

Raising Interoperability among Base Registries: The Evolution of the Linked Base Registry for Addresses in Flanders.

Raf Buyle^a, Ziggy Vanlshout^b, Serena Coetzee^c, Dieter De Paepe^a, Mathias Van Compennolle^d, Geert Thijs^b, Bert Van Nuffelen^e, Laurens De Vocht^a, Peter Mechant^d, Björn De Vidts^b and Erik Mannens^a

^aimec IDLab - Ghent University, Ghent, Belgium

^bInformatie Vlaanderen, Flemish Government, Brussels, Belgium

^cCentre for Geoinformation Science, Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

^dimec MICT - Ghent University, Ghent, Belgium

^eTenforce, Leuven, Belgium

ARTICLE INFO

Article history:

Received 17 Feb 18

Received in revised form 27 Aug 18

Accepted 26 Oct 18

Keywords:

Semantic Web,

Ontology,

e-government,

Address Registry,

Linked Open Data,

Semantic Interoperability

ABSTRACT

The transformation of society towards a digital economy and government austerity creates a new context leading to changing roles for both government and private sector. Boundaries between public and private services are blurring, enabling government and private sector to collaborate and share responsibilities. In Belgium, the regional Government of Flanders embedded the re-use of public sector information in its legislation and published a data portal containing well over 4000 Open Datasets. Due to a lack of interoperability, interconnecting and interpreting these sources of information remain challenges for public administrations, businesses and citizens. To dissolve the boundaries between the data silos, the Flemish government applied Linked Data design principles in an operational public sector context. This paper discusses the trends we have identified while 'rewiring' the Authentic Source for addresses to a Linked Base Registry. We observed the impact on multiple interoperability levels; namely on the legal, organisational, semantic and technical level. In conclusion Linked Data can increase semantic and technical interoperability and lead to a better adoption of government information in the public and private sector. We strongly believe that the insights from the past thirteen years in the region of Flanders could speed up processes in other countries that are facing the complexity of raising technical and semantic interoperability.

1. Introduction

The transformation of society towards a digital economy created a new context leading to changing roles where boundaries between public and private services are blurring [24]. This enables the government and the private sector to collaborate and share responsibilities [21]. In Belgium, the Flemish Government is undertaking ambitious reforms to further transform public services. Flanders is the northern federated state of Belgium¹ and has over 6 million inhabitants, or about 60% of the population of Belgium. The policy letter of the Vice Minister President states that Open Data is 'the norm' and that the government should focus on economic added value and close collaboration with the private sector [34]. The Flemish Government is already working with the private sector to co-create and co-finance the development and maintenance of open data sources. The Large-scale Reference Database (LRD) is an official Flemish data source with precise and detailed location information on buildings, parcels, roads, watercourses and railroads, identifying millions of objects in Flanders and providing a source for address positions. The LRD is the result of a public-private

partnership between the Flemish Government and the utility sector, with a substantial setup cost of EUR 93 Mio and an annual maintenance budget of EUR 7 Mio [20].

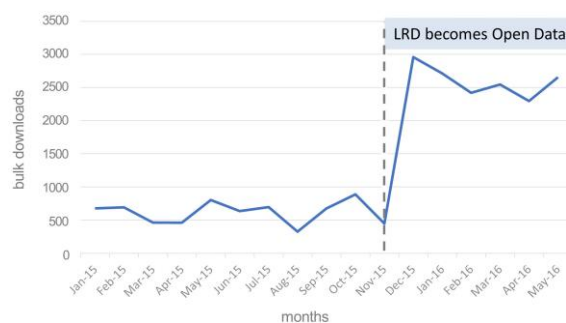


Fig. 1. The impact of opening LRD data (in terms of downloads), showing the increase of downloads due to opening-up the dataset (L. De Wolf, Flanders Information Agency).

¹ <https://www.vlaanderen.be/en/discover-flanders>

Since the LRD became Open Data in November 2015, more than 2500 bulk downloads (Fig. 1) and over 2.9 Mio requests were successfully processed each month.

Re-use of government data is considered to be an enabler of Open Government [39]. The **problem statement** of this article is: even though the Government of Flanders embedded the re-use of public sector information (PSI) in legislation [7, 11] and published a data portal² containing over 4000³ Open Datasets, interconnecting and interpreting these sources of information remains challenging for businesses, citizens and public administrations alike.

Public administrations struggle to deliver interconnected and cross-sectoral services due to sectoral specialisation or “departmentalisation” [39]. Due to interoperability problems, including adequate semantic standards and scarcity of web-oriented architecture, private partners struggle to reuse Public Sector Information (PSI). Interoperability is the ability of organisations to share information and knowledge, through the business processes they support, by means of the exchange of data between their ICT systems [28]. Public Sector Information is often modelled from a single perspective and therefore cannot be integrated with other information sources, applications and business processes [15]. The lack of standards causes high costs due to data transformations and mapping [39]. To overcome these hurdles, we need to address multiple interoperability levels; namely the legal, organisational, semantic and technical level [22].

This has led to a demand for stable, governed data standards [41, 4], which are “technical documents designed to be used as a rule, guideline or definition. They are consensus-built, repeatable ways of doing something”⁴. The Flemish Government launched an interoperability programme: Open Standards for Linked Organisations (OSLO) [22] which builds upon the principles of the European Interoperability Framework (EIF) [27]. Interoperability Frameworks assume a hierarchy in terms of maturity with regard to layers of interoperability [41]. This means organisational and legal interoperability can only be achieved when standards for semantic and technical interoperability have successfully been implemented. Therefore OSLO addresses both semantic and technical interoperability.

Incompatibilities between legislation in different policy domains and legal frameworks make working together more complex. Legal Interoperability refers to aligned legislation between different organisations. Organisational interoperability refers to aligned business processes between public administrations. This implies integrating business processes and related data exchange. [22, 27].

Semantic interoperability focusses on the meaning of data elements. As inter-organisational information systems only work when they communicate with other systems and interact with people, it includes developing vocabularies to describe data exchanges and ensures that data elements are understood in the same way by different parties when communicating [22, 33]. According to EIF, semantic interoperability also covers the syntactic aspect which refers to the grammar and format, such as HTML or XML.

Technical Interoperability is often centred on (communication) protocols and the infrastructure needed for those protocols to operate [27]. OSLO builds upon the design principles of Linked Data. The term ‘Linked Data’

refers to data which is published on the Web and, apart from being machine-readable, it is also linked to other external datasets [1], using the Resource Description Framework⁵ (RDF) as a flexible and extendable data model.

The main goal of this article is to unfold the process of reaching semantic and technical interoperability among base registries based on the principles of Linked Data. The address registry is presented as a case study of a base registry. The European Commission defines a base registry (BR) as a trusted and authoritative source of information which can and should be digitally re-used by others, where a single organisation is responsible and accountable for the collection, use, updating and preservation of information. ‘Authoritative’ here means that a base registry is considered to be *the* ‘source’ of information, i.e. it shows the correct status, is up-to-date and is of the highest possible quality and integrity [27].

The Flemish Government administration aligns its base registries with this definition but introduces three additional requirements:

- The base registries are part of a semantic coherent system of uniform identified base-objects and relations, which are in line with the OSLO-standards.
- A base registry reuses the identifiers of the base-objects in other base registries.
- The base registries are obliged to maintain the Life-cycle and History of the base objects.

Public data often has a location-based component. According to Garson and Biggs [32] “It is estimated that 80% of the informational needs of local government policymakers are related to geographic location” (p. 87). The central address registry is one of the Flemish base registries and of significant value to the public and private sector. The address registry includes geographical coordinates [54] and is released under the Flemish Open Data licence⁶ which enables re-use, including commercial re-use, for free. Examples of re-use of this authoritative source in the private sector are a more accurate address-position in navigation applications than commercially available data and a better address quality in administrative processes than locally managed datasets resulting in a lower mail bounce and lower costs.

In 2013, a pilot project ‘Interconnecting Belgian National and Regional Address Data’ was carried out in the context of the ISA Programme of the European Commission. The pilot published data from the Belgian federal level and the three Belgian regions as Linked Data. Results of this pilot indicated that the public sector had not yet tapped into the full potential of its address registries. The obstacles: (i) address data fragmentation (ii) heterogeneous address data formats and (iii) a lack of common identifiers. The research reported in this paper reveals how these obstacles can be overcome so that the full potential of address registries and other base registries can be realised [16]. Interoperability in the public sector is influenced by internal as well as external politics. Internal, organisational politics includes dealing with issues involving organisational members. External organisational politics is about how public administrations relates to their council, board, or other organisations [48]. With the majority of the government data having a location-based component, it is important for

² <http://opendata.vlaanderen.be/>

³ <http://opendata.vlaanderen.be/dataset>

⁴ <https://www.cen.eu/work/ENdev/whatisEN/Pages/default.aspx>

⁵ <https://www.w3.org/2001/sw/wiki/RDF>

⁶ <https://overheid.vlaanderen.be/open-data-bij-de-vlaamse-overheid#modellenticies>

Government that all government administrations and partners link to the ‘authoritative’ addresses avoiding redundant and incomplete shadow-databases.

Information management specialists often lack knowledge about how to deal with political aspects of information management, and consequently are ineffective. The outcome of this paper is valuable for researchers, public administrations and public sector ICT service suppliers that aim to raise interoperability in complex data ecosystems.

