IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 28, 2020, accepted March 16, 2020, date of publication March 31, 2020, date of current version April 28, 2020.
Digital Object Identifier 10.1109/ACCESS.2020.2984691

Business Rationale for Linked Data at Governments: A Case Study at the Netherlands' Kadaster Data Platform

ERWIN FOLMER[1,2], STANISLAV RONZHIN[1,2,3], JOS VAN HILLEGERSBERG[1], WOUTER BEEK[2,4], AND ROB LEMMENS[3]
[1]Behavioral, Management, and Social Sciences, University of Twente, 7522 NH Enschede, The Netherlands
[2]Kadaster Product and Process Innovation, 7311 KZ Apeldoorn, The Netherlands
[3]Geo-Information Science and Earth Observation, University of Twente, 7514 AE Enschede, The Netherlands
[4]Knowledge Representation and Reasoning Group, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands
Corresponding author: Stanislav Ronzhin (s.ronzhin@utwente.nl)

The authors are thankful to Kadaster and the University of Twente for providing the research opportunity.

ABSTRACT Linked Data is an innovative approach for publishing heterogeneous data sources on the web. As such, it can transcend the traditional confines of separate databases, as well as the confines of separate institutions. At the same time, businesses and governmental organizations alike are trying to cope with ever-increasing quantities of heterogeneous data that must be used across multiple departments, manufacturing locations, and governmental bodies. Linked Data would, therefore, be a great technological solution for today's organizational problems. However, we observe that a serious gap exists between Linked Data research and business research. While Linked Data research is almost exclusively technologically oriented, the business research literature has not devoted much attention to the use of Linked Data solutions yet. In this paper, we seek to bridge this gap, by introducing a real-world use case where Linked Data technologies are applied in large-scale government settings. We argue in detail that Linked Data provides a major contribution to the business vision of a modern governmental institution based on the experience of the Netherlands' Cadastre Land Registry and Mapping Agency (Kadaster), so far, the largest implementation of Linked Data in the governments of the Netherlands.

INDEX TERMS Government, information management, data systems.

I. INTRODUCTION

More than 15 years were needed for the Linked Data (LD) and the Semantic Web (SW) technology to evolve from a mere envision presented in [1] to a mature technology residing in the plateau of production of the Gartner diagram [2].

In Osterwalder's terms [3], the main value propositions of implementing LD (adopted from [4]) are decreased costs and increased flexibility of data integration and management. This leads to improved data quality and gives rise to new services. This can be observed from many inquiries indicating that depending on the scale and scope of an LD project the saving potential in the management and reutilization of data can be noteworthy (e.g. [5]– [7]).

However, despite the plethora of reported use cases, the focus is often on technical decisions leaving behind business aspects. A research gap can be observed from the number of LD-related publications in the business and computer science domains. Table 1 presents the number of publications containing the keywords "Linked Data" or "Semantic Web" found by Scopus in the business and computer science domains. As can be seen from Table 1 the number of publications in the business, management and accounting domain is one order of magnitude smaller than in the computer science domain.

LD as a disruptive technology affects organizations and creates challenges (and opportunities) for business development. These challenges have a complex multifaceted nature touching organizational, social and business aspects. However, it is still difficult to estimate the cost-effectiveness of LD implementations since it requires quantification of non-economic benefits and risks. All of these add uncertainty to the business rationale of organizations when it comes to implementing LD [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Wajahat Ali Khan.

A. PROBLEM STATEMENT

Through the five years of running the Platform Linked Data Netherlands (PLDN, https://pldn.nl), a community project

TABLE 1. The number of publications containing keywords "linked data" or "semantic web" found by SCOPUS in the business and computer science domains

Search request [a]	Domains	Number of publications
(KEY (linked AND data) OR KEY (semantic AND web)) AND (LIMIT-TO (SUBJAREA , "COMP"))	Computer Science	33770
(KEY (linked AND data) OR KEY (semantic AND web)) AND (LIMIT-TO (SUBJAREA , "BUSI"))	Business, Management and Accounting	1262

[a] Search service https://www.scopus.com/search/form.uri

of LD practitioners and researchers from both private and public sectors, we watched many LD implementations and projects both in industry and government, but also in particular in the spatial domain. Ordnance Survey UK was an early adopter, followed by other large implementations in amongst others Belgium, Swiss, and the Netherlands. However, we observed that policymakers were struggling in deciding whether they should define policies on LD or not. This was especially the case in the context of e-government. We interpreted this as a piece of evidence supporting the existence of a knowledge gap in the business rationale behind the implementation of LD. Therefore, we aim to bridge this gap, by providing in-depth knowledge gained from the case study at the Kadaster Data Platform, a major development within Kadaster, the Netherlands' Cadastre Land Registry and Mapping Agency. Our goal is to answer the following research question:

- Up to what extent can the Linked Data technology contribute to the business vision of a governmental institution like Kadaster?

B. RESEARCH APPROACH

This is qualitative research that combines the case study [8] and the action research [9] methodologies. It qualifies best as an exploratory case study research because it explores those situations in which the intervention being evaluated has no clear, single set of outcomes. Analysis of the business case of LD implementation at Kadaster was the basis for the identification of the business reasons to evoke an LD project.

The personal involvement of one of the authors in setting up the business case and in the process leading to its approval gave us an opportunity to analyze and describe the identified business reasons in detail. The uniqueness of this situation is a good argument for a single case study [8], as it enables to perform a deep dive. An alternative would have been to apply a multi-case study approach and analyze implementations in, for instance, UK, Belgium, and the Netherlands; this might be interesting for follow-up research. Authors personal involvement is also one of the key characteristics of action research: collaboration between researcher and organization in order to solve organizational problems. Traditional action research follows a plan, act, observe and reflect phases.

In the planning phase, a set of Business Requirements (BR) was identified based on the analysis of the business vision of Kadaster. After that in the act phase, the platform was developed using the business requirements as guiding principles. In the current phase, the platform is a working infrastructure, therefore, we observe the role of LD technology in supporting identified BRs. The roles of LD presented in this paper were identified as a result of analysis of BRs and LD capabilities supported by our own experience of developing the platform.

The reported case is of interest not only because of the availability of detailed information, but also because it is the largest Linked Data implementation in the Netherlands so far, and one of the largest implementations among governments in the world.

The following sections introduce the main concepts (Section 2) used within the paper and give the organizational context (Section 3). The business vision approved by the Kadaster's board of directors as well as BRs are presented in Section 4 when Section 5 explains how LD support the vision and meet the requirements. Section 6 discusses the outcomes in a broader sense. Conclusions are drawn in Section 7.

II. LINKED DATA AND OTHER CONCEPTS

In this section, we provide an overview of the most important concepts used in the paper. First, we explain the LD technology, the main patterns of LD utilization, value creation and business assets associated with LD. Second, we introduce the notion of a data platform.

A. LINKED DATA IN A NUTSHELL

According to [10], the Web is an elegant publication platform for documents, but it is not possible to search for data at a sub-document level. For example, it is easy to search and retrieve documents about the registration of certain buildings that are regularly published by Kadaster. However, it is not possible to search e.g., for the oldest building among these documents. Even though the year of construction does occur in the aforementioned registrations, the original web does not allow this information to be encoded in such a way that it can be uniquely identified.

The concept of the Semantic Web was envisioned by Tim Berners-Lee [1] to tackle this exact flaw of the original Web. For the above example, this implies that each data attribute that appears in the registration document is individually recognizable, retrievable, and combinable into aggregate statistics. However, the SW must be filled with data that is machine-readable and processable [10]. Therefore, the Linked Data initiative [11] took place promoting the use of semantic standards for representing and publishing information on the Web at the data level.

This can be done by encoding information using the Resource Description Framework (RDF) [12]. This standard is based on mature technologies: the graph data model [13] and the Hypertext Transfer Protocol (HTTP) [14]. The former allows instances and concepts, represented by nodes, to be related to one another by relationships, represented by arcs between the nodes. Through HTTP Universal Resource Identifiers (URIs) these data elements (nodes and arcs) become globally accessible, referenceable [15] and queryable by the means of SPARQL query language [16].

Apart front the declaration of the data model and its example serializations, the RDF standard introduces the RDF Schema (RDFS) vocabulary [17] a set of basic semantic primitives to capture the meaning of the concepts used in data. These primitives are used to construct systems of concepts and relations between them called ontologies or vocabularies (often used interchangeably) [18]. Vocabularies can be complex of heterogeneous granularity and are able to represent an entire knowledge domain.

B. VOCABULARIES

By decoupling the meaning of the data from the schemata the RDF standard enables an ecosystem of reusable vocabularies on the Web. Even though technically speaking, LD can be published without referring to any Web-accessible description of the used vocabulary, queryability and consequently usability of such data would be questionable since no one would able to learn about the meaning of the data. Reference [19] encourages data owners, engineers, and practitioners to publish and use vocabularies on the Web by introducing 5 stars for Linked (Open) Data vocabulary use.

The Linked Open Data Vocabulary (LOV, https://lov.linkeddata.es/) [20] project indexes existing vocabularies and it maintains a discovery portal over more than sixty-three thousand terms from almost seven hundred vocabularies (as for 1 June 2019).

