Interoperability

IT system interoperability and its objective, i.e., IT system *integration* [76], form a major strand of digital government efforts [25] and digital government research [62, 41, 42]. We explain interoperability via the transformation of data governance architecture in service of strengthening data governance principles as follows. In principle, there is no need for interoperability. Each authority of the primary institutional design could collect and hold all the citizen data it needs. In such a trivial, ad-hoc data governance architecture it is likely to have many data assets redundantly held in several authorities, resulting in significant issues: overall lower data quality, potential inconsistencies, violation of the once-only principle, higher risk of violation of data protection principles (minimality, transparency), difficulties in consent management; hand-in-hand with an increased amount of stakes/efforts in accountability. If an authority stops to collect and hold citizen data, it has to request the data it needs (as transactional data) from peer authorities that grant them access to these data assets, i.e., *interoperability* becomes necessary. The introduction of interoperability can change the data governance architecture to a better one. We see that interoperability cannot be simply explained as the result of *legacy system integration*, instead, it can be shaped along with the design space of the data governance architecture.

Interoperability has several key aspects:

- *Physical Access to Data.* For one organization/system, in order to access data of another organization/system, special parts of the IT systems have to be built. Typically, these come in the form of APIs (Application Programming Interfaces) enabling one system to query specific data from another.
- *Access Management.* Since the data to be transferred is typically not public, both systems must be able to verify the identity of the other system and the existence of actual rights or agreements for transferring data.

– *Semantic Interoperability.* Both the nomenclature, the meaning and the way data are encoded in one system may be different from how it is understood or encoded in another. To facilitate the use of external data, special translation systems have to be developed.

The Estonian governmental data exchange platform X-Road described later in Sect. 5.1 targets both the physical access and access management aspects, but not the semantic interoperability aspect. Indeed, as the number of different systems and their interconnections grows, the semantic interoperability is becoming the most expensive part of interoperability. Due to the immaturity of the technology devoted to semantic aspects of data, this component is typically implemented by non-automated programming, i.e., costly analysis and development work which is repeated anew for most new connections made between different systems.

4.4 IT Service Provision

Multi-tenancy, IT service provision, cloud computing: different terminology for almost the same thing (differences are in the decade, in which the terminologies have been used mostly; and in the technology stack usually connected to them). An IT service provider can increase resilience (this way decreasing risk) and take over responsibility, this way lowering the accountability stake. Often, IT service provision is assessed as the exact opposite by accountable stakeholders, i.e., as increasing risk. This is so, if the stakeholder does not trust or cannot trust the IT service provider (e.g., due to the lack of a sufficient regulatory framework – think of the "safe harbor" debate alone), i.e., if he or she needs to consider the IT service provider as the risk in itself. For example, it is common for governments to require that data managed by the state's organizations has to be stored in the servers physically located in the country. Even more, often government prefer to store data in data centres over which they have direct control: these data centers, on the other hand, may or may not be owned and managed by companies on behalf of public-private partnership.

The introduction of an IT service provider makes this provider a player in the digital government ecosystem. The introduction of an IT service provider changes the data governance architecture, but more fundamentally as in the *interoperability* dimension that we described in Sect. 4.3. The *interoperability* dimension is about shaping accountabilities only; the *provisioning* dimension is indeed about changing the institutional architecture, as indicated by the dashed arrow in Fig. 1. These changes are usually conservative changes, i.e., they extend the existing institutional architecture without changing roles of existing institutions and the interplay between existing institutions. For example, the establishment of a national central data center that hosts the data of the state agencies might not be considered a change to the institutional architecture, but it actually is. The role of the data center needs to be fixed and legally underpinned. The interplay of the data center with the other agencies need to be regulated and established. Now we are in a dilemma. Once the data center is fully established

as an organization, provisioning turns out to be interoperability in our framework (and technically, i.e., disregarding the distinction in our framework, ICT provisioning is a form of interoperability anyhow). So, how can we distinguish ICT provision from interoperability any longer? And: should we we distinguish ICT provision from interoperability at all? We could distinguish provisioning from interoperability by introducing a notion of *genuine* functions of the state as opposed to *digital-government*-related functions. Actually, it is fair to state, that provision of digital solutions itself is a distinguishable function of the state. With respect to the question, whether we should maintain the distinction between provisioning and interoperability; yes, we think it is important. Actually, it is a distinction that is of utmost importance in practice, for example, it shows in the distinction between *data controllers* and *data processors* in the European GDPR regulation (General Data Protection regulation), as we will explain in due course in Sect. 4.6. Also, the differences are often overlooked or neglected, when it comes to discussion of alternative architectural approaches and styles, in particular, in the analysis of centralization vs. decentralization – we will delve into a discussion of centralization vs. decentralization again in Sect. 4.6.

4.5 Evolving Digital Government

The digital government solution architecture shapes the digital administrative processes and the e-service delivery of the government authorities. Digital administrative processes run inside the authorities and inter-organizational, between the authorities. e-Services are delivered to citizens and companies and allow for triggering digital administrative processes. However, digital administrative processes need no e-service to be triggered. Administrative processes can be digitized (typically always for the sake of making them more efficient and effective) without re-shaping the interaction with the citizens and companies. In public perception and discussion, digital government is often identified with e-service delivery. However, many digital administrative processes actually run without being triggered through e-services. The ratio of administrative processes that are triggered by e-services is actually an interesting, however, often hard to assess or even hard to estimate, indicator for the maturity of digital government, compare with Sect. 5.1.

The state's institutional architecture does not change quickly. Despite smaller adjustments, changes to the institutional architecture are severe, as they reflect changes to the state's functions. A change to the institutional architecture can immediately trigger changes in the solution architecture. Often, in practice, we observe, that changes to the digital government architecture can then be very cost-intensive or even fail. The reason for this is, in general, in the efforts needed to adjust the interoperability and the provisioning dimension of the data governance architecture. Digital government projects that deal with administrative processes inside a single authority can be executed, in principle, without changing the data governance architecture. Still, those projects often fail. But the reason for such this is only in lack of ICT maturity, i.e., with respect to IT governance, IT management, used ICT technologies etc. These are practical problems

that are not specific to digital government but can concern all large-scale ICT projects. The question, whether public organization typically rather have a lower ICT maturity (as compared to private companies) is independent of that.