Our **research questions** consider the problem statement from an, semantical, technical, and organisational point of view:

- How can public administrations raise semantic interoperability to ensure data from Base Registries can be interpreted by all communicating parties?
- How to maintain the agreed semantics on a technical level and design a programmable interface that can be well interpreted by both humans and machines?
- How to rewire existing processes and data products to become interoperable within an operational public sector context?

The **contributions** of this paper, which build upon the principles of the semantic web, are valuable for semantic web researchers and organisations that aim to publish interoperable authoritative data sources.

The remainder of this article is organised as follows: Section 2 discusses relevant related work: we discuss the nature of the Flemish, Dutch and Danish address registries. We illustrate the Linked Data strategy in Flanders and outline a comparative study between the current Universal Resource Identifiers (URIs) strategies in Europe. Section 3 gives an overview of the address vocabulary and discusses the ‘fitness for use’ of the prevailing European address vocabularies for modelling the address registry in Flanders. Section 4 outlines the key success factors of a real-time Linked Data architecture. We will point out how the addresses, mined from the municipalities, are published at the SPARQL-endpoint in ‘nearly real-time’. Finally, we will elaborate on the deployment strategy which should allow other agencies to reuse some of the technical components when

refactoring their own base registries. Section 5 discusses how the Linked Address Base Registry facilitates the adoption of Addresses as Linked Data in the public and private sector. Section 6 provides an evaluation of the Linked Data approach in the address registry. Section 7 presents conclusions on how to raise interoperability in the public sector and an outline for future work.

2. Background and related work

This section discusses the nature of the Flemish, Dutch and Danish address registries. By providing an insight into the events that influenced the evolution of these address registries, we can untangle the different interoperability levels. Next, we outline the Linked Data strategy in Flanders and describe the process and methodology we have used to create a reusable vocabulary for addresses. Finally, we provide an evaluation of the current Universal Resource Identifiers strategies in Europe which lead to the Flemish URI-strategy.

2.1. The Central Reference Address DataBase in Flanders (CRAB)

In 2011, the Flemish government and the Flemish Geographical Information Agency (now Flanders Information Agency), developed an authentic source for addresses, referred to as ‘Central Reference Address DataBase’ (CRAB), containing well over 4.5 million addresses as well as address positions (xy-coordinates). According to the European INSPIRE directive, the overall concept of an address data specification is that it has a “locator”, e.g. a house number that enables a user to distinguish it from neighbouring addresses; and a geographic position, which enables an application to locate the address spatially. To identify an address in Flanders, it must be associated with a number of “address components” represented by a spatial identifier. These components are defined in the CRAB decree as *streetname, house number and box number, postal code and municipality* [8, 10].

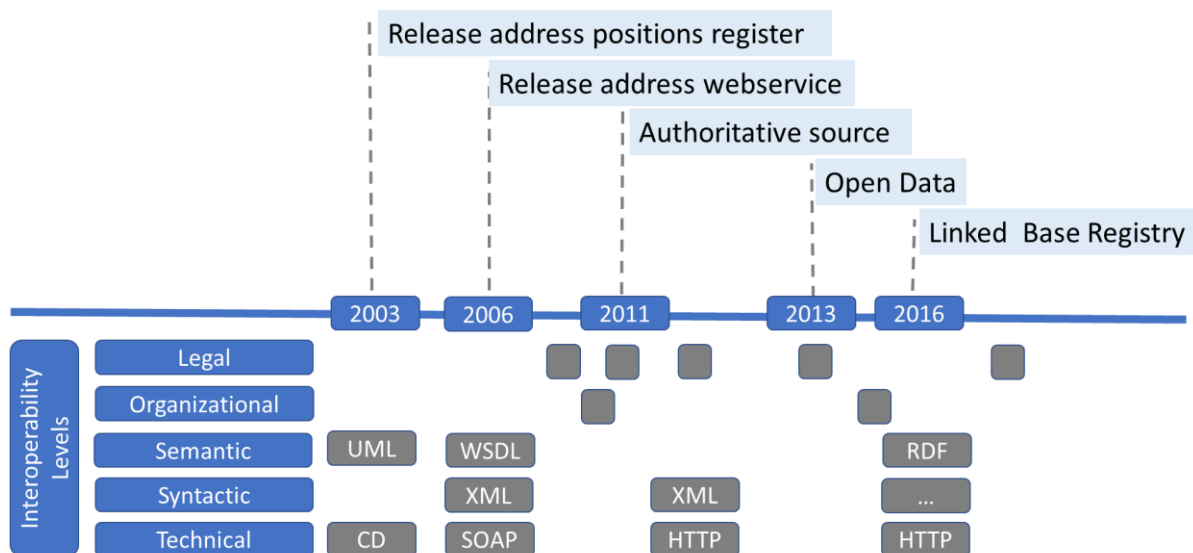


Fig. 2. Key influences on the address registry showing the impact on all interoperability levels, evaluated using the European Interoperability Framework [22] and inspired by an ISA evaluation framework in "How Linked Data is transforming eGovernment" [39].

Key influences

- The municipalities are responsible for the creation of addresses and management of the address components.
- Governments in Flanders are obliged to use the address registry and to provide feedback in case they detect an error.
- At the European level, the INSPIRE Directive aims to create a data infrastructure for the purposes of EU environmental policies and policies or activities which may have an impact on the environment [37]. By the end of 2017⁷ member states are obliged to provide harmonised address information by means of compatible services.

In the past two decades, many events influenced the development of the Central Reference Address Registry (CRAB) in Flanders. Figure 2 (Fig. 2) depicts the milestones evaluated using the European Interoperability Framework (EIF) [12, 15]. EIF is a set of recommendations which specify how administrations, businesses and citizens can communicate with each other across borders and within the EU. These interoperability levels are defined as *legal, organisational, semantic and technical* within a political context.

When the Flemish GIS administration (Ondersteunend Centrum GIS Vlaanderen, now Flanders Information Agency) was founded in 1995, there was no authoritative source for addresses. Although local governments have been responsible for the creation and maintenance of street names since 1977 [6], the lack of formal rules for addressing resulted in duplicate street names within the same municipality, misspelling of street names and many flavours of house numbers and box numbers to identify individual apartments in a building. The first initiatives towards a harmonised data model at regional and Federal level were initiated in 1999 (Geocodi). An important milestone was the release of the ‘address positions database’ in 2003, distributed to administrations in Flanders using a CD-rom as carrier, accompanied by a formal data specification in UML [47] ratified by a decision of the Flemish Government. In 2006 the address database became available via the Internet, by means of SOAP-webservices. A shared strategy for managing the lifecycle of addresses among the regional and local administrations (VLAR-address, 2008) paved the way towards an authoritative data source. The following year the INSPIRE directive from the European Commission [37] was converted into Flemish legislation [9]. The regulations are embedded into the CRAB [8, 10] and SDI [52] Decree, which are laws of the Flanders Region and the Flemish Community and set the scene for the legal interoperability level. The same year, an important hurdle on semantic interoperability was taken: a formal agreement on shared semantics was reached which eventually lead to shared semantics at Federal and regional levels in 2015. These events that intervened on all interoperability levels eventually lead to the formal approval of the address registry as the first authentic data source in the region of Flanders.

Driven by the Once Only principle embedded in Flemish registration [53] an authoritative source (also referred to as an ‘authentic’ source) builds upon following principles (i) administrations are obliged to use the authentic source, to avoid requesting information from citizens more than once, raise data consistency and reduce administrative burden, (ii) the

source is recognized by the Flemish government and (iii) administrations are obliged to report errors to the administrator of the source. In 2013, under an impulse of the European Directive on the re-use of public sector information [36] the CRAB Address registry became available as free Open Data. In 2014 The Flemish Government decided to focus on interoperable base registries [34], as defined in Section 1.

The Flemish government is managing the address registry conformant to best practices in all European Interoperability Framework-domains [25], more specific at the level of technical, semantic and organisational interoperability and conformant to legal requirements. At the organisational level, an important step was taken in 2015 by merging the GIS and e-Gov administrations into a new Information Agency. At the legal level, Flanders Information Agency has the ambition to embed the concept of Base Registries in a decree. The new agency launched the programme “OSLO” that focuses on the semantic interoperability level and extends the ISA CORE and INSPIRE Vocabularies in order to facilitate the integration of base registries with one another and their implementation in business processes of both the public and private sector. Finally, in 2016 the Flanders Information Agency rewired the Authentic Source for addresses and published it as Linked Open Data. The next step shall be taken in 2018, when the Federal and Regional administrations are obliged to use regional address registries (BeSt-Add⁸) in all their processes. The European Commission wanted to overcome the fragmentation of the address data. The aim was to tackle following hurdles: (i) address data fragmentation caused by the different isolated databases at the various government levels and the lack of a single access point, (ii) heterogeneous address specifications and standards, (iii) and a lack of common well-formed identifiers which are the cornerstone to link the different addresses (Colas et al., 2013). In 2011, the European Commission initiated a pilot project in Belgium, a federal state with three communities, three regions, and four language areas. The goal of the project was to interconnect the Belgian and Regional Address Data. The pilot was built upon the principles of Linked Data (Fig. 3) to make addresses more interoperable and lovable in Belgium. “The term Linked Data refers to a set of best practices for publishing and connecting structured data on the Web using international standards of the World Wide Web Consortium” [55]. The design principles⁹ as asserted by Tim Berners-Lee in 2006 (Fig. 3) were adopted by the address pilot project [16].