C. LINKED DATA IN USE

Reference [21] reviewed a wide range of LD applications and concluded that LD became a commonly deployed industrial technology. However, as pointed out by [22] not much has been published on Linked Data ecosystems, frameworks, and analytics. Nevertheless, three scenarios of LD utilization can be identified in the industry [23]:

1) LD is used internally to consolidate existing disjoint infrastructures, to overcome the legacy issues derived from them without altering existing systems and workflows.
2) Organizations reuse external data sources for content enrichment either for internal use or for creating reusable information products (e.g. see [5], [6]).
3) LD publishing allows an organization to become part of the Linked Data Cloud (see [24]) and thus participate in the surrounding and maintaining ecosystem.

Linked Data allows an organization to integrate its data from separate tables and hierarchies, to integrated networks as a principle action to support data and knowledge organization [25]– [27]. As such, the primary value proposition of LD lies in its ability to transfer data across contexts, while still preserving its original meaning, thereby generating network links at the data level [28]. This ability gives rise to several innovative business models described by [29]. The series of best practices for publishing data on the web provides actionable cookbooks for LD practitioners [30], [31].

D. LINKED DATA: VALUE CREATION AND BUSINESS ASSETS

Reference [32] describes the value creation process of LD usage patterns. This process starts with raw data, which can be provided in various ways, including tabular (spreadsheet), relational (SQL), tree-shaped (XML), markup (SGML), and binary (PDF) form. In the next step, the source data is translated into RDF and exposed for further consumption. After that, LD is being consumed and processed by an LD application. Finally, the end-user consumes human-readable data via functionally-extending applications and services. This pattern can be found in many Open Government Data projects as illustrated by [4] and [33] addressed to this that besides the value creation process itself, LD creates an environment enabling added value services.

Reference [23] distinguishes six business assets that appear in LD, albeit technically oriented: (1) instance data, (2) metadata, (3) vocabularies, (4) content, (5) services, and (6) technology. Besides, [34] pointed out that LD can be seen as a channel for computer-assisted communication, e.g. between data providers and users, where data and metadata are ''the explicit top of a pyramid of implicit acts of interpretation, observation, and construction'' (also see [35]).

E. DATA PLATFORMS

Data platforms are built to manage an increasing amount of data coming from an increasing number of applications in an increasing number of data formats. A modern open data platform should have several key capabilities (adopted form [36]):

1) Open Data Access: Data must be digestible and consumable via industry-standard protocols. The data platform must be truly independent and future proof so users have confidence that they can get to the data whenever and however they choose.
2) Virtual Data Consolidation: A modern data platform must virtually unite disparate data locations and formats by providing consistent management, operations and navigation of data sets.
3) Metadata & Provenance: Metadata with arbitrary granularity is the foundation for more intelligent control of data, since it allows access to data at different levels, and better usage of the data (see [37]).
4) Lifecycle Data Services: A modern data platform should transparently orchestrate and automate the life cycle, copy management, compliance and governance of data.
5) Data Value Delivery — Any platform's mission is to provide value to its users. Data value is more than just analytics or simple data visualization, rather it is the ability to match information with the user's needs.

In case a data platform also publishes non-open (proprietary, paid, privacy-sensitive) data, then comprehensive data security should be added as a key capability.

One of the new data platforms that implement the above-mentioned capabilities is *data.world* [38]. it uses LD as the main underlying technology,[1] but for end-users, the data and functionality are presented in a user-friendly way without confronting them with LD. Another example is an open data platform of the British Ministry of Housing, Communities & Local Government,[2] which is based on an architecture called PublishMyData [39]. Other data platforms (such as in the province of Groningen, the Netherlands) have been studied, resulting in the conclusions that users should be involved in setting up data platforms [40]. Within the industry, Bloomberg[3] recently launched an LD platform allowing customers to connect their enterprise systems directly to Bloomberg's comprehensive historical data archives.

III. KADASTER: CONTEXT AND BACKGROUND

The Netherlands' Cadastre Land Registry and Mapping Agency – in short Kadaster – collects and registers administrative and spatial data on property and the rights involved. This also goes for ships, aircraft and telecom networks. Doing so, Kadaster protects legal certainty.

A. DATA AND SERVICES

Kadaster publishes many large authoritative datasets including several key registers of the Dutch Government (e.g. Basic Registry of Topography (In Dutch: Basisregistratie Topografie; Dutch acronym: BRT) [41], Basic Registry of Addresses and Buildings (in Dutch: Basisregistratie Adressen en Gebouwen; Dutch acronym: BAG) [42]. Furthermore, Kadaster is also developing and maintaining Publieke Dienstverlening op de Kaart (Dutch acronym: PDOK; www.pdok.nl) containing shared services in a web portal where more than 150 spatial datasets coming from different Dutch government organizations are being published in several formats.

These data include an incredible number of geospatial objects. These objects are spatially and/or conceptually related but are maintained by different data curators. As a result, these datasets are syntactically and structurally disjoint. And currently, it requires non- trivial human labor to use them together. For these reasons, Kadaster started publishing its data assets as Linked Open Data. Together with the already existing data services on PDOK, it has become a complete offering of data supplied via many interfaces e.g. Web Feature Service (WFS), Web Map Service (WMS), Atom Feeds, SPARQL, REST API.

B. LINKED DATA AT KADASTER

Let us imagine that Kadaster registered an object, a building of the Saint Catharine church erected in 1900 in Eindhoven with a certain registration ID. Figure 1 depicts this as a plain text. This information can be decomposed into 3 facts: (1) the object is a church, (2) it has a name, and (3) it was erected in 1900. In Figure 1 these facts are shown as a graph with green rectangles as nodes and arrows as relations between them.

In Figure 1, the blue circles and rectangles represent the same graph but with all the arbitrary wording replaced by standardized notions and their URIs. Notions within a collection share the same namespace and are often abbreviated. For instance, in Figure 1 ''rdf'' is a namespace prefix for the basic RDF vocabulary. If there is a URI to represent a concept (e.g., *bag:AddressableObject*) it is depicted as a circle; literal values are shown as rectangles. This is done to emphasizes that only URIs can be linked.

Standardization of semantic descriptions and use of URIs allow linking data items between data sets. The Saint Catharine church is an outstanding building it appears in many datasets. In Figure 1, a dashed arrow represents a relation (*owl:sameAs*) between two representations of the same church in BAG (blue shapes) and in BRT (yellow shapes). Even though the building is classified differently in these datasets (as an addressable object in BAG and as a church in BRT) by linking them together we can infer additional knowledge e.g., that a church is an addressable object. In this way, previously disconnected data items being linked together via persistent URIs form a knowledge graph [43].

IV. KADASTER DATA PLATFORM

The Kadaster Data Platform (KDP) is the result of the road that Kadaster has taken from 2015 onwards. In this time, Kadaster realized that, in essence, ''data'' is the core of the organization. For this reason, the board of directors requested to set up data strategy and align it with the general ambition of Kadaster to design sustainable and future proof business vision. As a result, Kadaster set up a business case for investment in the new data platform. The platform was expected to support the already existing PDOK shared service and to provide extended infrastructure for future developments.

The KDP was realized and it is available now [44]. The data persistence layer of the platform is comprised of a document store and a triple store. Both storages are synchronized and are kept up to date by an Extract, Transform and Load (ETL) procedure that automatically loads incremental updates of the data assets from existing data infrastructures. The platform provides access to these data via queryable REST APIs (powered by the document store) and a SPARQL endpoint (triple store). On top of this endpoint, the KDP implemented various front-end functionalities allowing tabular browsing and hierarchical browsing of the data, as well as graph-based navigation.

However, regardless the all technologies that were to be used, the original business case was formulated to be technology agnostic and business-driven. It defined a vision on the horizon rather than a set of concrete specifications. In the following subsections, we present this vision (Section 4.1)

[1] See https://meta.data.world/linked-data-on-data-world-23f5cd60ce63

[2] http://opendatacommunities.org/home

[3] https://www.programmableweb.com/news/bloomberg-launches-open-data-and-linked-data-website/brief/2018/09/12