Small, ad-hoc changes to the solution architecture are always possible – as long as these changes do not require changes to the data governance architecture. Such small changes do not lead to direct changes in the institutional architecture. They can contribute, in the long run, to a change of the state's functions (by changing people's awareness, attitudes, minds), and on behalf of that to the institutional architecture. Take participatory budgeting as an example. Surely, a single agency in a single municipality could easily realize some participatory budgeting. (At least, technically and organizationally they could easily; whether they are allowed in regards of the surrounding regulatory framework, i.e., the institutional environment, is exactly again a different question). But a single agency in a single municipality introducing participatory budgeting would make no huge difference at state level, not even at municipality level. Only a systematic, state-level participatory budgeting initiative would also lead to a change in the institutional architecture.

These considerations also set limits to the notion of *disruptive technology*. As soon as institutional architecture is critically involved, technology itself cannot be disruptive. A technology can disrupt a market, but only in the boundaries of the established "formal rules of the game" [74]. A technology can never disrupt a state or a society as a whole, as long as the respective state or society is *non-dysfunctional*.

4.6 Data Governance Architecture in Practice

The Case of the European GDPR Regulation Understanding the *interoperability* dimension and the *provisioning* dimension as described in Sects. 4.3 and 4.4 is a key to architecture of digital government ecosystems. For example, the dimensions are also reflected in the European General Data Protection Regulation (GDPR) [24]. The GDPR introduces the notions of *data controller* and *data processor* as follows:

(i) "A data controller is a key decision makers [sic]. They have the overall say and control over the reason and purposes behind data collection and over the means and method of data processing."[6]
(ii) "If two or more controllers have the control over purposes and processes, then they are joint controllers. However, this doesn't [sic] apply if they are using the same data for different purposes."[6]
(iii) "A data processor will act on behalf of the controller. They only operate via instructions from the controller."[6]
(iv) "Individual users can make claims for compensation and damages against both [-a] processors and [-b] controllers."[6]

[6] https://www.gdpreu.org/the-regulation/key-concepts/data-controllers-and-processors/

In our context, the data controller (i) is an entity (authority, organization, company etc.) that collects/holds/uses data. (i) also assigns the accountability with respect to the *minimality principle* to the data controller, where the accountability emerges from (iv-b). (ii) is about the interoperability dimension; it clarifies the accountability of entities that exchange data; however, only for the case that the data are used for the same purpose. (iii) is about the *provisioning* dimension – we would call a data processor simply an IT service provider; and again, (iv-a) clarifies the accountability in the *provisioning* dimension.

From the GDPR example we also learn the following. It is important to understand that the *interoperability* dimension and the *provisioning* are crucial, in general. But then it is also crucial to understand their details when it comes to concrete regulatory frameworks.

Centralization vs. De-Centralization Often, in digital government system design efforts, it comes to a discussion of *centralization* vs. *decentralization*, e.g., in the implementation of data exchange solutions. In such discussions, arguments are often misleading, e.g, because they mix technical with organizational arguments at levels that do not fit, or neglect complex relationships between the technological design and institutional design of large-scale systems. A typical argument (that we heard occasionally) might be for example:

> "*We cannot use an ESB (Enterprise Service Bus) implementation, because citizens cannot trust a centralized government.*" (3)

In a practical project, we cannot ignore a statement such as (3) either, as it surely expresses an important concern.

The point in digital government ecosystem architecture is to systematically decouple considerations from each other along the described *interoperability* and *provisioning* dimension, always against the background of a well-understood data governance architecture. Meaning: when we discuss *centralization*, we first need to clarify and create awareness of the context of the discussion (the same with decentralization). Are we discussing a centralization in the *institutional architecture* or are we discussing a centralization in the *digital government architecture*, compare again with Fig. 1? If we discuss a centralization in the digital government architecture, are we discussing it with respect to the *interoperability* dimension or with respect to the *provisioning* dimension? For example, a centralization of entities in the primary institutional design is completely different issue from nominating an organization as a *data steward* in the *interoperability* dimension of a data governance architecture. And this is again different from establishing an organization in the *provisioning* dimension, for example, an organization that hosts a message-oriented middleware component or that acts, as yet another example, as the certification authority (CA) of a public key infrastructure (PKI).

5 Established and Emerging e-Government Technologies

5.1 The X-Road Data Exchange Platform

X-Road [3, 33, 34, 73, 4, 35, 55, 61][7,8,9,10] is the data exchange platform of the Estonian digital government ecosystem. X-Road is the only data exchange platform that is mentioned in the UN e-Government Survey 2020 [67]: "The data exchange platform in Estonia (X-Road) is administrated centrally to interconnect government information systems and databases and allow government authorities and citizens to securely send and receive information over the Internet within the limits of their authority."

The Estonian regulation on X-Road [59] defines: "1) the data exchange layer of information systems (hereinafter X-Road) is a technical infrastructure and instance between the members of X-Road, which enables secure online data exchange, ensuring evidential value".

X-Road is a peer-to-peer data exchange system teaming together

- a PKI (public key infrastructure),
- sophisticated software components for secure data exchange,
- a nomenclature of metadata items associated with each message along the core representation language and structure of messages,
- systematic (regulated [59]) organizational measures.

The main technical component of X-Road is the *security server*. An instance of the security server is installed by each authority and organization participating in X-Road (called X-Road members), i.e., there are many security servers running that together realize the secure data exchange in a decentralized manner. The security servers encrypt and decrypt messages, check the identity of other servers and their access rights and preserve a log of messages. Each member registers its e-services in a centrally administered directory. Each member grants access to its e-services itself via the access right management of the security server, i.e., access management remains with the member, determining which other members are allowed to access which of its services and data assets.