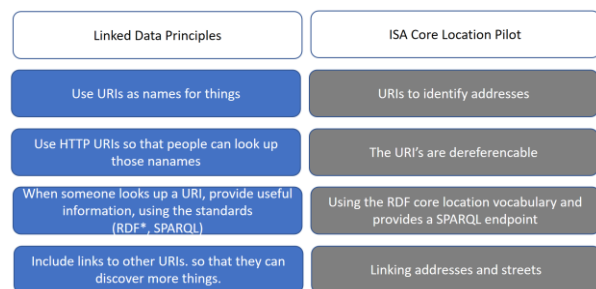


Fig. 3. Design principles of Linked Data and the specific implementation in the Core Locations Pilot, using the design principles as asserted by Tim Berners-Lee as a framework and mentioned in by ISA in "How Linked Data is transforming eGovernment" [39].

⁷ <http://inspire.ec.europa.eu/inspire-roadmap/61>

⁸ <https://overheid.vlaanderen.be/CRAB-Belgie-BeSt-Add>

⁹ <https://www.w3.org/DesignIssues/LinkedData.html>

The first principle states “Use URIs as names for things”. All addresses and streets were given a universally unique identifier which can be looked up via the web.

The second principle points to the “use HTTP URIs so that people can look up those names”. To create these stable identifiers best practices from the ISA programme for creating persistent URIs [3] were applied: including avoiding stating ownership in the URI, avoiding version numbers and implementing HTTP response code 303 to redirect from the real object to a document which describes the address or street.

The third principle: “When someone looks up a URI, provide useful information, using the standards RDF and SPARQL” is all about interoperability. RDF is the data model for Linked Data. According to W3C, “RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a “triple”). Using this simple model allows structured and semi-structured data to be mixed, exposed, and shared across different applications.”¹⁰ This pilot is using RDF as data model, the ISA core location vocabulary¹¹ provides the schema. The Core Location vocabulary is a context neutral, extensible data model derived from the INSPIRE address representation. We will discuss the differences between both vocabularies in section 3. This vocabulary was set up as a canonical data format, to bridge between the different non-interoperable address data models at the federal level and the three Belgian regions. The pilot implemented a SPARQL endpoint, which allowed querying the RDF data sources via the web. SPARQL is a recursive acronym for SPARQL Protocol and RDF Query Language.

The last principle “Include links to other URIs so that they can discover more things”. A dataset becomes more useful when it links to other resources, which allows humans and machines to discover more information following the links. Other parties on the web can link to the addresses, making their data more useful. The addresses core dataset itself does not link to other resources but enables other parties on the web to link to the addresses, making their data more useful. In Flanders, the domicile of a citizen and the Energy Performance Certificate (EPC) are both linked to an address.

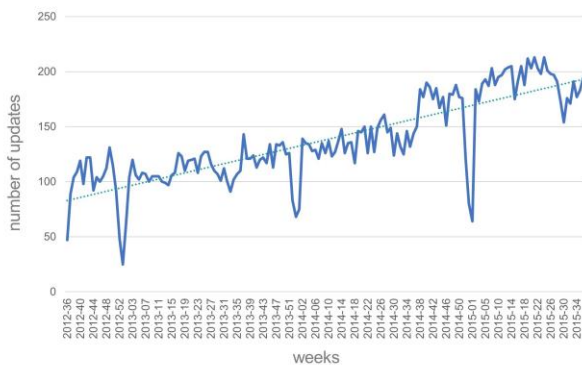


Fig. 4. Overview of the number of updates in the CRAB Address Registry each week (J.Laporte, Flanders Information Agency).

Because both resources are linked to the same URI, the government administration is able to give citizens insight in the energy efficiency of their domicile by following the links.

This pilot demonstrated the viability of a Linked Data infrastructure for addresses in Flanders, which can provide interoperability across the regions in Belgium. Despite this proof-of-concept paved the way towards a sustainable implementation for addresses in Flanders, some important complexities towards addresses still needed to be tackled, including the different numbering conventions.

As the CRAB Address registry is embedded into the core processes of government administrations and administrations are obliged to report errors, the number of updates is growing over 30% per year (**Fig. 4**). Figure 5 (**Fig. 5**) depicts the increasing use of the address registry, based on the transactions on the address Web Feature Service¹² (WFS).

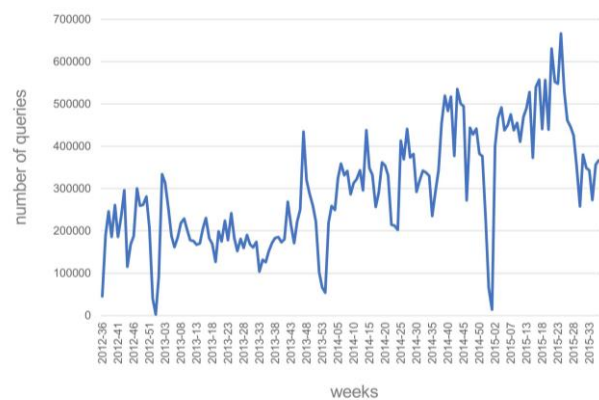


Fig. 5. the increasing use of the product based on registered users (J.Laporte, Flanders Information Agency).

2.2 The Address Registry in The Netherlands

The Netherlands are going through a similar process. The ‘Cadastré, the Dutch Land Registry and Mapping Agency’ (Kadaster) is at the helm of administrative and spatial data on addresses and buildings. The municipalities are responsible for the creation of addresses and management of the addresses components and have to register updates within four days [14] in the ‘Basic registry of Addresses and Buildings’ (BAG). The ‘Cadastré’ has shifted the focus from the address object to the physical objects that are addressable. These addressable objects can be ‘residence objects’ (dwellings or offices), mooring places or places for the permanent placement of mobile homes. The residence objects refer to buildings, individual apartments or offices within a large complex. These objects are registered in the base registry for addresses and buildings. The address is a property of a physical object and consists of following components as defined in the BAG object-catalogue [13]: address locator name, the thoroughfare name (street name, “naam openbare ruimte”) and the municipality. Since 2012, the address registry is published as Open Data. The Dutch Cadastré published the base registry with cadastral

¹⁰ <https://www.w3.org/2001/sw/wiki/RDF>

¹¹ https://joinup.ec.europa.eu/asset/core_location/description

¹² <http://www.opengeospatial.org/standards/wfs>

information as Linked Open Data in 2016. The address registry was published as Linked Open Data in 2017¹³.

Together with Geonovum, responsible for developing and maintaining standards for geographical information, administrations in the Netherlands joined forces to develop an Open Data strategy [31]. They have brought together experts from the public and private sector in the steering committee ‘Platform Linked Data Netherlands’¹⁴ (PILOD). The PILOD-platform has published an experimental version of the Dutch Building and Address registry as Linked Open Data¹⁵. A crucial step in the development was that all addresses and streets were given a persistent universally unique identifier which can be looked up via the web. They developed a similar URI strategy, both building upon best practices of ISA and INSPIRE [46]. Whereas the pilot in Flanders is extending the RDF using the ISA core location vocabulary as a basis, the Dutch pilot on addresses has developed an extension¹⁶ which is more in line with the existing BAG data model.

2.3 The Danish Address Registry

The Danish Address Programme is a sub-programme of the Basic Data Programme¹⁷ which is a collaboration between a number of National governmental bodies, the associations of Local and Regional Governments. The Danish Agency for Data Supply and Efficiency¹⁸ is the responsible authority for addresses in Denmark as well as for the Danish address registry, which was a part of Building and Dwelling Registry [17] but is now an independent registry¹⁹. The Danish administrations have focused on eliminating shadow databases by identifying base registries’ which they refer to as ‘GRUNDDATA’.

The Danish government has a long history in the standardisation of addresses which goes back to 1978 when a standard address structure was introduced [42] followed by the first address registry in 1980. The first address-structure consisted of a municipal code, street code, address number, floor and door identifiers. In 1990 the addresses were rewired from ‘Address as an Attribute’ in the different registers (such as the central person register and central business register), to an ‘Address as a common asset’ where the address becomes a fully-fledged object [42] which allows other registries to link to the Address Registry. These addresses were georeferenced and harmonised voluntarily by the local municipalities. This included a harmonisation of the ‘property data registers’ and the ‘municipality technical base maps’ towards a building and dwelling registry.

The 98 municipalities have the practical authority for the assignment of addresses (Danish Enterprise and Construction Authority, 2010). In 2000 the Danish Government appointed the building and dwelling register as a base register which was published in 2002 as Open Data. Although registries, data models and ministerial initiatives existed, there were no official guidelines for the infrastructure model [35]. In 2004, the report “Basisdata, forståelsesramme og analysemodel til kategorisering af basisdata” [40] was published, which focused on data quality and the

unambiguous connections between the registries. [35, 40]. Additional requirements were introduced [35, 42]:

- The base registries are part of a semantic coherent system of uniformly identified base objects and relations. The data model has similarities with the conceptual data models behind INSPIRE.
- Core Objects contain provenance information, which defines when the information was registered and for which period of time it is valid, often referred to as bitemporal data.
- A base registry should reuse the identifiers of the base objects of other base registries. The addresses have persistent identifiers, which are a globally unique identifier (GUID).
- The base registries are obliged to maintain the Life-cycle and History of the base objects. A telling example is the data about a building that is under construction or has been demolished.