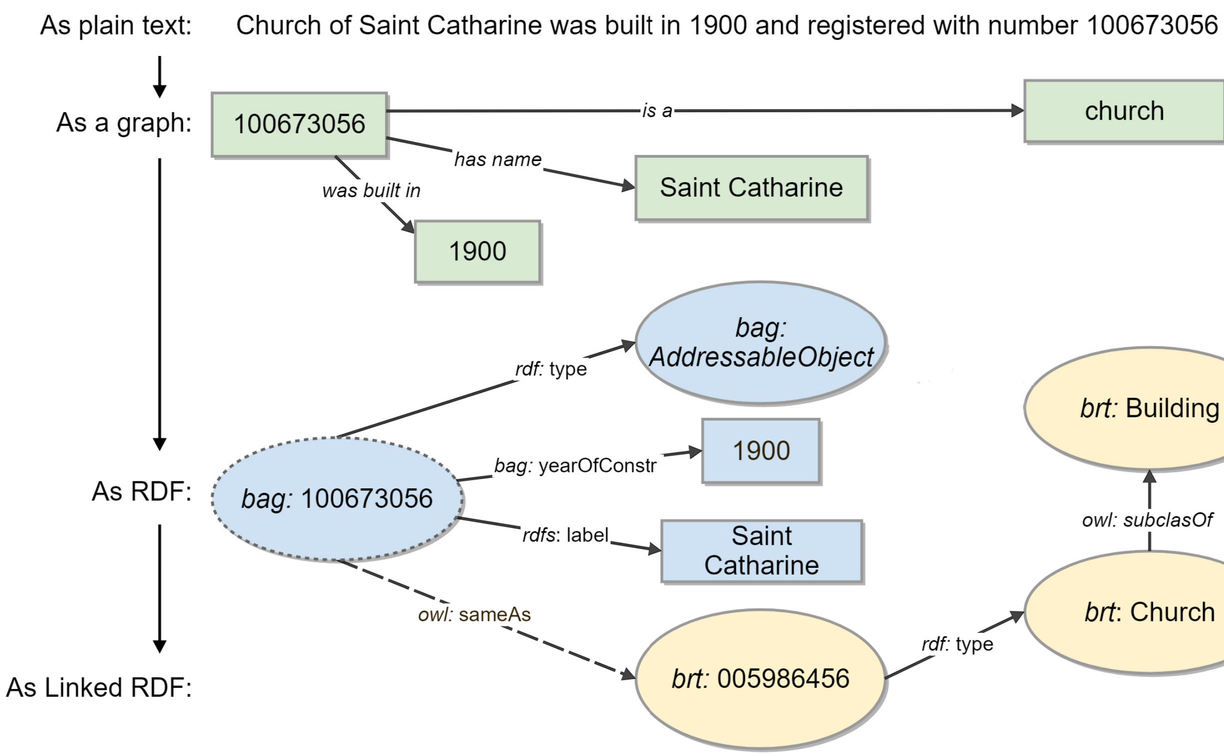

FIGURE 1. **Representation of facts about a registered object, the Saint Catharine church erected in 1900 in Eindhoven using RDF. The plain text given at the top of the figure can be decomposed into 3 facts: (1) the object is a church, (2) it has a name, and (3) it was erected in 1900. These facts represented as a graph under the plain text. In the graph, green rectangles denote nodes and arrows represent relations between nodes. The blue ellipses(concepts) and rectangles (literal values) depicted in the center of the figure represent the same graph but with all the arbitrary wording replaced by standardized notions and their URIs. The URIs are shortened using namespace prefixes. Yellow shapes represent data items from another dataset (BRT) which are linked to blue ones (BAG) forming a part of a knowledge graph. Even though the building is classified differently in these datasets (as an addressable object in BAG and as a church in BRT) by linking them together we can infer additional knowledge e.g., that a church is an addressable object.**

and discuss how the vision can be operationalized via a set of BRs of what to-be the KDP (Section 4.2-4.3).

A. VISION AND AMBITIONS

The goals of the KDP are related to the current five-year vision of Kadaster.[4] This vision consists of four ambitions supported by eleven business requirements. These ambitions are as follows.

Ambition 1 (Certainty and Legitimacy): Since the core mission of Kadaster is to provide certainty and support legitimacy of properties on ground-level and beneath, therefore in all situations, users of the data should be able to trust the data and make a responsible decision based on the data.

Ambition 2 (Spatial Data Provider): The second ambition of Kadaster is to become the spatial platform of the Netherlands, a place to discover and consume spatial data. This ambition can be explained by the long history and trusted brand of Kadaster, next to the five-year experience of hosting the PDOK web portal.

Ambition 3 (Increase Data Value): Recent usage analysis of PDOK has discovered about 90% of hits were related to less than 5% of the data sets [44]. Therefore, more data does not automatically lead to more use. The data itself should have value for the users, therefore, a focus on more valuable data is needed. A realistic view is that the KDP will provide access to integrated services (such as querying) allowing federation from different sources in the background.

From the end-user perspective, the distinction between open and paid (proprietary) data is less interesting than from the suppliers' perspective. The KDP should not be limited to open data and will need the functionality for access management and billing.

Ambition 4: (Use-Case Oriented): Kadaster as a government body should deliver societal and economic benefits. In line with a popular belief that one size fits none, different user groups should be accommodated with different data formats. Therefore, there should be a change from current supply-driven offerings to demand-driven ones.

B. BUSINESS REQUIREMENTS

The ambitions presented in Section 4.1 define the Kadaster vision of the KDP. The following presents BRs that detail the vision.

BR 1 (Metadata as Part of The Data): To prevent misinterpretation of data, users need to be provided with more context about data. The current practice of data portals to

[4]https://www.kadaster.nl/meerjarenbeleidsplan-2018-2022 [DUTCH]

publish metadata relies on standards for dataset and service descriptions (e.g., [45]) containing relevant information e.g., how to access the data. Therefore, before the data can be accessed it needs to be discovered in catalogues containing dataset or service level descriptions. Keeping in mind that descriptions can be ambiguous, and catalogues can contain up to dozens of thousands of datasets (e.g., European Data Portal has 1061275 datasets as of March 2020), the task to find data suitable for reuse in a user application became very difficult.

For the KDP, metadata should be integrated and presented to the user together with the data. Metadata should include not only descriptive information about the dataset but also data element definitions, value lists, provenance information about how this dataset was constructed, how it was transformed, its potential use cases and limitations.

BR 2 (Data at The Source): Current practice often requires copying and transforming Kadaster data to fit the system requirements of data (re)users. This is a common workflow for other governmental organizations using data from the key registers maintained and served by Kadaster. As a result, there is a certain risk that data copies stored in the systems of consumers evolve differently from the original data.

For dynamic datasets (such as key registries), the copied data become outdated immediately after the copying has been finished. Although it is possible to have an update process running on the copied data, in practice most copies have a limited update process if any at all. It needs no explanation that it becomes a risk when copied outdated data is used for emergency services, making formal government decisions, etc.

Another issue, but of less importance, is the additional costs associated with having all those copy databases (including licenses and database administrators).

For the KDP it means that nearly all reasons for copying data into the user's data stores should be eliminated. Most of the user arguments are related to the lack of easy to use services, and the absence of a Service Level Agreement (SLA) defining service performance, downtime, etc. The KDP plans to meet these demands by deploying a constellation of developer-friendly customizable data APIs.

BR 3 (Knowledge Graph): Coined by Google in 2012, the term Knowledge Graph (KG), in a broad sense, refers to a graph-based representation of general world knowledge. By harnessing the SW technology, KGs allow going beyond a keyword search paradigm in information search and retrieval. "Things, not strings" as it was put by Google. Knowledge Graphs are fuel for intelligent systems and agents that would be able to answer complex questions such as: "Can I build a shed in my backyard? And if no, what additional requirements do I need to meet?". KGs are constructed by putting together heterogeneous cross-domain data repositories. Since 2018, Knowledge Graphs have also been included in the Gartner Hype Cycle for Emerging Technologies.

In the new approach, all the dispersed Kadaster datasets would be glued together forming a Kadaster Knowledge Graph (KKG). After that, the graph can be enriched with external open data resources from the Web. However, the integration of multiple data sets into a Knowledge Graph requires alignment of the data models between sets. Even though Kadaster owns all the data and schemas it is still not trivial to perform this, due to the complexity and heterogeneity of meaning.

BR 4 (Linkability): In contrast to the third business requirement (interlinking data within Kadaster), this requirement focuses on providing means for linking external resources to Kadaster resources by external organizations.

Kadaster as a government organization publishing authoritative data with legal weight and defined quality is interested in promoting the use of its data in a wider context. For example, crowdsourcing projects like Wikipedia when publishing information on administrative division can directly refer to the administrative units as these are published by the Government. The problem, however, is two-fold. First, data in key registers are not indexed by Google and therefore are not searchable in the search engine. Second, identifiers for data elements used in key registers make sense only in the local scope of that service. As a result, there is no straightforward way to reference these data. Ed Parsons, a Google's Geospatial Technologist and a member of the Board of Directors of the Open Geospatial Consortium put it: "Information that is not linkable is not used, information that is not used is not valuable" [46].

BR 5 (Quality for Purpose): The notion of quality is context-dependent (e.g., [47]). In short, quality characteristics can vary depending on the application. Therefore, Kadaster understands quality as "fitness for use" rather than a system of measurements (e.g. accuracy, completeness, relevance, consistency etc.) with defined threshold characteristics. In practice, this means data should be published with the quality that fits the context of the users. For some cases a dataset to be considered of high quality because it conforms to a specific standard (e.g., ISO/TS 8000-1:2011 Data quality [48]), however, for other use cases the quality of the data would depend on the degree of interlinking with other datasets. Therefore, Kadaster aims at providing users with enough information, so they can decide by themselves if the data is good enough for the intended use.

Content and services are the main business assets concerned with quality. For content, the priority is to establish a trust to the data by publishing details about quality measures and methods, data collection and versioning history. We argue that services without proper SLA's hardly have value. Data is never perfect, but transparency in providing quality measurements and data provenance can be achieved.