X-Road Usage Patterns The official statistics page of X-Road[11] indicates, as of January 2021, that there are almost 3000 different services on X-Road, altogether answering approx. 162 million queries per month. Only approx. three percent of the X-Road queries were initiated by private persons for their own informational needs, whereas the absolute majority – 97% – were initiated by a small category of very specific businesses and service providers as well as a

[7] X-tee in Estonian; in English: originally pronounced as 'crossroad', nowadays pronounced as 'x road'
[8] https://x-road.global/
[9] https://www.niis.org/
[10] https://x-road.global/
[11] https://www.x-tee.ee/factsheets/EE/

State Authority	Queries per month
Employments Registry	≈ 40 million
System for Drug Prescriptions	≈ 10 million
e-Health System	≈ 10 million
Population Registry	≈ 10 million
Medical Insurance System	≈ 10 million

Table 3. Top five Estonian data providers (X-Road) as of January 2021.

large number of the several state's organizations. To be more concrete, the five top data providers for queries during the last month (January 2021) are listed in Table 3. An important context information for these numbers is the population size of Estonia with approx. 1.3 million people.

In our understanding, almost all the queries to the Employments Registry and the Population Registry are expected to be made for the functioning of e-administration, i.e., intra-organizational data exchange. However, the rest of the top queries (drug prescriptions and health information) are mostly generated by pharmacies selling prescription drugs and doctors writing, storing and using personal medical data and prescribing drugs. Almost all the drug prescriptions in Estonia are electronic and use the X-Road for data exchange. Soon after the introduction of the system in 2010, it has been one of the most frequent sources for X-Road queries.

Looking at the more detailed query statistics along with the X-Road visualization tool[12], we see that the next most active groups of X-Road users after the ones listed in Table 3 are:

- Bailiffs: mainly querying the information from the tax and customs board.
- The Police: mainly querying information from the police databases and car insurances.
- Ridango (a private company managing the sales of most of the public transport e-tickets): querying information about student status, age and similar information directly influencing the price of tickets.

Is X-Road De-Centralized? The technical basis of X-Road is decentralized. In the actual data exchange, i.e., in sending messages, there is no middleware component involved, as we would find, e.g., in ESB (enterprise service bus) technology, see [16]. In the actual data exchange, there is no man-in-the-middle involved, as we know it, e.g., from the value-add networks (VAN) back in the days of EDI (electronic data exchange), again see [16]. The messages are sent directly between members; but sending of messages is streamlined by the joint protocol of X-Road, which is enforced through the obligatory usage of the security server (either the available implementation, or an own implementation that adheres to the X-Road security server specification). This does not mean, that there is no *centralization* at all in X-Road. First, there is a state-managed central

[12] https://logs.x-tee.ee/visualizer/EE/

organization and the certification authority (CA) for establishing the PKI (public key infrastructure) [65, 26]. Also, the information systems of members (that are accessed then via the e-services of the members) have to be published and confirmed by a registry maintained by the central authority, see the Estonian "X-Road regulation"[59]. This registration process aims to enforce the *minimality principle*, the *data quality* principles and the *once-only* principle, compare with Sects. 4.1. Enforcement of these principles is carried out by different state authorities who have been assigned the task of auditing these aspects of all the information systems registered as users of the X-Road system.

All of this concerns the basic data exchange mechanism provided by X-Road, i.e., the X-Road platform. It does not mean that X-Road prevents *centralized services*. Centralized services can be implemented on top of X-Road. The Estonian *Document Exchange Center* (DEC) resp. Dokumendivahetuskeskuse (DVK) [21, 57, 18], was a perfect example for this – interestingly, the document exchange center has been deprecated and supersesed by a *de-centralized* document exchange protocol, i.e., DHX (Dokumendivahetusprotokoll)[13]. As another example for adding a centralized service, it is interesting to look at the concept of X-Rooms, which is described in the vision document of Estonia's Government CIO Office on the next generation of digital government architecture [70]. An X-Room is a publish-subscribe service, a standard pattern in message-oriented middleware. If it is just a recommended architectural pattern for realizing e-services, it is not necessarily about adding a centralized service; whereas, if it comes with provisioning, it leads to a centralized service. Adding message-oriented middleware components to a decentralized IT system architecture is standard in enterprise computing. When adding a centralized component, this amounts to adjusting the data governance architecture along the *provisioning* dimension as described in Sect. 4.4.

Is X-Road based on Blockchain? X-Road is not based on blockchain. The fact that the X-Road security server might exploit cryptographic data structures and algorithms that are also used by blockchain technology [51] (such as Merkle trees [48] for implementing audit logs) does not make it a blockchain. A blockchain – as introduced with the cryptocurrency Bitcoin by Satoshi Nakamoto in 2019 [50] – is a peer-to-peer network that implements a distributed, replicated database that achieves consensus via an entirely de-centralized consensus protocol [5, 9]. X-Road makes no efforts to achieve consensus, except for authentication, despite there is no centralized ledger. Although X-Road is not based on blockchain technology, it has been sometimes perceived as such by the media. Therefore, in 2018, NIIS (the official product owner of X-Road) launched an official statement[14] that "there is no blockchain technology in the X-Road". What is true however (and what might have contributed to the fact that X-Road has been perceived as blockchain-enabled or even as blockchain) is the fact, that

[13] https://www.ria.ee/dhx/EN.html

[14] https://www.niis.org/blog/2018/4/26/there-is-no-blockchain-technology-in-the-x-road

many of the Estonian state registries are secured by a so-called KSI blockchain (keyless signature infrastructure) that we will describe further in due course in Sect. 5.3.