2.4 The Linked Data strategy in Flanders

The Flemish government is focusing on a sustainable strategy for linked base registries. In the previous section, we have learned that it is essential to focus on all interoperability levels. In this section, we will focus on the semantic interoperability strategy and how it influenced the development of the Central Reference Address Registry (CRAB) in Flanders.

The OSLO-program, which started as a grassroots initiative at the level of the local governments²⁰, increased awareness on the need for semantic interoperability. OSLO created an ontology in three main domains of interest: (i) persons and organisations [15], (ii) locations, and (iii) public services²¹, in a setting where stakeholders are focusing on their similarities rather than on their differences [56]. This was achieved by implementing a process and methodology for developing semantic agreements, based on the ISA methodology [29]. In 2015 the ownership of OSLO and the responsibility for the governance and life-cycle management of the ontology was transferred to the regional government²², which started the follow-up project OSLO².

OSLO² provides a policy framework for technical topics, including the URI-strategy (Fig. 6), and domain-specific topics including a context neutral model for addresses.

The steering committee (represents the ‘authority’) agreed upon the specification process and the selection of the different working groups. Each thematic working group develops a context neutral vocabulary, by extending EU standards (ISA and INSPIRE) with specific local concepts that support the processes of the different governmental levels in Flanders. Working groups discuss relevant entities, relations and attributes, which are iteratively refined and formalized. At the end of the process the steering committee validates the specifications and the vocabulary. The working group consists of over 70 experts representing local governments, Flemish administrations, telecom providers, utility companies (water, energy), the real estate sector and non-profit organisations. To derive the requirement

¹³ <https://data.pdok.nl/datasets>

¹⁴ http://www.pilod.nl/wiki/Wie_we_zijn

¹⁵ http://www.pilod.nl/wiki/Bag_dataset

¹⁶ <http://lod.geodan.nl/vocab/bag/index.html>

¹⁷ <http://grunddata.dk/english/>

¹⁸ <http://sdfe.dk>

¹⁹ <http://danmarksadresser.dk/>

²⁰ <https://joinup.ec.europa.eu/community/semic/news/flemish-oslo-standard-become-local-extension-isa-core-vocabularies>

²¹ <https://purl.org/oslo>

²² <https://overheid.vlaanderen.be/OSLO-Wat-is-OSLO2>

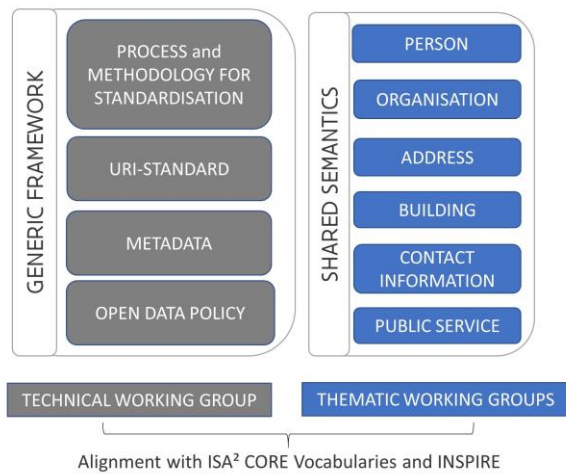


Fig. 6. Overview the OSLO² working groups [50].

for the high-level domain model, the participants start by specifying use-cases. In a next step the working group or a subgroup defines the attributes. The umbrella working group integrates the results into the global domain model. Finally, a conformance statement is created.

2.5 Overview of the Danish, Dutch and Flemish Address Registries

Comparing the strategy and development in Denmark with the Netherlands and Flanders **Table 1**. Overview of the Danish, Dutch and Flemish Address Registries shows that all initiatives focus on a semantic coherent system of uniform identified base objects and relations and that all the registries are available as Open Data. The Danish Government is adopting the principles of Linked Data²³ to raise semantic interoperability. At the time of writing, the Danish address registry was not published as Linked Data.

Table 1. Overview of the Danish, Dutch and Flemish Address Registries

	Danish	Dutch	Flemish
Open Data	YES	YES	YES
Reuse of vocabularies	INSPIRE	-	ISA and INSPIRE
Unique Identifier	GUID	URI	URI
Linked Data	NO	YES	YES

2.6 URI strategy in Flanders in relation to W3C, ISA and other EU member states

In this section, we will compare the different prevailing URI standards which influenced the Flemish strategy (**Fig. 7**). The OSLO URI working

group has developed a URI standard for persistent identifiers that supports government administrations in Flanders by providing guidance that ensures that HTTP URIs are future proof. By providing a standard, an accompanied guidelines document and individual guidance on request, Flanders Information Agency ensures a consistent URI structure for all publishers. The URI strategy is based on principles from W3C documents²⁴ ISA guidelines²⁵ and a comparison of the URI strategy in the Netherlands [45], the latter being inspired by the UK recommendation ‘Designing URI Sets for the UK Public Sector’ [18,19].

The W3C group ‘Cool URIs for the Semantic Web’ discusses the 303-redirect (a way to redirect web applications to a new URI) and hash-URI approach to identify real-world objects and the use of content negotiation but does not suggest or impose any kind of URI structure. Some examples do include ‘id’ and ‘doc’ that are also used in other URI strategies that are evaluated in the paper. The document does not suggest a specific approach regarding redirect versus hash-URIs, as the best approach is case dependent. Three approaches are described: (1) when using redirects: `http://www.example.com/id/alice` 303-redirects to `http://www.example.com/doc/alice`, (2) in case using fragment identifiers: `http://www.example.com/about#alice` is automatically truncated to `http://www.example.com/about`, the last approach (3) is combining both approaches, so that there is a 1 to 1 mapping between a real-life object and its describing page: `http://www.example.com/bob#this`.

As **Fig. 7** shows, the compared strategies mostly share the same approach. The most important differences are related to the interpretation of ‘type’ and ‘concept’. The ISA type can be one of following values: id or item for real-world objects; doc for documents that describe those objects; def for concept definitions; set for datasets; or a string specific to the context, such as ‘authority’²⁶ or ‘dcterms’.²⁷ The Dutch ‘type’ indicates the kind of the URI: id: identifier of a real-life object in a registry; doc: documentation about the real-life object by this registry and def: definition of a term in an ontology. The Dutch type is in line with the UK strategy, also using id, doc and def [18]. The OSLO type describes the nature of the referenced resource. It has to be chosen from the following list: id: when referencing a non-information resource; doc: document (that describes a non-information resource and, the target of a 303-redirect after resolving an id resource); ns: taxonomies, ontologies or vocabularies. Additionally, the “id” class does not demand the described resource to originate from a registry. Finally, the Flemish OSLO standard recommends that the 303-redirect approach is used (as is recommended by ISA)²⁸. Resources in the “ns” class are allowed to use fragment identifiers. The UK URI strategy refers to the W3C Cool URI rules when it comes to resolving URI’s and considers both options: fragment identifiers and 303-redirects.

²³ <https://arkitektur.digst.dk/model-rules-english>

²⁴ <http://www.w3.org/TR/cooluris/>

²⁵ <https://joinup.ec.europa.eu/community/semic/document/10-rules->

[persistent-uris](#)

²⁶ <http://publications.europa.eu/code/en/en-130100.htm>

²⁷ <http://dublincore.org/documents/dcmi-terms>

Interoperability Solutions for European Public Administrations (I)	United Kingdom (II)	The Netherlands (III)	Flanders, OSLO ² (IV)
<code>http://{domain}/{type}/{concept}/{reference}</code>	<code>http://{domain}/{type}/{concept}/{reference}</code>	<code>http://{domain}/{type}/{concept}/{reference}</code>	<code>http://{domain}/{type}/{concept}/{reference}</code> *
Domain: the combination of the host and the relevant sector.	Domain: based around the data.gov.uk domain, split by sectors as sub-domains.	Domain: a combination of an internet domain and a path.	Domain: a thematic subdomain is optional
Type: id or item for real world; doc for documents; def for concept definitions; set for datasets; a string specific to the context	Type: URI type (id for Identifier URI, doc for Document URI), Representation URI (def for Ontology URI, set for Set URI)	Type: id for identifier of a real world; doc for documents; def for definition of a term in an ontology	Type: describes the nature of the referenced resource; id: non-information resource; doc: document describing a non-information resource; nir: used instead of id and doc when using hash-based ids; list; raw; ns; meta
Concept: collectionID, the type of real world object or the name of the concept scheme.	Concept: A word or string to capture the essence of the real-world 'Thing' that the set names.	Concept: indicates the kind of thing the URI identifies. Usually it is a class name.	Concept: specifies the type of the resource. This can be interpreted using a simple trick: {resource} is a {category} {class}
Reference: refers to a specific item, term or concept.	Reference: A string that is used by the set publisher to identify an individual instance of concept. The reference should match the way that it is used in normal use.	Reference: usually a string which is used in the existing register already.	Reference: a repeating pattern that allows for hierarchical structure. Each element follows one of the following structures: {base-reference}; {base-reference}/{version}

Fig. 7. Comparing the URI strategy of ISA, United Kingdom, the Netherlands and Flanders. The recommended pattern is (a variation on) 'http://{domain}/{type}/{concept}/{reference}'. The domain is the combination of the host and the relevant sector. It is a matter of choice whether the sector is defined as a sub-domain of the host or as the first component of the path.