BR 6 (Approachable Spatial Data): Geospatial information has great value for the economy and society alike (e.g., [49], [50]). However, traditionally, geodata were meant to be used mostly within dedicated Geographic Information Systems (GIS). As a result, the emerged ecosystem of geoinformation standards, tooling and related education is focused on the GIS users.

However, due to a great demand for geospatial from outside of the GIS community, the KDP should ensure that its data is findable and (re)usable for other groups such as web developers in general, SW specialists and Business Intelligence (BI) analysts. In this context, it is wanted to support users in searching for information with arbitrary granularity (e.g. within a dataset or a data model).

Regarding findability, the goal is to allow data discovery in a web search engine, which is not the case now. With its recent announcement, Google has introduced a portal for dataset searching (comparable with Google Scholar for scientific papers). It will change the landscape of open data registers. From the data supplier perspective, it will become a necessity to be listed in the Google Datasearch.

BR 7 (Community and Support): The supply-driven culture of open data is focused on guiding users only up to the moment when a data set is discovered. After that, often, users are left alone with all the struggles of using an unknown data source, without any support from the data provider. To improve usage, minimize misinterpretation and misuse, it is important to give good support to the users. This support can be helpdesk, forum, documentation, workflows, and tools. User feedback is essential for creating and improving services offered by the KDP. The understanding of usage patterns enables the user-centric design of data services.

BR 8 (Attractiveness for Data Suppliers): The KDP, as the spatial data platform of the Netherlands, should be an interesting opportunity for data suppliers for data publishing. On the one hand, this implies a business model with low costs to data suppliers. On the other hand, data owners should be able to claim ownership of their data, i.e. dataset appearance and branding should be flexible based on the demands of the owner.

BR 9 (Environmental act): One of the most ambitious projects of the Dutch Government is the revision of the Environment and Planning Act (de Omgevingswet in Dutch; https://aandeslagmetdeomgevingswet.nl/) [51]. The act will replace 15 existing laws (and in the future even 8 more) and was planned to take effect in 2021 (but is expected to get some delay). Apart from being a legal effort, it leads to one of the biggest IT projects in the Netherlands. The KDP should be supportive of the new Environment and Planning Act. It requires data is published in line with the IT architecture supportive for the environmental act [74].

BR 10 (Data Analytics): Evidence-based decision and policymaking is a long-standing trend. GIS tools provide a rich functionality to gain insights based on Kadaster data. However, the recent evolution of web-based software tools has brought functionality of standalone desktop GIS into the web browsers of lay users. The KDP should assist users in sensemaking from its data by supporting web-based analytical tools. This is especially relevant in the context of transparent data journalism and verification of fake news.

Business Intelligence (BI) tools (e.g., Tableau (https://www.tableau.com), Power BI (https://powerbi.microsoft.com/)) and their users is another recognized target for the KDP. Ease of data integration and access to real-time data from within BI tools prevent unwanted copying of data and thus improve the quality of decision and policymaking [52].

Meaningful data exploration requires the KDP to support the visualization of data models and query result sets with proper techniques. Different types of data require different visualization e.g. lists, tables, node-link visualization for graph data, map-like interfaces for spatial information.

BR 11 (Interoperability): Improvement of interoperability has many societal and economic benefits (e.g., [53]). The value of data increases with links to other data, benefiting from the network effect. Interoperability allows data to become part of the web, increase its (re)utilization leading to the growth of both economic and societal benefits.

C. RELATION BETWEEN BUSINESS AMBITIONS AND BUSINESS REQUIREMENTS

The business ambitions and requirements presented in Sections 4.1 and 4.2 together form a foundation of the KDP business case. Figure 2 provides a resulting overview of the relations between the elements of the case. In the figure, grey shapes represent four Business Ambitions that together formulate (solid arrows) the vison of Kadaster. The ambitions are supported (dashed arrows) by eleven BRs.

As shown in Figure 2, certainty and legitimacy is related to metadata as well as to ensuring access to the data from the source, linkability, and quality information. The organization improves its role of a service provider by having rich sources of approachable data, that can be discovered and used by a wide audience. Data analytics adds to the role of a service provider and supports the use-case oriented vision of the Kadaster mission. In general, data interoperability and knowledge graphs are required to cover a broad range of use cases, and in particular the Environmental Act.

V. ROLE OF LINKED DATA

In Section 4 we outlined the reasons for Kadaster to undertake the development of the new data platform. This section elaborates on the role of the LD technology in reaching business ambitions. We distinguish two groups of roles. The first group is comprised of LD capabilities that directly meet the BRs. The second group represent those LD capabilities that were possible before, but LD made them more accessible and cost-effective.

A. DIRECT ROLE: ENABLING TECHNOLOGY

As an innovation, LD is able to drive radical change in the capabilities of information systems and organizations. LD directly influences the way we understand organizational data by enabling an ecosystem of methods and tools that support seamless and simultaneous access to knowledge dispersed between data silos.

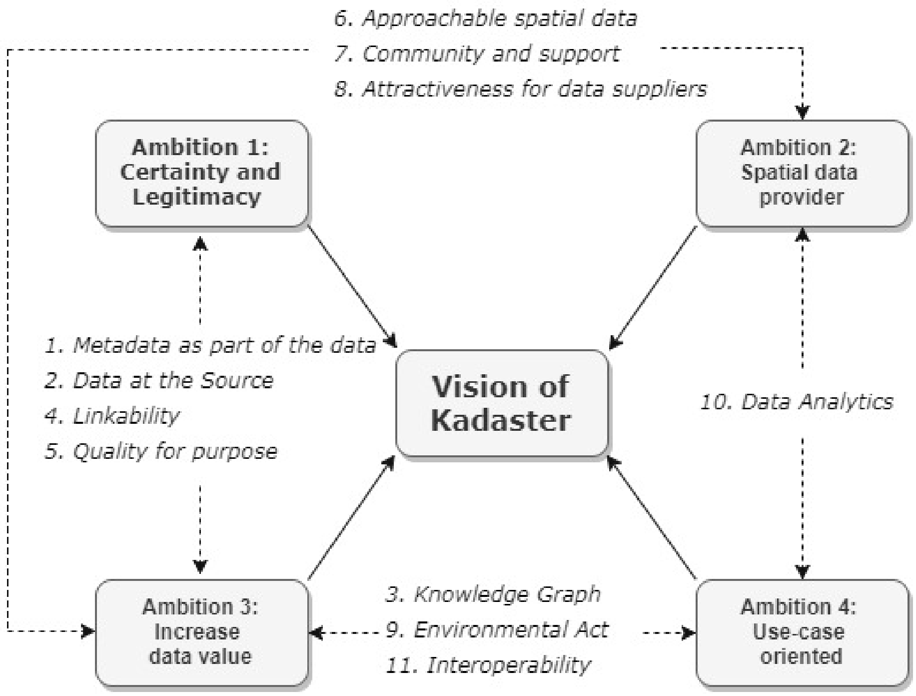

FIGURE 2. Relationships between the business ambitions and business requirements of kadaster. Four business ambitions represented as grey shapes formulate (solid arrows) the vision of kadaster. Eleven business requirements are formulated to support (dashed arrows) the business ambitions.

1) DATA ACCESS ON INSTANCE LEVEL: METADATA AND QUALITY

In contrast to relational databases, where dataset descriptions are stored separately, LD is self-descriptive: linked metadata are just additional triples that are stored together with other data triples. First, it allows publishing metadata on data level and second, it enables querying metadata and data at the same time. The former ensures the capability of capturing complex context for data items including information about data quality [54]. This supports users in understanding if data fit the intended use. Being able to query rich metadata for every data item across multiple datasets allows merging data discovery and data access into one step, thus, streamlining data exploration [55], [56].

However, these advantages can be realized if metadata descriptions are harmonized and well understood by users. There are several standardized metadata vocabularies for describing data at the dataset level. The most prominent examples developed by W3C as recommendations for publishing data on the Web include the Vocabulary for Interlinked Datasets (VoID) [57] and the Data Catalog Vocabulary (DCAT) [2014]. These vocabularies are complementary and are deployed by many open data providers. Extension of DCAT, namely DCAT-AP (application profile) is supported by the European Data Portal (https://www.europeandataportal.eu/).

Quality of available Linked Open Data is very diverse [59]. Existing approaches towards the assessment of LD and metadata quality were overviewed in [60] which resulted in the identification of 18 quality dimensions (e.g. availability, consistency etc.) and 69 metrics (e.g. correctness of facts, adequacy of semantic representation and/or degree of coverage). Non-programming domain experts are able to formalize quality requirements using e.g. a domain-specific language developed in [61]. W3C published the Quality Vocabulary [62] that provided a set of RDF classes and properties to capture and represent the evaluation of a given dataset (or dataset distribution) against a specific quality metric. The vocabulary also provides the mapping between quality dimensions of ISO/IEC 25012 [63] and ones from [60].