X-Road Federation In 2014, Finland and Estonia decided to cooperate tightly in developing further their digital government ecosystems. The Nordic Institute for Interoperability Solutions NIIS[15] was founded as a joint agency of Finland and Estonia and was made the official product owner of the X-Road code base [27, 60]. In the sequel, X-Road was deployed as a data exchange platform also in Finland and joint efforts were started to realize cross-border, federated digital government services. For a discussion of challenges of and approaches to federation of digital government ecosystem in the context of the case Finland-Estonia, see [27].

Impact of X-Road The data exchange platform X-Road is mentioned in the UN e-Government Surveys 2016, 2018 and 2020. The UN e-Government Survey 2018 uses X-Road to explain the concept of what they call "Government as an API" [66], as follows:

"*Estonia created X-Road, an application network for exchanging data among agency systems so that all government services are effectively available in one spot. In addition to offering querying mechanisms across multiple databases and supporting the secure exchange of documents, X-Road seamlessly integrates different government portals and applications.*

The private sector can also connect with X-Road to make queries and benefit from access to a secure data exchange layer.

X-Road has made it possible to bring 99 per cent of public services online. On average, 500 million queries per year are made annually using X-Road. Indeed, its use has been estimated to save as many as 800 years of working time. The solution has been equally successful in its roll-out to Finland, Azerbaijan, Namibia, as well as the Faroe Islands. Furthermore, cross-border digital data exchanges have been set up between Estonia and Finland, making X-Road the first cross-border data exchange platform. "([66], Box 8.2. Government as an API, p. 184)

It is important to note, however, that X-Road is used primarily as a tool of e-administration: it is not meant for nor is it used for directly providing e-services or increasing citizen participation. The "public services" mentioned in the last quote should be understood as either inter-organizational data exchange services or information services for citizens, the software of which uses X-Road for obtaining necessary data from other organizations.

Additionally, dozens of countries have used X-Road to implement digital government data exchange[16].

[15] https://www.niis.org/
[16] https://x-road.global/xroad-world-map

5.2 Other Data Exchange Platforms

Cybernetica UXP Cybernetica[17] is a spin-off from the Estonian research institute Küberneetika[18] (1969-2006, since 2007: Department of Cybernetics and Department of Software Science of Tallinn University of Technology). Cybernetica designed and implemented crucial parts of the first versions of X-Road in 2001, including architecture, protocols, security solutions, security server. Nowadays, Cybernetica offers its own data exchange platform UXP as a product[19], which is partly based on the same prototype as X-Road version 6.

NLX NLX[20] is a data exchange platform that is implemented on behalf of an initiative of municipalities in the Netherlands. "NLX is an open source peer-to-peer system facilitating federated authentication, secure connecting and protocolling in a large-scale, dynamic API ecosystem with many organizations."[21]. The system architecture of NLX is oriented towards X-Road, see Sect. 5.1.

GAIA-X In September 2020, GAIA-X has been founded as a non-profit organization by eleven companies from Germany plus eleven companies from France under the aegis of the German Federal Ministry for Economic Affairs and Energy (BMWi). According to GAIA-X, it wants "to create the next generation of data infrastructure for Europe, its states, its companies and its citizens."[22]. GAIA-X targets federation; consequentially, *semantic representation* is among its architectural principles [22]. Semantic interoperability is the key to cross-border digital government solutions. A similar institutional design can be found in the digital government ecosystems of several countries (several states of a federal country). Where, e.g., the format of data assets of municipalities can be centrally standardized respectively prescribed, this is not any more possible in federated scenarios. Here, the differences in data format are a major obstacle to successful digital transformation. A key success is in systematic efforts in semantic description and semantic mapping of data assets.

5.3 Auxiliary Technologies for Digital Government

Keyless Signature Infrastructure Document timestamping solutions are mission-critical in many organizational contexts. Organizations want to have tamper-proof and provable document logs not only in the communications with other organizations; they also want to be safe against failure (accidental or intentional) of their own members/employees. Equally, the state wants to trust the

[17] https://cyber.ee/
[18] https://cyber.ee/
[19] https://cyber.ee/products/secure-data-exchange/
[20] https://nlx.io/
[21] https://docs.nlx.io/understanding-the-basics/introduction/
[22] https://www.data-infrastructure.eu/

operations of its authorities and, again, the authorities want to trust the operations of their employees. Since 2007, Guardtime offers a document time stamping solution as a service, i.e., the so called KSI blockchain (keyless signature infrastructure blockchain). In the Estonian digital government ecosytem, the solution is successfully used to secure the healthcare registry, the property registry, the succession registry, the digital court system and the state gazette [44].

The KSI blockchain achieves a practical implementation of an idea that goes back to Stornetta et al. in 1993 [6], i.e., it stores timestamped document hashes in a Merkle tree [48] and publishes the root hash of the tree *periodically* (e.g., once a month) in a newspaper (e.g., in the Financial Times, among others, in case of the Guardtime solution) [10, 8], see also [9].

Solid Solid[23] (Social Linked Data) is a technology originally designed and advocated by Tim Berners Lee. Given the frustration over the state of data sovereignty in the World Wide Web – with players such as Facebook and Alphabet and data scandals such Cambridge-Analytica, Solid targets to free data from applications: to reconnect people to their data, so to speak. The key concept of Solid is the *pod* (personal online data store). A user stores his or her data in a *pod* and grants applications access to these. The user decides where the *pod* is stored and who can access it. In [11], Solid has been used in a pilot study of local and regional governments of Flanders (one of the federated states of Belgium) to "empower citizens in reusing their personal information online in different contexts such as public services, banking, health insurance, and telecom providers."[11]. As such, this project provides a rigorous consent management approach. Therefore, in the context of digital government ecosystems, the limits to such approach are in the limits of consent management *itself*, as described in Sect. 4.1.