The ISA concept is a collectionID, the type of real-world object identified (e.g. a road), the name of the concept scheme (e.g. 'language'). Finally, the reference refers to a specific item, term or concept. The Dutch concept indicates the kind of thing the URI identifies, usually being a class name. Lastly, the {reference} is usually a string which is used in the existing registry already. When using dates to refer to versions, they advise using the W3C DateTime format. The OSLO concept specifies the category of the resource. This becomes part of the identifier as follows: {resource} is a {concept} {type}. For example 'https://data.vlaanderen.be/doc/adres/3706808'²⁸ can be read as: '.../id/adres/3706808' is an identifier for an address'.

ISA allows a collectionID to be used as well, while this is not allowed in the Flemish rules. We have detected no conflict with the standard in the Netherlands. The new part is referred to as the 'reference', a repeating pattern that allows for a hierarchical structure. Each single reference element follows either one of the following structures: (1) {base-reference} or (2) {base-reference}/{version}. The base reference allows for specification of a resource. It should be interpreted in the context of all preceding references. Finally, the version can represent a specific absolute date (in W3C format), a relative date (e.g., "latest") or a version (e.g., 1.2). Versions are only allowed where different versions of resources are allowed to co-exist. ISA provides little details about the specific format of the reference. There are no conflicts between ISA and the Flemish rules.

3. Core vocabularies for addresses

This section discusses the prevailing European ISA and INSPIRE address vocabularies and evaluate their 'fitness for use' for modelling the CRAB address registry. This "rewiring" of the CRAB happened as part of the development of a Building Registry which required an important revision of the address registry. Finally, we discuss the rewiring of the address model in depth. The resulting model bridges to other administrative domains like Persons, Businesses and Public Services.

3.1 INSPIRE Data Specification for the spatial data theme Addresses

Driven by 'the Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007' to establish an Infrastructure for Spatial Information in the European Community (INSPIRE), the 'INSPIRE Data Specification for the spatial data theme Addresses' was released in 2009. The data specification was developed by the Thematic Working Group on Addresses, in which the Flemish Agency for geographical information (now Flanders Information Agency) participated.

INSPIRE [38] defines an address as the "Location of properties based on address identifiers, usually by road name, house number, postal code" INSPIRE and serve different purposes including location, identification, jurisdiction, sorting and ordering and emergency response (e.g. CRAB is

²⁸ <https://data.vlaanderen.be/id/adres/3706808>

used by the emergency services). The INSPIRE address defines following address components [38]: administrative unit name (the name of a country, municipality or district), address area name (a non-administrative area or the name of a natural feature like a lake), thoroughfare name (a street name, the name of waterway) and postal descriptor. The vocabulary is documented using the Unified Modeling Language²⁹ and the structure and constraints of the XML representation are described using the XML Schema Definition Language³⁰.

3.2 ISA Core location vocabulary

The ISA core location vocabulary was developed following ‘the Directive Decision ‘COM/2008/0583 final - COD 2008/0185’ of the European Parliament and the Council on interoperability solutions for European public administrations (ISA).

According the European Commission in 2011, there were two main principles which defined the Core vocabularies [25]: (i) highly reusable: the specification is simple and captures basic and generic characteristics of an information entity, regardless of the context this entity is used and (ii) extensible: domain specific specializations can be drafted on top of the core representation.

The Core Vocabularies Working Group (Location Task Force) defined the Core Location Vocabulary in 2012 as a minimum set of classes and properties for describing a location represented as an address, a geographic name, or a geometry [23]. A set of commonly agreed Core Vocabularies supported by the EU Member States provides a concrete starting point for promoting semantic interoperability among European public administrations. The Core location vocabulary refers to the INSPIRE Data Specification for Addresses: the granular address properties (PO Box through to Post Code) are taken from the INSPIRE address guidelines. The vocabulary is documented both in a HTML³¹ and a Turtle³² specification that describes an RDF graph in a compact and natural text form³³. An address that is provided using these properties will be INSPIRE conformant. The additional property of full address is not part of the INSPIRE Address guidelines” [23].

3.3 A comparative survey of the ISA³⁴ and INSPIRE³⁵ address models

Although the ISA Core Location Vocabulary is driven by work on the INSPIRE directive [23], a mapping exercise revealed that INSPIRE and ISA have some significant differences. The ISA Core Location Vocabulary (Fig. 8) models its definition of “Address” on AddressRepresentation from INSPIRE. AddressRepresentation (Fig. 8) is a composite datatype and not an entity as the ISA-model suggests.

The fully-fledged INSPIRE address is intended to integrate the address life-cycle information with master-data management systems such as geographic information in support of Environmental Policy. The INSPIRE address-representation and the ISA address are light-weight representations of an address. They are intended to be used in situations

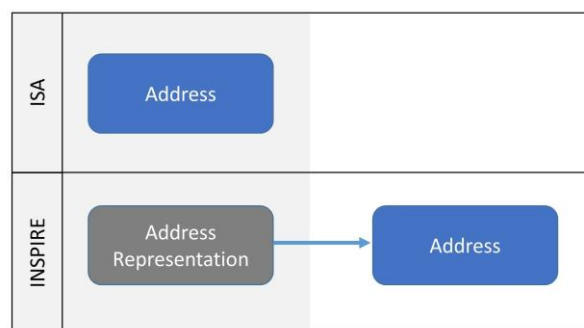


Fig. 8. Comparing the ISA and INSPIRE address model: Address and AddressRepresentation.

where the address is no more than an attribute of another entity, e.g., a Person. In the ISA Core Location Vocabulary, an Address is intended to be used as an attribute of Person:

`http://data.vlaanderen.be/ns/persoon#gezinsadres`. Moreover, it has no identifier of its own; there is only a pointer to the original structured address from which it was derived. Another problem is that ISA calls this entity/datatype “Address” and not “AddressRepresentation”. In doing so it created a schema-conflict with INSPIRE of the type “homonym” [26]. The impact of using the same name for different concepts is that user could expect the ISA-address to be a fully-fledged address instead of merely a lean representation without a unique identifier.

ISA added some attributes to INSPIRE’s AddressRepresentation. FullAddress is one of them and appears to be an implementation of the “ordered” constraint on INSPIRE: AddressRepresentation. Other attributes (PO-Box, AdminUnitL1-2) are specialisations of corresponding INSPIRE properties. As ISA limits itself to an address-as-an-attribute, certain properties of structured addresses are missing in the Core Location Vocabulary. Most noteworthy is address position. It can be argued that this property is also lacking in the INSPIRE:AddressRepresentation and that it is, therefore, permissible for ISA to leave it out of their definition. But as ISA renamed AddressRepresentation to Address one could ask if such an important aspect of structured addresses can be ignored. On the other hand ISA sports a Location entity/datatype which can have a geometry and address at the same time (if needed) and in that way an ISA:Address can be loosely coupled to a position. A relation with the parcel/building/building unit to which the address was assigned by the registering authority also lacks but is understandable as these addressable objects are currently not part of the Core Vocabularies.

On a more generic level, there are some additional differences to be mentioned, most important one being the definition of “identifier”. In the ISA Core an identifier is a composite datatype comprised of four attributes: the actual identifier or key and some metadata about the key, more specifically source, date and type. INSPIRE however leaves out the metadata and splits the key into its parts: a namespace, a localid and an optional versionid. Although namespace and source are somewhat

²⁹ <http://inspire.ec.europa.eu/data-model/approved/r4618-ir/html/index.htm?goto=2:1:1:7062>

³⁰ <http://inspire.ec.europa.eu/schemas/ad/3.0/Addresses.xsd>

³¹ <https://www.w3.org/ns/locn>

³² <https://www.w3.org/TR/turtle/>

³³ <https://www.w3.org/ns/locn.ttl>

³⁴ https://joinup.ec.europa.eu/asset/core_location/description

³⁵ <http://inspire-twg.jrc.ec.europa.eu/data-model/approved/r937/fc/>

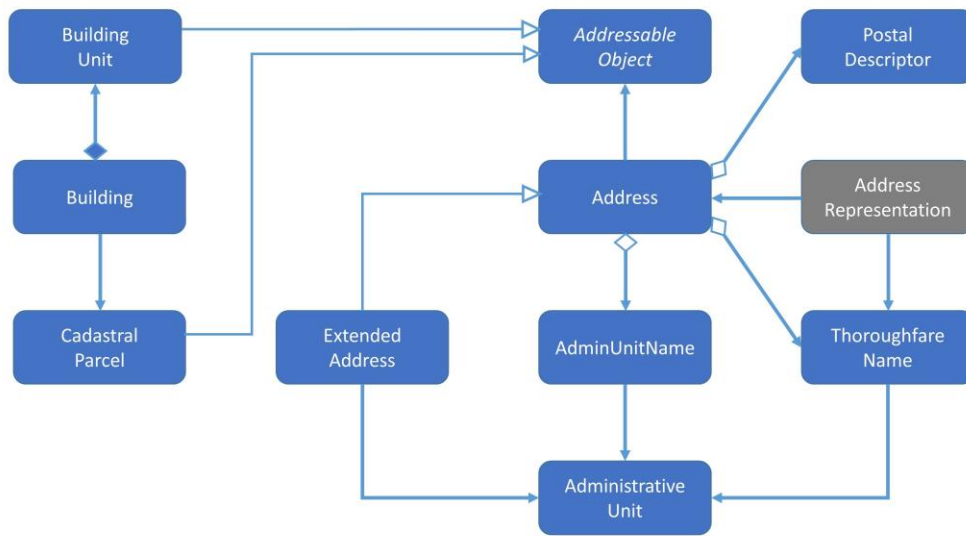


Fig. 9. OSLO² Conceptual Address Model.

overlapping concepts it seems that we have two different approaches to represent the identifier: one with metadata and an undivided key and one without metadata and a subdivided key.