The above-mentioned approaches are of academic nature and require extensive research and development efforts to adapt them to a particular case. Users increasingly rely on computational assistance when dealing with data because of the increase in volume, complexity, and creation speed of data. The FAIR Data Principles [64], in contrast, emphases the role of machines in Finding, Accessing, Interoperating, and Reusing data with none or minimal human intervention.

2) DESILOFICATION: KNOWLEDGE AND GAPS

Exposing data in a graph-based format, such as RDF, is an important prerequisite for building KG's [65], however, it does not enable seamless out of the box reasoning over these data. What it does, is by taking down technical barriers between data silos it exposes knowledge gaps between divisions of the government. Therefore, in order to build a KG, these gaps need to be bridged. However, how to identify

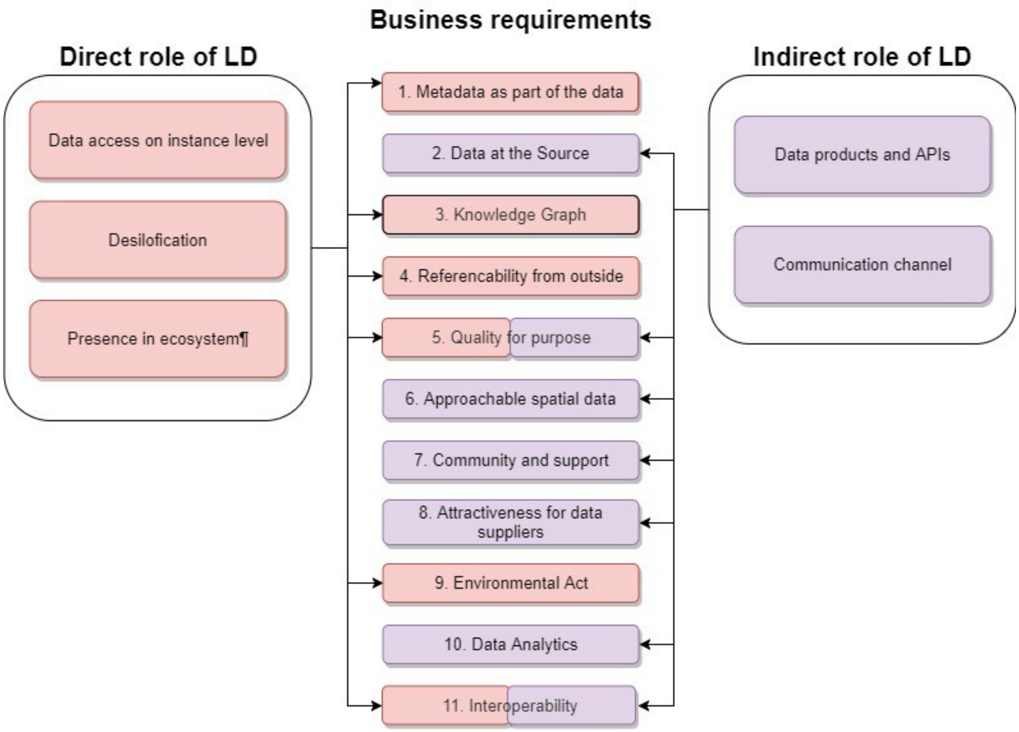

FIGURE 3. Role of LD in business requirements. The colours denote the role linked data play for each of the business requirements. Purple represents an indirect role, pink - direct role.

them if they lie in areas outside of departmental knowledge? This is a chicken and egg problem. Gaps cannot be identified upfront before constructing a KG which in turn cannot be created without identification of gaps. Even though LD can't solely solve this challenge since it is more of organizational nature rather than technical, it creates an environment where these gaps can be formally identified and represented.

Governments own and control systems of legal definitions. These systems are hierarchical and therefore their structure can be traced top-down to identify the precise meaning of relations on ontology level. However, this approach does not help in defining instance-level relations because their number and complexity grow very fast with every instance added to a KG. The network effect makes it difficult to foresee and formalize all possible relations. Instead, a bottom-up use case-driven approach allows defining arbitrary relations between instances.

Naturally, Kadaster data are rich with spatial and temporal information. Space and time are fundamental sources of contextual information and therefore, they allow linking data instances that are semantically disjoint on the ontological level. This is especially relevant in cases when the top-down approach is hindered (or even not possible) due to existing semantic heterogeneity of legal definitions and terminology between independent governmental agencies. In this context, one particular area of interest is the ownership of so-called "link sets", datasets that link other datasets (e.g. formal government registries). Who will take the lead, who is responsible for wrong links, who will pay? Link sets are essential for building large KG's, but so far not that many "owned" linked sets are published.

Governmental data is used for making legal decisions. This put additional requirements to the accuracy of semantic relations between data items in KG's that go well beyond the capabilities of *owl:sameAs* and *rdfs:seeAlso* (e.g., see [66], [67]). On the other hand, desilofication is a perfect opportunity to identify existing discrepancies between key registries. Moreover, running SPARQL queries across key registers allow finding outliers in data. Comparison with external resources can highlight inner inaccuracies as well.

3) PRESENCE IN ECOSYSTEM: POINT OF CRYSTALLIZATION FOR PUBLIC DATA

By publishing and maintaining Linked Data, Kadaster becomes a part of the Web of Data and thus participates in the surrounding data ecosystem. Projects like Wikidata (https://www.wikidata.org) crowdsource interlinking and publish link sets between common resources. Interlinking with and from resources on central hubs (e.g. DBpedia) of this ecosystem promotes the use of Kadaster data among a wide range of Web users. By traversing the links, users can discover more resources in the "follow your nose" fashion. However, to be able to discover data via Google search would be an ultimate solution. Google uses the schema.org and DCAT vocabularies for building their KG. Therefore, providing a mapping to this vocabulary is an essential prerequisite for being indexed by google crawlers.

TABLE 2. Summary of linked data contribution to business requirements

Business requirements	The role of Linked Data	Reference
1. Metadata as part of the data	Enables publishing metadata on data instance level allowing simultaneous data discovery and access.	[37], [54], [55], [56], [57], [58]
2. Data at the source	Facilitates a flexible and cost-efficient process of creating and managing use-case oriented data products and services.	[2], [73]
3 Knowledge Graph	LD is an enabling prerequisite for building KG's.	[18], [43], [65], [66]
4. Linkability	Use of HTTP URIs for data elements enables data to be a part of the global data ecosystem, thus making it globally accessible and referenceable.	[10], [46], [68]
5 Quality for purpose	Information about quality is a part of metadata. LD makes this information available at the data instance level, so users can see if data are suitable for reuse.	[4], [34], [59], [60], [61]
6 Approachable spatial data	LD facilitates the improvement of discoverability and simplifies the creation of custom dissemination channels for diverse user communities.	[70]
7. Community and support	LD explicitly defines meaning by the means of vocabularies, thus, creating a common language for referring to data items.	[34], [35]
8 Attractive for data suppliers	Enables seamless access (for users) to datasets originated from different suppliers. Allows flexible branding of data services.	[69], [72]
9. Environmental Act	Capability to see a multitude of data sets as a whole is required by the New Environmental Act. LD is seen to be the main facilitating technology.	[74]
10. Data Analytics	LD enables a flexible mechanism for the creation of different dissemination channels (e.g. for BI tools). Semantic queries can be stored and exposed supporting transparency of analysis.	[16], [22] [52], [56]
11. Interoperability	LD standards ensure interoperability on semantic and syntactic levels.	[4], [75]

By allowing semantic and syntactic interoperability, LD helps to transcend barriers not only between departmental data repositories but also between different governmental bodies. Similar to DBpedia [68], KDP acts as a central hub for the publishing and interlinking governmental data on a national level. In this sense, the strong spatiotemporal component of Kadaster data is seen as an important competitive advantage [69]. It provides information dimensions needed for interrelating data that have very little in common otherwise. This is also valid for semantically heterogeneous data that is voluntarily produced by the public. These data can be linked and structured around existing geospatial resources.

B. INDIRECT ROLE: FACILITATING TECHNOLOGY

Points discussed in Section 5.1. represent new capabilities that were hardly possible without LD technology. In this section, we discuss roles that can be fulfilled with already existing technologies. However, LD provides more accessible and cost-effective solutions.

1) DATA PRODUCTS AND APIS

By cutting off options of bulk download (in the future) Kadaster will entirely rely on data APIs for data dissemination. Therefore, the ability to create a custom data API in a timely and efficient manner is fundamental for meeting requirements of the business case.

On the one hand, in this context, LD is merely yet another channel of data dissemination (like e.g. WFS) targeting specific needs of a user community (in this case SW developers). On the other hand, LD facilitated the creation of APIs in a flexible and efficient way. It creates an environment where siloed data can be meaningfully (re)combined into use-case oriented data products to feed the APIs [70].