6 Discussion

6.1 On Step-by-Step Emergence of Digital Government Solutions

Despite the formulation of all the digital government strategies, visions, agendas, and declarations, today's digital governments are rather *arbitrarily emerging* instead of *systematically evolving*. See Fig. 2. A first digital solution is built. During the design of the solution, the project learns, step-by-step, about the limitations imposed by the underlying institutional architecture. The resulting data governance architecture has the scope of the solution, it is a sub-architecture encompassing only those organizations involved in the solution. A lot of effort is usually invested into provisioning decisions, with the many different stakeholders developing ad-hoc opinions on-the-fly. Slight adjustments are made to the institutional architecture on behalf of provisioning, that seem to have little impact, but actually embody new unpredictable constraints and limitations on

[23] https://solidproject.org/

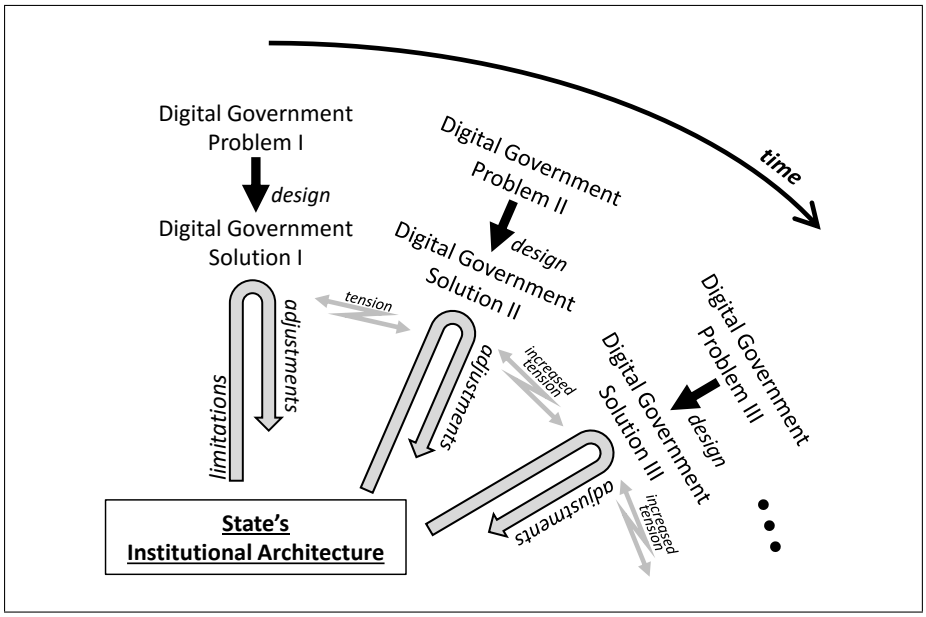

Fig. 2. Step-by-step emergence of digital government solutions in a digital government ecosystem.

future digital government solutions. Next, a further solution project is started for a next digital government problem (problem/solution II in Fig. 2). All problems in shaping the data governance architecture are approached from scratch. Overall, in this process, there is no learning curve. In general, the stakeholders in the second project are different from those in the first; and there is no knowledge transfer from the first to the second project, as there is no systematic *state-level* knowledge management process. This means that the project needs more efforts (is more costly) as it needed to be. A more severe problem is that the design decision in the new project usually follow different design rationales leading to different kinds of decision. The latter creates tension with in between the projects; on the one hand, indirectly via the adjustments to the institutional architecture and, one the other hand, directly – if the projects continue to run in parallel (in future maintenance, follow-up projects etc.). The more digital solutions emerge, the higher rise the tension and, actually, the project costs of new projects.

It can be hoped, that the described frictional losses can be mitigated with a deliberate state-level approach to systematically manage the emergence of the single digital government solutions, this way turning this arbitrary process into a systematic evolvement of digital government. We are convinced, that the elaboration of an architectural framework (such as ours in Sect. 4.2) would be a crucial step towards such systematically evolving digital government.

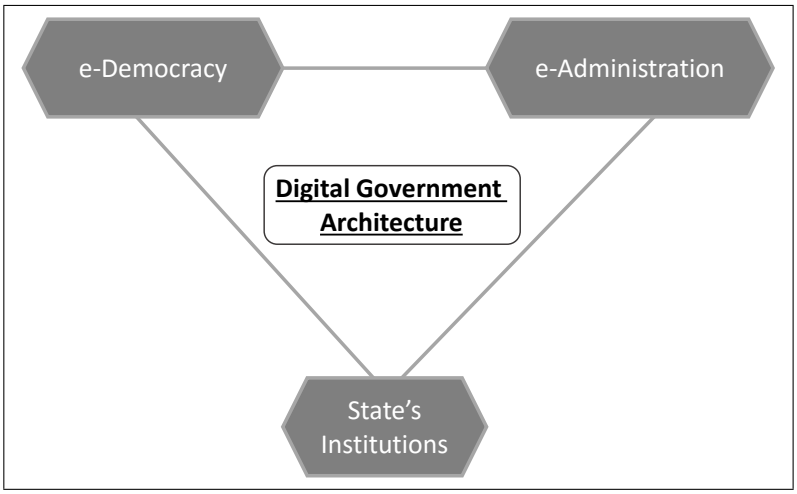

Fig. 3. Digital government architecture in the tension between e-democracy and e-administration.

6.2 On e-Democracy and e-Administration

The field of digital government has always been dominated by two major, well distinguishable strands that we have called *democratic* and *technocratic* narrative in [19]. In the technocratic strand, digital government is about increasing the *efficiency* and *effectiveness* of e-services offered to the *citizen as a customer*. In the democratic strand, digital government is about strengthening our democracies embracing the *citizen as a citizen* with enabling participation and fostering transparency. However, we should not assume that these narratives are a perfect reflection of the underlying forces and motivation driving the development of digital government. Indeed, a strong motivational force is automation of work – similar to industrial automation – even if no e-services are offered or participation encouraged. We will start using the notion of *e-administration*, defined as the automation aspect of digital government. Now, e-administration can be opposed to *e-democracy* with typical digital initiatives such as open government data[24], e-consultation, participatory budgeting, and i-voting [39], see Fig. 3.