The INSPIRE program is working on an RDF specification of the address data model. At the time of writing this article, the specification was still experimental³⁶.

3.4 OSLO address model

Based on the insights of section 3.3, we conclude that the INSPIRE Address specification is better adapted to the needs in the fields of all EIF levels, while the ISA Core vocabularies provide a crucial integration with other administrative domains like Persons, Businesses and Public Services. In order to achieve a feasible model, and to benefit from the best of both worlds, the Flemish Government has adapted the ISA address towards INSPIRE. The vocabulary is documented both in a HTML <https://data.vlaanderen.be/ns/adres> and a Turtle specification <https://data.vlaanderen.be/ns/adres.ttl> Most distinguishing feature of this rewiring was the instantiating of addresses as objects like BAG and OSLO, as discussed in section 2.2. In the past only the address components were registered as separate objects. Part of this effort was a tighter alignment of the conceptual model to the INSPIRE:Address specification. The INSPIRE model is very generic, designed to be applicable to all address definitions that can possibly exist in Europe. So part of the alignment actually boiled down to mapping the more generic INSPIRE terms to existing terms in the Belgian context: <http://data.vlaanderen.be/ns/adres#Adres> (e.g. a Belgian housenumber is actually a LocatorDesignator). Eventually, the resulting model was not only adopted on a regional level but also on a national level where it was published as BeSt-Add³⁷ as a result of a joint effort between the three regions and important federal institutions like the National

Registry (of Persons), the postal services, the cadastral agency, the National Geographical Institute amongst others amongst others.

Apart from making the generic aspects from the INSPIRE specification more concrete, the most difficult part of deriving a Core Vocabulary for addresses was modelling the clear distinction between the official or registered elements of an address and other more optional properties like sitemames or flooridentifiers. These were moved to a subclass AddressExtension: <http://data.vlaanderen.be/ns/adres#Adresuitbreiding>

Since the OSLO Address Vocabulary (Fig. 9) is to be considered as a local implementation of the ISA and INSPIRE specification we aligned on the Belgian definition of Address: BeSt-Add.

However, to accommodate for foreign addresses we reused and extended the Address definition of <http://www.w3.org/ns/locn#Address> to denote an AddressRepresentation. The locn vocabulary was designed for capturing all kinds of addresses within the world. To keep the link with the Belgian addresses in structured form, a relationship from the addressrepresentation to the BelgianAddress is included.

As we mentioned before we discovered that ISA and INSPIRE support different definitions of “identifier”. To accommodate both approaches we defined two subclasses for the key: simple key and composite key. We then added the metadata attributes from ISA. The concept of alternative identifier was incorporated since more than one identifier can be exchanged. For other generic elements like “geometry” <http://data.vlaanderen.be/ns/adres#positie> we aligned the OSLO Address Vocabulary to INSPIRE, although the requirement to use GML to serialise the geometry was replaced with a more generic construct also allowing well-known text representation of coordinate reference systems³⁸ (WKT) and such. The ISA Core vocabularies nor the INSPIRE specifications mention much about metadata on an object-level. ISA explicitly leaves out metadata properties and INSPIRE mentions it only on a dataset-level although the available lifecycleInfo properties are in fact metadata.

³⁶ <http://inspire-eu-rdf.github.io/inspire-rdf-guidelines/>

³⁷ <https://overheid.vlaanderen.be/CRAB-Belgie-BeSt-Add>

³⁸ <http://docs.opengeospatial.org/is/12-063r5/12-063r5.html>

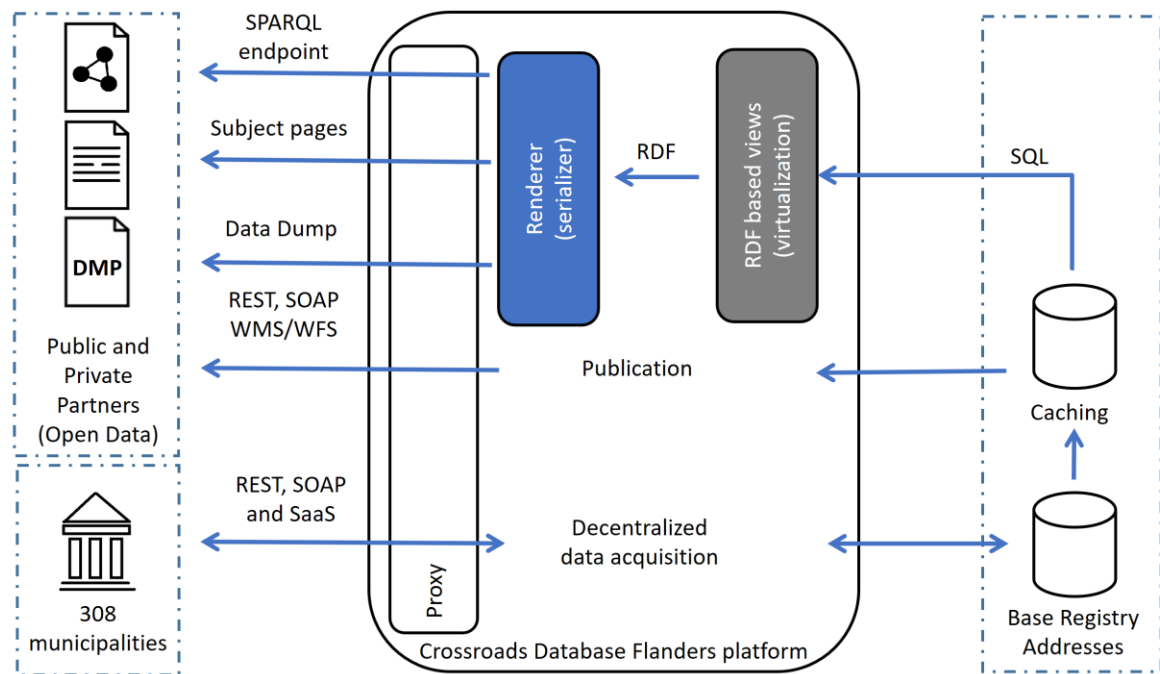


Fig. 10. Crossroads Database Flanders Architecture.

The OSLO Address Base Registry Vocabulary 39 is intended to exchange addresses. The application profile is a more elaborated vocabulary with additional restrictions (such as cardinality and code lists to be used) intended for a specific application. There is one application profile for addresses defined: namely the one for address registries. Examples of additional restrictions to the vocabulary are: the municipality name and street name are obliged, and all code lists must be aligned on INSPIRE.

The address representation, which is an Address defined in <https://www.w3.org/ns/locn>, is extended with properties that are specific for Belgian Addresses, such as a ‘busnummer’, which is Dutch for ‘letterbox’, that identifies a dwelling and is identified by <http://data.vlaanderen.be/ns/adres#Adresvoorstelling.busnummer>. The ‘busnummer’ is a specialisation of the W3C locator designator <http://www.w3.org/ns/locn#locatorDesignator> that consists of number or a sequence of characters that uniquely identifies the locator within the relevant scope. We define the ‘busnummer’ property to have a domain of <http://www.w3.org/ns/locn#Address> and a range of <http://www.w3.org/2001/XMLSchema#string>. This can be compared to object-oriented programming which commonly defines a class address with an attribute called ‘busnummer’ of the type String. The Range defines that the value of a “busnummer” must be a String as defined by <http://www.w3.org/2001/XMLSchema#string>.

The Model of the Flemish Address Register (<https://data.vlaanderen.be/ns/generiek>) applies the W3C PROV (<https://www.w3.org/TR/prov-o/>) ontology for modeling provenance. This structured metadata records the origin of addresses including when, why,

and who created or modified the information. This allows users of the address register to evaluate if the information can be trusted and to integrate the provenance information with other information sources.

4. Design principles of Linked Data applied to the Address Base Registry

In 2008, the regional Government of Flanders initiated the: “MAGDA” data platform, which supports organisations to publish and access public sector information. Due to the need of providing more integrated, interconnected and cross-sectoral services [39], Flanders Information Agency started-up a pilot project and applied the principles of Linked Data because they expected them to increase interoperability. In 2018, the MAGDA-platform was rebranded to Crossroads Database Flanders (CDF).