Moreover, the content returned by APIs can be enriched with dereferenceable URIs and semantic descriptions (e.g., by means of JSON-LD, see https://json-ld.org) thus allowing us to provide users with the best from both worlds: a data service with high availability and short response time which serves semantically unambiguous data. In a similar way, existing Spatial Data Infrastructures can be enriched

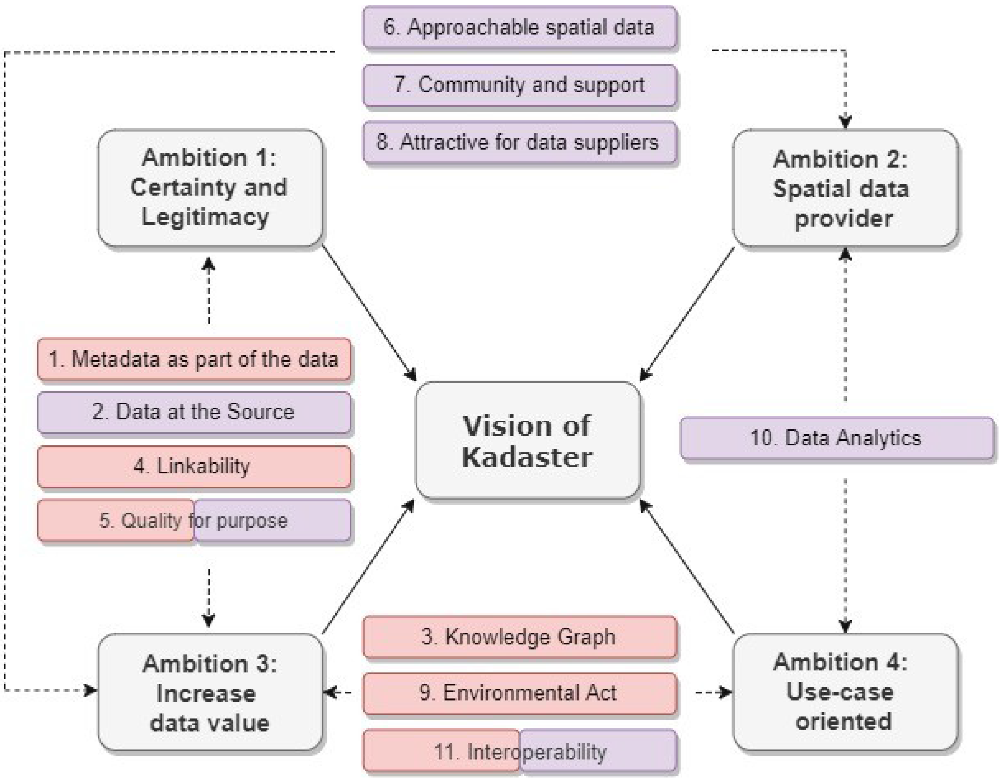

FIGURE 4. The role of LD in the business vision of Kadaster. Grey shapes represent four Business Ambitions that together comprise (solid arrows) the vison of Kadaster. The ambitions are supported (dashed arrows) by eleven BRs. Purple shapes represent BRs where LD plays an indirect role, pink - direct role.

with URIs to provide users with better metadata for already familiar resources [71].

Adjusting URI strategies [72] to reflect the origin of the data allows flexible and cost-efficient branding of data. In this way, LD combines efficiency advantages of a central system, but from the outside, it can be presented as different datasets with different branding in different domains, which lowers the barriers for data suppliers to add data [73].

2) EXPLICIT MEANING

Formalization of meaning is required by LD. Ontologies can be seen as shared systems of interpretation of meaning for users. This allows establishing identity, and, thus, allows different users to refer to the same information in an explicit way. As pointed out by [34], LD act as a means of machine mediated communication between data providers and users meaning is required

C. ROLE OF LINKED DATA: SUMMARY

The role of the LD technology in the business requirements is graphically summarized and depicted in Figure 3. Table 2 provides a concise description of LD contribution and lists related work.

VI. DISCUSSION

The discussed Kadaster Data Platform as a project, team and system is an internal name at Kadaster. For the branding to the outside world, the PDOK name is used. Current version to the PDOK portal provides access to three key registers of the Dutch Government as LD, namely BRT, BRK and BAG. Apart from the Kadaster datasets, the platform hosts three external datasets. They can be accessed via an open SPARQL endpoint at https://data.pdok.nl/sparql.

BRs related to the business ambition of being a spatial data provider (BR6, BR7 and BR8) have an indirect relation with LD (see Figure 3), therefore the ambition can be fulfilled without the use of LD. The validity of this observation is supported e.g. by the experience of the PDOK shared service that successfully provided spatial data using the traditional approach of XML-based Spatial Data Infrastructures. However, this approach seemingly has already reached its full potential. It exhibited limits, especially when it came to widening user community and ensuring interoperability outside of GIS world. As was explained in Section 5.2., LD makes it easier to reach new customer groups.

Although the ambition of being a spatial data provider is quite Kadaster specific, because of its role in the PDOK shared service, all three other ambitions are relevant for many government organizations. Therefore, we argue that the reasons to invest in the development of a data platform powered by the LD technology are generic and are not limited to the Kadaster case only.

Another argument for the general applicability of the results is based on the mission of Kadaster namely, to provide societal benefits. This mission is not unique for Kadaster and is relevant to at least all the government organizations. Then it is also likely that the business rationale presented in this paper will hold to a large extent for all government organizations.

When summarizing all of the Kadaster ambitions to evoke the effort of building a new data platform, it comes down to ''achieve the maximum level of data interoperability''. From literature, it is well known that open standards such as from W3C, contribute best to interoperability. It is also known that interoperability contains many layers and aspects, such as technical, semantical, organizational [75]. LD can be seen as an integrated approach in which open standards, in cohesion, cover a broad range of interoperability aspects and layers.

VII. CONCLUSION & FURTHER RESEARCH

The exploratory case study presented in the paper investigated the business rationale behind implementing LD in a government context. Analysis of BRs and the capabilities of LD technology revealed that LD provides a major contribution to the design of the data platform. This allowed us to conclude that LD can support the business vision of governmental organizations. Therefore, the topic of LD should be part of business discussions, including vision and ambitions statements.

Figure 4 presents a composite visualization of the role of LD in the business vision of Kadaster defined as four ambitions and eleven business requirements (as a summary of Figures 2 and 3). LD plays a direct role in three out of four business ambitions. As such, it enables a use-case oriented vision, increases the value of the data and ensures the certainty of the data and legitimacy of the organization.

LD adds most to the Kadaster ambitions of providing certainty and legitimacy, to increase data value and to be use-case-oriented. The ambition to become the spatial data provider of the Netherlands is not directly influenced by LD, but without LD it will be difficult to reach users outside the GIS community. In the end, LD puts data interoperability to the next level, a desire by many organizations, including Kadaster.

Further efforts should be focused on the research of the adoption of LD within government organizations. One possible direction is the topic of Knowledge Graphs, collections of interlinked LD datasets that have many potential use cases in the context of e-government.

In the context of the KDP, further research is twofold. We plan to further investigate the quality aspect of the published LD and carry out a reflection phase to see up to what extent the platform satisfies the Business Requirements after one year.

REFERENCES

[1] T. Berners-Lee, J. Hendler, and O. Lassila, ''The semantic Web,'' *Scientific Amer.*, vol. 284, no. 5, pp. 34–43, May 2001, doi: 10.1038/scientificamerican0501-34.

[2] S. Ronzhin, E. Folmer, and R. Lemmens, ''Technological Aspects of (Linked) Open Data,'' in *Open Data Exposed*, vol. 30, B. van Loenen, G. Vancauwenberghe, and J. Crompvoets, Eds. Hague, The Netherlands: T.M.C. Asser Press, 2018, pp. 173–193.

[3] A. Osterwalder and Y. Pigneur, *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*. Hoboken, NJ, USA: Wiley, 2010.

[4] P. Archer, M. Dekkers, S. Goedertier, and N. Loutas, *Study on Business Models for Linked Open Government Data*. Brussels, Belgium: European Union, 2013.

[5] I. Mitchell and M. Wilson, ''Linked data: Connecting and exploiting big data,'' Fujitsu Services Ltd., Basingstoke, U.K., White Paper, Mar. 2012. [Online]. Available: https://www.fujitsu.com/uk/Images/Linked-data-connecting-and-exploiting-big-data-(v1.0).pdf

[6] G. Kobilarov, ''Media MS emantic Web—How the BBC uses DB pedia and linked fata to make connections,'' in *The Semantic Web: Research and Applications*, vol. 5554, L. Aroyo, P. Traverso, F. Ciravegna, P. Cimiano, T. Heath, E. Hyvonen, R. Mizoguchi, E. Oren, M. Sabou, and E. Simperl, Eds. Berlin, Germany: Springer, 2009, pp. 723–737.

[7] Y. Raimond, T. Scott, S. Oliver, P. Sinclair, and M. Smethurst, ''Use of semantic Web technologies on the BBC Web sites,'' in *Linking Enterprise Data*. Boston, MA, USA: Springer, 2010, pp. 263–283.

[8] R. K. Yin, *Case Study Research and Applications: Design and Methods*. Newbury Park, CA, USA: Sage, 2017.

[9] R. L. Baskerville and A. T. Wood-Harper, ''A critical perspective on action research as a method for information systems research,'' *J. Inf. Technol.*, vol. 11, no. 3, pp. 235–246, 1996.