We have to understand the evolutionary forces behind the development of digital government. Following Bruno Latour, we can look at the modern government authorities as a network of people and automated software systems (e-administration), with the e-democracy component enabling feedback. Since administrations have a tendency to grow [56], but the financial and personnel limits are severe, the growth is naturally channeled into the development of more and more software systems, i.e., a more elaborate e-administration. That is, the motivation behind e-administration is not so much simplifying the work of people working in the organization, as it is growing the power of the organization

[24] https://opengovdata.org/

via automation. We speculate that it is easier to channel regularly necessary change into e-administration than it is to change the traditional human part of the organization. If so, then a more automated digital government could be more flexible while the pace of organizational change would slow down. This hypothesis needs to be confirmed yet.

As the e-administration component grows, so does the need for data: first, automated systems manage to handle practically unlimited amounts of data, second, machine learning and A.I. create a possibility to predict risks in various spheres and thus to both optimize organizational processes and to take a more proactive stance. As examples, consider public transport and road network optimizations, distribution of firefighting and police resources, selecting tax audit targets etc. Since more power needs more feedback for stability, we might expect the e-democracy component to be driven by the growth of e-administration.

7 Conclusion

Currently, the level of digital government implementation, i.e., the pervasiveness of digital government, is discussed prominently in terms of data, i.e., as proceeding form ICT-driven to data-centric digital government. This comes as no surprise, given the substantial developments in the data technology sector, with Big data and data science. Indeed, there are still vast opportunities to exploit data at state level to better the government's effectiveness, accountability and inclusiveness. At the same time, the risk that citizens' data are misused for citizens' surveillance and control will never vanish. The challenge is in getting the data governance structure right. And this challenge needs to be understood early in all digital government design issues and, therefore, needs to be reflected in each approach to digital government architecture.

Technologies come and go. Emerging technologies drive change. Emerging ICT technologies drive digital transformation. In three decades now, the field of digital government has always shown a particularly optimistic approach to be determined by emerging technologies. Digital government has always been ICT-driven. But what is the *function* of digital government? It should be more than making public administration more efficient and effective. It should be in *connecting* governments with citizens. And what is the *form* of digital government? Government has an institutional design. This institutional design gains primacy in the architecture of digital government ecosystems. We argue that the architecture of any digital government ecosystem can be identified, essentially, with its data governance architecture, which links data assets with accountable organizations, supporting a range of data governance principles. We are convinced, that such viewpoint not only helps to analyze existing digital government ecosystems, solutions and technologies alike; but is also a key to shaping the next generation of digital government.

References

1. Al Gore: Access America: Reengineering Through Information Technology – Report of the National Performance Review and the Government Information Technology Services Board. Vice President of the United States (1997)
2. Andersen, R.: The panopticon is already here. The Atlantic **September** (2020)
3. Ansper, A.: E-State From a Data Security Perspective. Tallinn University of Technology, Faculty of Systems Engineering, Department of Automation, Tallinn (2001)
4. Ansper, A., Buldas, A., Freudenthal, M., Willemson, J.: High-performance qualified digital signatures for X-Road. In: Nielson, H.R., Gollmann, D. (eds.) Proceedings of NordSec 2013 – the 18th Nordic Conference on Secure IT Systems. Lecture Notes in Computer Science, vol. 8208, pp. 123–138. Springer (2013)
5. Antonopoulos, A.M.: Mastering Bitcoin: Programming the Open Blockchain. O'Reilly (2017)
6. Bayer, D., Haber, S., Stornetta, W.: Improving the efficiency and reliability of digital time-stamping. In: Capocelli, R., De Santis, A., Vaccaro, U. (eds.) Sequences II, pp. 329–334. Springer (1993)
7. Bharosa, N., Lips, S., Draheim, D.: Making e-government work: Learning from the Netherlands and Estonia. In: Proceedings of ePart 2020 – the 12th IFIP WG 8.5 International Conference on Electronic Participation. pp. 41–53. LNCS 12220, Springer (2020)
8. Buldas, A., Kroonmaa, A., Laanoja, R.: Keyless signatures infrastructure: How to build global distributed hash-trees. In: Proceedings of NordSec'2013 – the 18th Nordic Conference on Secure IT Systems. LNCS 8208, Springer (2013)
9. Buldas, A., Draheim, D., Nagumo, T., Vedeshin, A.: Blockchain technology: Intrinsic technological and socio-economic barriers. In: Proceedings of FDSE'2020 – the 7th International Conference on Future Data and Security Engineering. LNCS 12466, Springer (2020)
10. Buldas, A., Saarepera, M.: Document Verification with Distributed Calendar Infrastructure. US Patent Application Publication No.: US 2013/0276058 A1 (2013)
11. Buyle, R., Taelman, R., Mostaert, K., Joris, G., Mannens, E., Verborgh, R., Berners-Lee, T.: Streamlining governmental processes by putting citizens in control of their personal data. In: Proceedings of EGOSE'2019 - the 6xt International Conference on Electronic Governance and Open Society – Challenges in Eurasia. pp. 346–359. CCIS 1135, Springer (2020)
12. Coase, R.H.: The problem of social cost. Economia **November**, 386–405 (1937)
13. Coase, R.H.: The nature of the firm. Law & Economics **3**, 1–44 (1960)
14. Di Maio, A.: Moving Toward Data-Centric Government. Gartner Group Report G00248186. Gartner (2014)
15. Draheim, D.: Business Process Technology – A Unified View on Business Processes, Workflows and Enterprise Applications. Springer, Berlin Heidelberg (2010)
16. Draheim, D.: The service-oriented metaphor deciphered. Journal of Computing Science and Engineering **4**(4), 253–275 (2010)
17. Draheim, D.: Smart business process management. In: 2011 BPM and Workflow Handbook, Digital Edition: Social BPM – Work, Planning and Collaboration under the Influence of Social Technology, pp. 207–223. Workflow Management Coalition (2012)
18. Draheim, D., Koosapoeg, K., Lauk, M., Pappel, I., Pappel, I., Tepandi, J.: The design of the Estonian governmental document exchange classification framework. In: Kő, A., Francesconi, E. (eds.) Electronic Government and the Information Systems Perspective. pp. 33–47. Springer (2016)