This section explains how the “CDF-platform” (Fig. 10) was enabled to publish Linked Data on the web. We will discuss the reference components: the proxy server serving as the entry point of the Linked Data Infrastructure, the renderer which serialises the RDF or creates a human-readable HTML subject page and the RDF store. We will outline how the addresses, mined from the local communities, are published at the SPARQL-endpoint and HTTP services in ‘nearly real-time’. Finally we will elaborate on the deployment strategy which allows other agencies to either reuse the complete setup as is, or recombine the components to another setup with minimal effort.

³⁹ <https://data.vlaanderen.be/id/applicatieprofiel/adresregister>

4.1 Crossroads Database Flanders-platform

In Flanders, the administration responsible for the ‘base registries’ hosts a single point of service delivery via the data portal “CDF” which allows citizens and businesses to access all government supplied services, regardless of which authority or channel provides them. At this moment 308 local governments, 978 public partners, over 2000 private partners and 1 out of 3 citizens are connected to the secure platform performing millions of requests on information objects about citizens, businesses, addresses, buildings and their locations [44]. The first product released on the “CDF-platform” according to the Linked Data principles is the Central Reference Address Database, containing over 3 million addresses and their geographical coordinates. The addresses are synchronised in real time between 308 local governments and the Linked Base registry.

The realised Linked Data Infrastructure for CRAB applies traditional, known, Linked Data publishing techniques. The spillover effect is the adoption of the URIs and vocabularies in traditional services including geographical Web Feature Services and SOAP Services, which facilitate interoperability between the different endpoints.

The solution is not dedicated towards solely publishing CRAB, but it is conceived with the ambition to support further exploitation beyond the owners of the CRAB registry. Other governmental bodies within the Flanders region are welcomed to reuse the

complete solution or components from the solution. By embracing this ambition in the design phase the future buy-in can be achieved. This ambition reflects in the chosen technologies that instantiate the solution. For each component commercial as well as open-source options are possible. An overview of the implementation details is published on the GitHub of the Flemish Government Administration: <https://github.com/Informatievlaanderen/Data.Vlaanderen.be/tree/master/documentation>.

4.2 Proxy-server

The proxy server is the entry point of the Linked Data Infrastructure. It implements the desired content negotiation. Information Flanders created a reference implementation that satisfies the URI strategy and the implementation guidelines of Information Flanders as a reusable component.

In addition to the effort gain, more coherency on the supported content negotiation is achieved. It is clear which headers are to be supported with which values. For instance the reference implementation supports the content types ‘application/rdf+xml’, ‘text/turtle’ and ‘text/html’. The implementation of the HTML subject page renderer is based on the mu.semte.ch⁴⁰ ecosystem. The reference implementation will provide answers to a special situation such as when the accept header is empty. Using content negotiation, the proxy will forward each valid URI to a subject page in human readable format (HTML) or to a machine-readable document (RDF in different serialisations). Human readable subject pages exist to provide useful information about the resource and to build a trust relationship with the user of the URI. The existence of a page that contains human interpretable data is key for the adoption of the URI as a trustworthy

identifier, in particular for users who are less familiar with Linked Data. It is intended that the contents of the human readable and the machine-readable formats are the same. But for reasons of easing the online exploration of the data, there might be slight differences to support better exploration of the data. For instance, data sources usually physically store relationships only in one direction.

For instance, a streetname belongs to a municipality. When exploring, the data users should not be restricted to follow only the directionality stored in the data source, but any direction should be offered. The provided reusable component for creating human-readable subject pages generates HTML (Fig. 11) according to the Flemish government style guide⁴¹ and is easy configurable to create alternative views based on the provided RDF data. For the machine-readable subject pages two alternative solutions are being provided: users can perform direct queries on the RDF SPARQL endpoint or they can retrieve stored content on disk often referred to as a caching.

Gegevens	
label	adres 3706808
versie	http://data.vlaanderen.be/doc/adres/3706808/v14
huisnummer	70
busnummer	
officieel toegekend	✓
straatnaam	Koningin Maria Hendrikaplein
gemeentenaam	Gent
postinfo	9000
geometrie	geografische positie adres 3706808
status	InGebruik

Fig. 11. Human readable representation of an address subject page: <http://data.vlaanderen.be/id/adres/3706808/>

4.3 The SPARQL endpoint and API

Public and private partners can query the Base registry using a service which accepts SPARQL queries. SPARQL⁴² is a query language for the RDF, similar to the Structured Query Language (SQL) for relational data. We use the HTTP message header (e.g. ‘application/rdf+xml’ or ‘text/html’) which indicates the request format by the client application to provide the matching representation, often referred to as content negotiation. The renderer serialises the RDF or creates a human-readable HTML subject page. The RDF store is implemented as a virtual triple store:

⁴⁰ <https://mu.semte.ch/>

⁴¹ <http://www.vlaanderen.be/nl/vlaamse-overheid/webuniversum>

⁴² <https://www.w3.org/TR/rdf-sparql-query/>

the Virtuoso⁴³ transformation engine transforms the data from the ‘Base Registry Cache’ (SQL) on the fly to RDF triples. The advantage is that the address updates from the municipalities, are published at the SPARQL-endpoint in ‘nearly real-time’, accessible via <http://data.vlaanderen.be/sparql>⁴⁴.

The API⁴⁵ of the Address Registry⁴⁶, which provides programmatic read and write access to Base Registries for Flanders is a classical XML-implementation for reasons of backwards compatibility, the semantics are in line with the RDF data model, and the service is using dereferenceable URIs as primary identifiers.

4.4 Deployment strategy

The Address Registry is deployed on a cloud infrastructure: Microsoft Azure⁴⁷. All the components are available as runnable services using Docker. Docker is “... an open-source engine that automates the deployment of applications into containers” [49]. This deployment strategy fits exactly the ambitions to provide a setup that is reusable by other governmental agencies. The infrastructure is setup using Terraform⁴⁸. The Terraform configuration describes the Docker Swarm⁴⁹ setup, that is responsible for executing and monitoring of the services. One can opt to run the complete setup as is, but one can also recombine the components to another setup with minimal effort. For instance, replacing the ONTOP component with a commercial edition simply means replacing the docker reference of the ONTOP solution with the docker providing the commercial edition. The Docker layer creates a transparent system architecture with well-documented flexibility points. In addition the Docker layer makes the provided solution independent of the hardware used: it can be run on a Container As A Service solution offered by a cloud provider such as Azure, but it can also be run on a single machine or a developer machine.

5. Applications of the Linked Address Base Registry

5.1 Adoption of Addresses as linked data in the private sector

An example of linked addresses as an interoperability facilitator is the case of Postbuzz⁵⁰ and Tenforce. This private initiative lets citizens and businesses discover what is buzzing in their neighbourhood including hyperlocal news. The location of the citizen, the businesses, the news and events play a crucial role in the location-based services. Tenforce has developed a new service which builds upon the Generic Information Platform for the Public Domain⁵¹ (GIPOD) and gathers all information concerning works or manifestations in the public domain. The GIPOD dataset contains URIs which identify addresses impacted by roadworks or manifestations (ongoing). Postbuzz provides personalized services by notifying businesses if their location cannot be reached, including detailed status information such as the start and end date of the manifestation. The personalized notifications are visualized as tiles (Fig. 12).

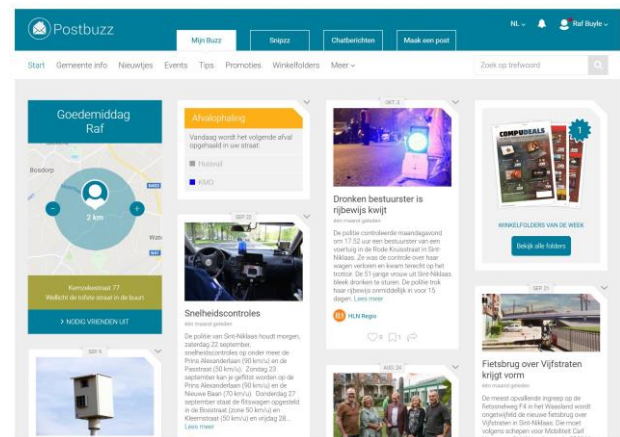


Fig. 12. Postbuzz lets citizens and businesses discover what is buzzing in their neighbourhood, including hyperlocal news.

Linked Address Data has the ability to combine data from the businesses and their addresses in the Postbuzz knowledge graph using the SPARQL endpoint with the impacted locations in GIPOD. When re-using public sector information, until now Postbuzz had to look-up the coordinates of the provided addresses. In the past, the city of Ghent in Belgium annotated their news with the local address identifier of the Flemish government. Because of a lack of context, the developers of Postbuzz missed out this valuable clue. Linked address data provides context information via the dereferenceable address identifiers (URIs) including a reference to the vocabulary and links to other useful information such as other addresses in the same street.