[10] T. Heath and C. Bizer, ''Linked data: Evolving the Web into a global data space,'' *Synth. Lectures Semantic Web, Theory Technol.*, vol. 1, no. 1, pp. 1–136, Feb. 2011, doi: 10.2200/S00334ED1V01Y201102WBE001.

[11] T. Berners-Lee. (2011). *Linked Data-Design Issues (2006)*. [Online]. Available: http://www.w3.org/DesignIssues/LinkedData.html

[12] World Wide Web Consortium, Recommendation. (Feb. 25, 2014). *RDF 1.1 Concepts and Abstract Syntax*. [Online]. Available: https://www.w3.org/TR/rdf11-concepts/

[13] A. Silberschatz, H. F. Korth, and S. Sudarshan, ''Data models,'' *ACM Comput. Surveys*, vol. 28, no. 1, pp. 105–108, 1996.

[14] R. Fielding, *Hypertext Transfer Protocol*, document RFC 2616, Jun. 1999.

[15] C. Bizer, T. Heath, and T. Berners-Lee, ''Linked data–the story so far,'' *Int. J. Semantic Web Inf. Syst.*, vol. 5, no. 3, pp. 1–22, Jul. 2009, doi: 10.4018/jswis.2009081901.

[16] J. Perez, M. Arenas, and C. Gutierrez, ''Semantics and Complexity of SPARQL,'' in *Proc. Int. semantic Web Conf.*, 2006, pp. 30–43.

[17] D. Brickley, R. V. Guha, and B. McBride, *RDF schema 1.1. W3C recommendation*. Cambridge, MA, USA: World Wide Web Consortium, 2014.

[18] F. Bauer and M. Kaltenbock, *Linked Open Data: The Essentials*, vol. 710. Vienna, Italy: Mono/Monochrom, 2011.

[19] K. Janowicz, P. Hitzler, B. Adams, D. Kolas, and C. Vardeman II, ''Five stars of linked data vocabulary use,'' *Semantic Web*, vol. 5, no. 3, pp. 173–176, 2014.

[20] P.-Y. Vandenbussche, G. A. Atemezing, M. Poveda-Villalón, and B. Vatant, ''Linked open vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web,'' *Semantic Web*, vol. 8, no. 3, pp. 437–452, Dec. 2016.

[21] J. G. Breslin, D. O'Sullivan, A. Passant, and L. Vasiliu, ''Semantic Web computing in industry,'' *Comput. Ind.*, vol. 61, no. 8, pp. 729–741, Oct. 2010.

[22] M. Lnenicka and J. Komarkova, ''Big and open linked data analytics ecosystem: Theoretical background and essential elements,'' *Government Inf. Quart.*, vol. 36, no. 1, pp. 129–144, Jan. 2019.

[23] T. Pellegrini, C. Dirschl, and K. Eck, ''Linked data business cube: A systematic approach to semantic Web business models,'' in *Proc. 18th Int. Academic MindTrek Conf. Media Bus., Manage., Content Services*, 2014, pp. 132–141.

[24] C. Bizer, T. Heath, D. Ayers, and Y. Raimond, ''Interlinking open data on the Web Demonstrations track,'' in *Proc. 4th Eur. Semantic Web Conf.*, Innsbruck, Austria, 2007, pp. 1–7.

[25] S. Halford, C. Pope, and M. Weal, ''Digital futures? Sociological challenges and opportunities in the emergent semantic Web,'' *Sociology*, vol. 47, no. 1, pp. 173–189, Feb. 2013.

[26] A. Blumauer, ''Linked Data in Unternehmen. Methodische Grundlagen und Einsatzszenarien,'' in *Linked Enterprise Data*. Springer, 2014, pp. 3–20.

[27] E. Folmer and D. Krukkert, ''Linked data for transaction based enterprise interoperability,'' in *Proc. Int. IFIP*. Berlin, Germany: Springer, 2015, pp. 113–125.

[28] T. Berners-Lee. *Web Architecture From 50,000 Feet*. Accessed: Dec. 9, 2003. [Online]. Available: http://www.w3c.org/DesignIssues/Architecture.html

[29] S. Brinker, ''Business models for linked data and Web 3.0,'' *Chiefmartec*, vol. 10, p. 21, Aug. 2010.