19. Draheim, D., McBride, K., Misnikov, Y., Hartleb, F., Lauk, M., Lemke, F., Nagumo, T., Pappel, I.: On the narratives and background narratives of e-Government. In: Proceedings of HICSS'2020 – the 53rd Hawaii International Conference on System Sciences. pp. 2114–2122. AIS (2020)
20. Draheim, D., Nathschläger, C.: A context-oriented synchronization approach. In: Proceedings of PersDB 200 – the 2nd International Workshop in Personalized Access, Profile Management, and Context Awareness: in Conjunction with the 34th VLDB Conference. pp. 20–27. ACM (2008)
21. Eesti Äriarhiivi: Requirements for electronic document management systems' functionality (Nõuded elektrooniliste dokumendihaldussstemide funktsionaalsusele). Eesti Äriarhiivi (2002)
22. Eggers, G., et al.: GAIA-X: Technical Architecture. Federal Ministry for Economic Affairs and Energy (BMWi) Public Relations Division, Berlin (2020)
23. European Commission: EU eGovernment Action Plan 2016-2020: Accelerating the Digital Transformation of Government – COM(2016) 179 final. European Commission (2016)
24. European Commission: Regulation 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 (General Data Protection Regulation). European Commission (2016)
25. European Commission: New European Interoperability Framework: Promoting Seamless Services and Data Flows for European Public Administrations. Publications Office of the European Union, Luxembourg (2017)
26. Felt, S., Pappel, I., Pappel, I.: An overview of digital signing and the influencing factors in Estonian local governments. In: Tran Khanh Dang, Roland Wagner et al. (eds.) Proceedings of FDSE'2016 – the 3rd International Conference on Future Data and Security Engineering. pp. 371–384. LNCS 10018, Springer (2016)
27. Freudenthal, M., Willemson, J.: Challenges of federating national data access infrastructures. In: Proceedings of SecITC 2017 – the 12th International Conference on Security for Information Technology and Communications. pp. 104–114. LNCS 10543, Springer (2017)
28. Holsapple, C., Whinston, A.B. (eds.): Decision Support Systems: Theory and Application. Springer (1995)
29. Huntington, S.P.: The Clash of Civilizations and the Remaking of World Order. Simon & Schuster (1996)
30. ISO/IEC JTC 1/SC 7: ISO/IEC 25012:2008: Software engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Data quality model. International Organization for Standardization (2008)
31. Janssen, M., Kuk, G.: The challenges and limits of big data algorithms in technocratic governance. Government Information Quaterly **33**(3), 371–377 (2016)
32. Jensen, M.C., Meckling, W.H.: Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics **3**(4), 305–360 (1976)
33. Kalja, A.: The X-Road: a key interoperability component within the state information system. In: Information technology in public administration of Estonia – yearbook 2007. pp. 19–20. Ministry of Economic Affairs and Communications Estonia (2008)
34. Kalja, A.: The first ten years of X-Road. In: Kastehein, K. (ed.) Information technology in public administration of Estonia – yearbook 2011/2012. pp. 78–80. Ministry of Economic Affairs and Communications Estonia (2012)

35. Kalja, A., Robal, T., Vallner, U.: New generations of Estonian eGovernment components. In: Proceedings of PICMET'2015 – the 15th Portland International Conference on Management of Engineering and Technology. pp. 625–631. IEEE (2015)
36. Kalvet, T., Toots, M., Krimmer, R.: Contributing to a digital single market for Europe: barriers and drivers of an EU-wide once-only principle. In: Proceedings of DG.O'2018 – the 19th Annual International Conference on Digital Government Research. pp. 45:1–45:8. ACM (2018)
37. Koppenjan, J., Groenewegen, J.: Institutional design for complex technological systems. International Journal of Technology, Policy and Management **5**(3), 240–257 (2005)
38. Krimmer, R., Kalvet, T., Toots, M., Cepilovs, A., Tambouris, E.: Exploring and demonstrating the Once-Only Principle: A European perspective. In: Proceedings of DG.O'2017 – the 18th Annual International Conference on Digital Government Research. pp. 546–551. ACM (2017)
39. Krimmer, R., Volkamer, M., Duenas-Cid, D.: e-Voting – an overview of the development in the past 15 years and current discussions. In: Proceedings of E-Vote-ID 2019 – the 4th International Joint Conference on Electronic Voting. pp. 1–13. LNCS 11759, Springer (2019)
40. Latour, B.: Reassembling the Social – An Introduction to Actor-Network-Theory. Oxford University Press (2005)
41. Layne, K., Lee, J.: Developing fully functional e-government: A four stage model. Government Information Quarterly **18**(2), 122–136 (2001)
42. Lee, J.: 10year retrospect on stage models of e-government: A qualitative meta-synthesis. Government Information Quarterly **27**(3), 220–230 (2010)
43. Lemke, F., Taveter, K., Erlenheim, R., Pappel, I., Draheim, D., Janssen, M.: Stage models for moving from e-government to smart government. In: Proceedings of EGOSE'2019 – the 6xt International Conference on Electronic Governance and Open Society - Challenges in Eurasia. Communications in Computer and Information Science, vol. 947. Springer (2019)
44. Martinson, P.: Estonia – the Digital Republic Secured by Blockchain. PricewaterhouseCoopers (2019)
45. Matheus, R., Janssen, M., Maheshwari, D.: Data science empowering the public: Data-driven dashboards for transparent and accountable decision-making in smart cities. Government Information Quaterly **37**(3), 1–9 (2020)
46. McBride, K., Draheim, D.: On complex adaptive systems and electronic government – a proposed theoretical approach for electronic government studies. The Electronic Journal of e-Government **18**(1), 43–53 (2020)
47. McNeely, C.L., on Hahm, J.: The big (data) bang: policy, prospects, and challenges. Review of Policy Research **4**, 304–310 (2014)
48. Merkle, R.: Protocols for public key cryptosystems. In: Proceedings of S&P'1980 – the 1st IEEE Symposium on Security and Privacy. pp. 122–122 (1980)
49. Mozur, P.: Inside China's dystopian dreams: A.I., shame and lots of cameras. The New York Times **July 8** (2019)
50. Nakamoto, S.: Bitcoin: A Peer-to-Peer Electronic Cash System (2008), available at: https://bitcoin.org/bitcoin.pdf
51. Narayanan, A., Clark, J.: Bitcoins academic pedigree. Communications of the ACM **60**(12), 36–45 (2017)
52. Nee, V., Ingram, P.: Embeddedness and beyond. In: Brinton, M., Nee, V. (eds.) The New Institutionalism in Sociology, pp. 2–45. Russell Sage Foundations (1997)
53. Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press (1995)