[30] L. van den Brink, P. Barnaghi, J. Tandy, G. Atemezing, R. Atkinson, B. Cochrane, Y. Fathy, R. García Castro, A. Haller, A. Harth, K. Janowicz, Å. Kolozali, B. van Leeuwen, M. Lefranãois, J. Lieberman, A. Perego, D. Le-Phuoc, B. Roberts, K. Taylor, and R. Troncy, "Best practices for publishing, retrieving, and using spatial data on the Web," *Semantic Web*, vol. 10, no. 1, pp. 95–114, Dec. 2018.
[31] *Best practices for publishing linked data*, W. W. W. Consortium, Cambridge, MA, USA, 2014.
[32] A. Latif, A. U. Saeed, P. Hoefler, A. Stocker, and C. Wagner, "The linked data value chain: A lightweight model for business engineers.," in *Proc. I-SEMANTICS*, 2009, pp. 568–575.
[33] T. Kinnari, "Open data business models for media industry-finnish case study," M.S. thesis, Dept. Inform. Service Economy, Aalto Univ., Espoo, Finland, 2013. [Online]. Available: https://aaltodoc.aalto.fi/bitstream/handle/123456789/10140/hse_ethesis_13166.pdf?sequence=1&isAllowed=y
[34] K. Janowicz, S. Scheider, T. Pehle, and G. Hart, "Geospatial semantics and linked spatiotemporal data-Past, present, and future," *Semantic Web*, vol. 3, no. 4, pp. 321–332, 2012.
[35] S. Scheider, K. Janowicz, and B. Adams, "The observational roots of reference of the semantic Web," 2012, *arXiv:1206.6347*. [Online]. Available: http://arxiv.org/abs/1206.6347
[36] C. Van Wagoner. (2017). *What is a Modern Data Platform*. [Online]. Available: http://www.dbta.com/Editorial/Trends-and-Applications/What-is-a-Modern-Data-Platform-118864.aspx
[37] L. McHugh, *Measuring the Value of Metadata*. Birmingham, AL, USA: Baseline Consulting, 2009.
[38] B. Jacob and J. Ortiz, "Data.world: A platform for global-scale semantic publishing," in *Proc. Int. Semantic Web Conf.*, 2017. [Online]. Available: http://ceur-ws.org/Vol-1963/paper605.pdf
[39] Swirrl. (2018). *PublishMyData Linked Data Publising Platform*. [Online]. Available: http://www.swirrl.com/
[40] E. Ruijer, S. Grimmelikhuijsen, M. Hogan, S. Enzerink, A. Ojo, and A. Meijer, "Connecting societal issues, users and data. scenario-based design of open data platforms," *Government Inf. Quart.*, vol. 34, no. 3, pp. 470–480, Sep. 2017.
[41] (2019). *Basisregistratie Adressen en Gebouwen (BAG)*. [Online]. Available: https://bag.basisregistraties.overheid.nl/
[42] (2019). *Basisregistratie Topografie (BRT)*. [Online]. Available: https://brt.basisregistraties.overheid.nl/
[43] L. Ehrlinger and W. Wöß, "Towards a definition of knowledge graphs," in *Proc. SEMANTiCS (Posters, Demos, SuCCESS)*, vol. 48, Sep. 2016. [Online]. Available: http://ceur-ws.org/Vol-1695/paper4.pdf
[44] E. Folmer, W. Beek, L. Rietveld, S. Ronzhin, R. Geerling, and D. den Haan, "Enhancing the usefulness of open governmental data with linked data viewing techniques," in *Proc. 52nd Hawaii Int. Conf. Syst. Sci.*, 2019. [Online]. Available: https://scholarspace.manoa.hawaii.edu/bitstream/10125/59728/0288.pdf
[45] *19115: 2003 Geographic Information-Metadata*, International Organization for Standardization, London, U.K., 2003.
[46] E. Parsons. (2017). *If You Can't Link to Does it Exist*. [Online]. Available: https://www.edparsons.com/2017/09/cant-link-exist/
[47] E. J. A. Folmer and J. Verhoosel, "State of the art on semantic IS standardization, interoperability & quality," UT, CTIT, TNO en NOiV, Enschede, The Netherlands, 2011.
[48] *Data Quality—Part 1: Overview (ISO/TS 8000-1:2011)*, International Organization for Standardization, London, U.K., 2011.
[49] A. Krek, "Geographic information as an economic good," in *GIS for Sustainable Development*. Boca Raton, FL, USA: CRC Press, 2005, pp. 105–124.
[50] W. T. Castelein, A. Bregt, and Y. Pluijmers, "The economic value of the Dutch Geo-information sector," *Int. J. Spatial Data Infrastruct. Res.*, vol. 5, no. 5, pp. 58–76, 2010.
[51] J. Bulles, B. Cartigny, and P. Bollen, "Analyzing the new 2019 dutch environment and planning act," in *Proc. OTM Int. Conf. Meaningful Internet Syst.*, 2017, pp. 163–172.
[52] Chen, Chiang, and Storey, "Business intelligence and analytics: From big data to big impact," *MIS Quart.*, vol. 36, no. 4, pp. 1165–1188, 2012.
[53] E. J. A. Folmer, "Quality of semantic standards," Ph.D. dissertation, Dept. Ind. Eng. Bus. Inf. Syst., Univ. Twente, Enschede, The Netherlands, 2012.
[54] O. Hartig and J. Zhao, "Publishing and consuming provenance metadata on the Web of linked data," in *Proc. Int. Provenance Annotation Workshop*, 2010, pp. 78–90.
[55] W. Beek and E. Folmer, "An integrated approach for linked data browsing," in *Proc. ISPRS*, Jul. 2017, pp. 35–38.
[56] E. Folmer, W. Beek, and L. Rietveld, "Linked Data Viewing as part of the Spatial Data Platform of the Future," in *Proc. ISPRS*, Jul. 2018, pp. 49–52.
[57] K. Alexander, R. Cyganiak, M. Hausenblas, and J. Zhao, "Describing linked datasets—On the design and usage of voiD, the 'vocabulary of interlinked datasets,"' in *Proc. WWW, Workshop Linked Data Web (LDOW)*, Madrid, Spain, 2009. [Online]. Available: http://ceur-ws.org/Vol-538/ldow2009_paper20.pdf
[58] F. Maali, J. Erickson, and P. Archer, *Data catalog vocabulary (DCAT)*, vol. 16. W3c Recommendation, 2014.
[59] J. Debattista, C. Lange, S. Auer, and D. Cortis, "Evaluating the quality of the LOD cloud: An empirical investigation," *Semantic Web*, vol. 9, no. 6, pp. 859–901, Sep. 2018.
[60] A. Zaveri, A. Rula, A. Maurino, R. Pietrobon, J. Lehmann, and S. Auer, "Quality assessment for linked data: A survey," *Semantic Web*, vol. 7, no. 1, pp. 63–93, Mar. 2015.
[61] J. Debattista, S. Auer, and C. Lange, "Luzzu—A methodology and framework for linked data quality assessment," *J. Data Inf. Qual.*, vol. 8, no. 1, pp. 1–32, Nov. 2016.
[62] World Wide Web Consortium, Working Group Note. (Dec. 15, 2016). *Data on the Web Best Practices: Data Quality Vocabulary*. [Online]. Available: https://www.w3.org/TR/vocab-dqv/
[63] *Data Quality Model (ISO/IEC 25012:2008)*, International Organization for Standardization, London, U.K., 2008.
[64] M. D. Wilkinson *et al.*, "The FAIR guiding principles for scientific data management and stewardship," *Sci. Data*, vol. 3, no. 1, Dec. 2016, Art. no. 160018, doi: 10.1038/sdata.2016.18.
[65] H. Paulheim, "Knowledge graph refinement: A survey of approaches and evaluation methods," *Semantic Web*, vol. 8, no. 3, pp. 489–508, Dec. 2016, doi: 10.3233/SW-160218.
[66] W. Beek, S. Schlobach, and F. van Harmelen, "A contextualised semantics for owl: Sameas," in *Proc. Eur. Semantic Web Conf.*, 2016, pp. 405–419.
[67] H. Halpin, P. J. Hayes, and H. S. Thompson, "When owl: Sameas isn't the same redux: Towards a theory of identity, context, and inference on the semantic Web," in *Proc. Int. Interdiscipl. Conf. Modeling*, 2015, pp. 47–60.
[68] C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann, "DBpedia–A crystallization point for the Web of data," *J. Web Semantics*, vol. 7, no. 3, pp. 154–165, Sep. 2009.
[69] J. Black, "On the derivation of value from geospatial linked data," Ph.D. dissertation, Dept. Comput. Sci. Eng., Univ. Southampton, Southampton, U.K., 2013.
[70] S. Abbas and A. Ojo, "Towards a linked geospatial data infrastructure," in *Proc. Int. Conf. Electron. Government Inf. Syst. Perspective*, 2013, pp. 196–210.
[71] S. Ronzhin, "Next Generation of Spatial Data Infrastructure: Lessons from Linked Data implementations across Europe," *Int. J. Spat. Data Infrastruct. Res*, vol. 14, pp. 84–106, Jun. 2019.
[72] World Wide Web Consortium, Interest Group Note. (Dec. 3, 2008). *Cool URIs for the Semantic Web*. [Online]. Available: https://www.w3.org/TR/cooluris/
[73] *Best Practices for Publishing Linked Data*, W. W. W. Consortium, Cambridge, Ma, USA, 2014.
[74] T. Sloos, M. Brattinga, and S. Oostenbrink. (2018). *Kaderstellende Notities URI-Strategie. Deelprogramma Digitaal Stelsel Omgevingswet*. [Online]. Available: https://aandeslagmetdeomgevingswet.nl/publish/pages/143597/uri-strategie_12_maart_2018.pdf
[75] G. Percivall, "The application of open standards to enhance the interoperability of geoscience information," *Int. J. Digit. Earth*, vol. 3, no. 1, pp. 14–30, Apr. 2010, doi: 10.1080/17538941003792751.

ERWIN FOLMER received the Ph.D. degree, in 2012. He joined TNO, in 2001, and became a Senior Scientist on the topic of interoperability and standards. Since 2009, he joined the University of Twente to start a Ph.D. research on the standardization topic, while continuing his work for TNO. His Ph.D. degree was based on the Quality of Semantic Standards thesis. From 2013 to 2014, he was a Visiting Researcher with the ERCIS/University of Munster. In 2015, he joined Kadaster, continuing the work on standards and interoperability with special focus on spatial data platforms. He is also chairing the Platform Linked Data Netherlands, an open community to support the Linked Data adoption.

STANISLAV RONZHIN received the M.Sc. degree from Utrecht University, in 2015. He is currently pursuing the Ph.D. degree with the Faculty of Geo-Information Science and Earth Observation, University of Twente. His primary interests include exploratory search and semantic enrichment of Linked Geospatial Data. He co-created the Living Textbook, a web-based collaborative environment for learning and authoring ontology-based concept maps. Apart from the University of Twente, he is heavily involved in the research and development of the Kadaster Data Platform at the Netherlands' Cadastre Land Registry and Mapping Agency (Kadaster). He also teaches the Linked Data technology for Business Process Integration to master students.

JOS VAN HILLEGERSBERG is currently a Full Professor with Business Information Systems. He is the Head of the Department of Industrial Engineering and Business Information Systems, University of Twente. He is contributing to several national and international projects on design of collaborative businesses and industrial networks applying ICT, such as data analytics, architecture transformation, agent technology, and sensor data. He is the Chairman of the Program Committee of the Dutch Research Institute for Advanced Logistics. Before joining the University of Twente, he was on the faculty of the Rotterdam School of Management at the Erasmus University, working on component-based software systems, IT management, global outsourcing, and agent systems for supply chains. He also worked for several years in business. At AEGON, he was Component Manager for the setup of an Internet Bank. At IBM, he worked on artificial intelligence and expert systems. His research deals with innovation of supply chains and business networks using ICT.

WOUTER BEEK is currently a Postdoctoral Researcher from VU University Amsterdam. He has co-developed the LOD Laundromat (Best Paper Award (ISWC) and the Best LOD Application Award, in 2015), LOTUS (2nd place LOD Award 2016 and ESWC 2016), LOD-a-lot (best paper nominee, SEMANTiCS 2017), sameAs.cc (best resource paper award, ESWC 2018). He is interested in the Semantic Web as a platform for knowledge-intensive applications, the deployment of large-scale knowledge bases for innovative reuse, and the interaction between Web semantics and pragmatics, including the empirical study of semantics. He is working together with Frank van Harmelen in the Knowledge, Reasoning and Representation (KR&R) Group. In addition to his scientific work, he is co-founder of the start-up company Triply, which sells large-scale Linked Data deployments to customers in geospatial services, galleries, libraries, archives, and museums, academic research, manufacturing, and eGovernment.

ROB LEMMENS received the Ph.D. degree in geoinformatics from the Delft University of Technology. In 2010, he worked as a Visiting Scientist with the Google Headquarters, Mountain View, USA, on Google Maps. In 2013, he was a Visiting Scientist with the Department of Geography, Hunter College of the City University of New York, focusing on visualization of Linked Open Data and ontologies for VGI. He has expertise in geo web technology, spatial data infrastructures and semantic modeling of distributed geo-information services, and ontology-based geo-information. He has led the ontology development in the SEMA project which bridges human sensor information with semantic context models. He is involved in the Platform Linked Data Netherlands and Co-Winner of the hackathon on data journalism, in 2016. In the COST ENERGIC project http://vgibox.eu he co-created the VGI Knowledge Portal. He is lead architect for ontology development for the Living textbook application, which is currently deployed in the http://www.eo4geo.eu/ project. He coauthored several publications on deploying ontologies in learning environments, geoprocessing, and geodata discovery. He supervises M.Sc. and Ph.D. student projects in the domain of geosemantic modeling and Linked Data.