54. North, D.: Institutions. Journal of Economic Perspectives **5**, 97–112 (1991)
55. Paide, K., Pappel, I., Vainsalu, H., Draheim, D.: On the systematic exploitation of the Estonian data exchange layer X-Road for strengthening public private partnerships. In: Proceedings of ICEGOV2018 – the 11th International Conference on Theory and Practice of Electronic Governance. pp. 34–41. ACM (2018)
56. Parkinson, C.N.: Parkinson's law. The Economist **Nov** (1955)
57. PricewaterhouseCoopers: Public Services Uniform Document Management – Final Report (Lõpparuanne Avalike teenuste ühtne portfellijuhtimine). Pricewaterhouse-Coopers (2014)
58. Project Management Institute: Organization Charts and Position Descriptions. In: A Guide to the Project Management Body of Knowledge (PMBOK Guide), 5th ed., pp. 260–261. Project Management Institute (2013)
59. Regulation no. 105: The Data Exchange layer of Information Systems. Government of the Republic of Estonia (2016)
60. Robles, G., Gamalielsson, J., Lundell, B.: Setting up government 3.0 solutions based on open source software: The case of X-Road. In: Proceedings of EGOV'2019 – the 18th IFIP WG 8.5 International ConferenceElectronic Government. pp. 69–81. LNCS 11685, Springer (2019)
61. Saputro, R., Pappel, I., Vainsalu, H., Lips, S., Draheim, D.: Prerequisites for the adoption of the X-Road interoperability and data exchange framework: A comparative study. In: Proceedings of ICEDEG 2020 – the 7th International Conference on eDemocracy & eGovernment. pp. 216–222,. IEEE (2020)
62. Scholl, H.J., Klischewski, R.: E-government integration and interoperability: Framing the research agenda. International Journal of Public Administration **30**(8–9), 889–920 (2007)
63. Shen, J., Cooley, V.E., Ma, X., Reeves, P.L., Burt, W.L., Rainey, J.M., Yuan, W.: Data-informed decision making on high-impact strategies: Developing and validating an instrument for principals. The Journal of Experimental Education **80**(1) (2012)
64. Tepandi, J., Lauk, M., Linros, J., Raspel, P., Piho, P., Pappel, I., Draheim, D.: The data quality framework for the Estonian public sector and its evaluation. Transactions on Large-Scale Data- and Knowledge-Centered Systems **35**, 1–26 (2017)
65. Tsap, V., Pappel, I., Draheim, D.: Key success factors in introducing national e-identification systems. In: Tran Khanh Dang, Roland Wagner et al. (eds.) International Conference FDSE 2017 – the 4th International Conference on Future Data and Security Engineering. pp. 455–471. LNCS 10646, Springer (2017)
66. UN Department of Economic and Social Affairs: United Nations E-Government Survey 2018 – Gearing e-Government to Support Transformation Towards Sustainable and Resilient Societies. United Nations, New York (2018)
67. UN Department of Economic and Social Affairs: E-Government Survey 2020 – Digital Government in the Decade of Action for Sustainable Development. United Nations, New York (2020)
68. UN Economic and Social Council: Relating the Principles of Effective Governance for Sustainable Development to Practices and Results – Note by the Secretariat, E/C.16/2019/4. United Nations (2019)
69. United Nations General Assembly: Transforming Our World: the 2030 Agenda for Sustainable Development – Resolution A /RES/70/1. United Nations (2015)
70. Vaher, K.: Next Generation Digital Government Architecture. Republic of Estonia GCIO Office (2020)

71. van Ooijen, C., Ubaldi, B., Welby, B.: A Data-Driven Public Sector: Enabling the Strategic Use of Data for Productive, Inclusive and Trustworthy Governance. OECD Working Papers on Public Governance No. 33. OECD (2019)
72. Walsham, G.: Actor-Network Theory and IS research: Current status and future prospects. In: Lee, A.S., Liebenau, J., DeGross, J.I. (eds.) Information Systems and Qualitative Research, pp. 466–480. Springer (1997)
73. Willemson, J., Ansper, A.: A secure and scalable infrastructure for inter-organizational data exchange and eGovernment applications. In: Proceedings of the Third International Conference on Availability, Reliability and Security 2008. pp. 572–577 (2008)
74. Williamson, O.E.: Transaction cost economics: how it works; where it is headed. De Economist **146**, 23–58 (1998)
75. Williamson, O.E.: The new institutional economics: Taking stock, looking ahead. Journal of Economic Literature **38**(3), 595–613 (2000)
76. Wimmer, M., Traunmüller, R.: Integration - the next challenge in e-government. In: Proceedings of EurAsia-ICT 2002: 1st EurAsian Conference Information and Communication Technology. pp. 213–218. LNCS 2510, Springer (2002)