5.2 Better adoption of Addresses within the public sector

The Flemish government administration applied the Address Registry in a high impact project “citizens portal” that will allow 6.4 Million citizens to access public services. The citizens profile enables local and regional government administrations to integrate different public services, regardless of the channel they use. The starting point is the citizen that has a secure login on the portal, identified by its electronic identity. The address identifier is key for several administrative processes: birth, domicile, place of residence, the death of the citizen and its relatives. Location-based information has rarely been used in transactional public services in Flanders because a lack of vocabularies that bridge between alphanumeric and geographical information. The RDF data model for addresses may facilitate the integration of addresses in several business domain services on a semantic level. While developing the citizens-portal, we noticed that

⁴³ <https://virtuoso.openlinksw.com/>

⁴⁴ <http://data.vlaanderen.be/sparql>

⁴⁵ <https://beta.basisregisters.vlaanderen.be/api/v1/adressen/200039>

⁴⁶ <http://beta.basisregisters.vlaanderen.be/Help>

⁴⁷ <https://azure.microsoft.com>

⁴⁸ <https://www.terraform.io/>

⁴⁹ <https://github.com/docker/swarm>

⁵⁰ <https://www.postbuzz.com>

⁵¹ <https://overheid.vlaanderen.be/producten-diensten/generiek-informatieplatform-openbaar-domein-gipod>

developers are familiar with JSON, which has a very simple syntax but no inherent tied semantics. Developers want a simple, extensible way to create an API that gets the job done and does not design them into a corner⁵². JSON-LD bridges between JSON and the formal OSLO data specification. The semantics of the address properties in this business-oriented services are a subset of the address in the base registry. By adding a dereferenceable URI, the analyst or developer can discover and use additional data which is not provided in the API. The data can be retrieved by dereferencing the URI and selecting a representation through content negotiation. The semantic agreements reached at the business level of the ‘citizens portal’ project, are modelled in the Unified Modeling Language (UML). The UML is automatically⁵³ transformed into an RDF model and linked to other vocabularies including addresses. The formal specification is then published at data.vlaanderen.be⁵⁴, including a JSON-LD context⁵⁵ which allows embedding the semantic agreements in business-oriented services. To preserve the “context” within these services, the shortcut terms in the service are mapped to the terms in the data.vlaanderen.be⁵⁶ RDF-vocabularies. This principle is known as “expanded term definition” and is accomplished by using a JSON-LD context⁵⁷. The term ‘verblijfsadres’ (place of residence) is mapped to the vocabulary ‘<http://data.vlaanderen.be/ns/persoon#Verblijfplaats>’ (Fig. 13). These JSON documents can then be interpreted⁵⁸ as Linked Data.

```

{
  "@id": "https://data.vlaanderen.be/id/persoon/00000014182",
  "@type": [ "GeregistreerdPersoon", "Inwoner" ],
  "heeftVerblijfplaats": [
    {
      "@type": "Domicilie",
      "verblijfsadres": {
        "straatnaam": {
          "@id": "http://data.vlaanderen.be/id/straatnaam/7003",
          "naam": "Lubbeekstraat"
        },
        "huisnummer": "20",
        "gemeentenaam": {
          "@id": "http://data.vlaanderen.be/id/gemeentenaam/3210",
          "naam": "Lubbeek"
        }
      },
      "land": {
        "@id": "http://publications.europa.eu/resource/authority/country/BEL",
        "naam": "Belgie"
      }
    }
  ]
}

```

Fig. 13. JSON-LD example used in the ‘citizens portal’ which links the citizens domicile to an address in the central reference address database.

6. Discussion

By focusing on Linked Data, the Flemish Information Agency aims to increase adoption of base registries by private and public partners via the presence on the web. As a spillover, this architectural approach has important benefits for the internal organisation. In this section, we will reflect on the added value of Linked Data on addresses and on how OSLO fosters semantic interoperability. We use the design principles as asserted by Tim Berners-Lee (Fig. 3) and an evaluation framework from the

European Commission [43] to reflect on the implementation of Linked Data on addresses in Flanders.

6.1 Understandability

We have learned that the formal structure of addresses is difficult to understand. A telling case is Postbuzz, that missed out the context of addresses due to a lack of context. Linked address data provides context information via the dereferenceable address identifiers (URIs) including a reference to the vocabulary and links to other useful information such as other addresses in the same street. In addition, it is crucial that the vocabulary is supported by the community of stakeholders. OSLO created a setting where the stakeholders from the thematic working group ‘addresses and buildings’ focused on their commonalities rather than on their differences. In an early stage, consensus building and a meet-in-the-middle approach are essential for a broad support of a semantic standard [15]. A strong focus on linked data and on reusable ontologies has important benefits for the internal information household.

6.2 Scattered Information

As 80% of the informational needs are related to geographic location [32] and spatial data often uses domain specific standards, address information is scattered on different information systems. By invoking the URI-standard, we have created universally unique and stable identifiers for addresses (first design principle). This allows linking addresses more easily with other datasets. Until now, the CRAB identifier was only unique on a system level, and implicit at the Flemish level. According to W3C Library Linked Data Incubator Group, working towards a Linked Data architecture, "can help organisations improve their internal data curation processes and maintain better links between, for instance, digitised objects and their descriptions... even where data is not entirely open" [5]. Where an address in relation to a person to date used a different definition than the authentic source for addresses, extending and adopting the ISA Core vocabularies enabled the Flemish Government to create a link between ‘persons’ and ‘addresses’ in an unambiguous way [23]. On top of the reuse of the vocabulary, persons can be linked to the authoritative address via a URI instead of duplicating the data.

6.3 Usability and Interoperability

Because it is difficult to link or integrate location based information, the Flemish government has created a architecture (second principle) which allows retrieving the address information via the web. This will allow users to link directly to an address using a lightweight HTTP service instead of duplicating the data in their information system. We expect this will avoid shadow databases, containing redundant and outdated information. OSLO is aligning the vocabularies of the different base registries, which are part

⁵² <https://json-ld.org/spec/latest/json-ld-api-best-practices/>

⁵⁶ <http://data.vlaanderen.be/ns/>

⁵⁷ <https://www.w3.org/TR/json-ld-syntax/#the-context>

⁵⁸ <https://www.w3.org/TR/json-ld-syntax/#interpreting-json-as-json-ld>

of a semantic coherent system. This allows integrating address information more easily in the various business processes.

6.4 Different user needs

By using content negotiation mechanisms it becomes possible to serve different representations of a resource from the same URI. This allows clients to specify which version best fits their needs. The renderer serialises the RDF and creates both a human-readable HTML subject page and various machine-readable formats, including ‘application/rdf+xml’ and ‘text/turtle’.

6.5 Machine readability and reasoning

When looking up an address via the web, we provide useful machine-readable information (third principle), using RDF as a data model. We have extended the RDF using the European ISA and CORE vocabularies. This enables addresses to become ‘self-describing’, which allows applications that are not familiar with the Flemish context to dereference the URI and find the definition [12]. Reuse of vocabularies can also lower the integration cost, particularly in case of reference data [2]. In addition, we have included links (fourth principle) to other URIs, so users can follow the links from the address to the street. Other parties on the web can link to the addresses, making their data more useful. By adding a dereferenceable URI, we provide context to analysts or developers who can discover and use additional data which is not provided in the API.

6.6 Future work

The Flemish government has invested in different Linked Data distribution strategies that are aligned with the needs of their clients. A single-file data dump has a low server-side complexity but does not allow live querying on the web. This introduces high costs for the clients. The SPARQL endpoint allows flexible live querying, but its availability is problematic [51]. Therefore the Flemish Government administration will explore the possibilities of Linked Data Fragments⁵⁹, a REST(ful) publishing strategy that allows efficient offloading of query execution from servers to clients through a lightweight partitioning strategy [51]. Representational State Transfer (REST) style, outlines how to construct network-based software applications having the same characteristics as the Web: simplicity, evolvability, and performance [30]. Figure 14 (Fig. 14) offers a uniform view⁶⁰ on the different HTTP interfaces for Linked Data in relation to the server and client effort based on the vision Linked Data Fragments [30].

Another obvious extension to this research is the publication of Linked Open Data which is linked to information under privacy regulations. Examples include questions such as: “is the Uniform Resource Locator referring to an object under privacy regulations also privacy sensitive information and how to cope with these challenges within an operational context?”.

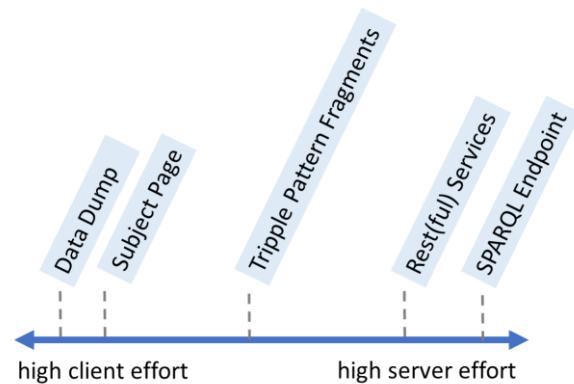


Fig. 14. An overview of the different HTTP interfaces for Linked Data in relation to the server and client effort [30].

7. Conclusions

In this paper, we presented insights on the implementation of a Linked Base Registry for addresses by unfolding the process followed towards raising semantic interoperability based on Linked Data principles. The Flemish Government administration aligned the base registry for Addresses with the design principles of Linked Data, because they expected this to increase interoperability. The approach goes beyond Linked Data initiatives which often take abstraction of the interoperability levels: namely on the *legal, organisational, semantic and technical level* [22]. The private sector “Postbuzz” initiative and the high impact project “citizens profile” were the first to adopt the Linked Base Registry for addresses. They indicate that Linked Data can indeed increase semantic and technical interoperability and can lead to the adoption of addresses in the public and private sector. For reasons of backward compatibility, the classical XML-webservices will be maintained. A spill-over effect of the Linked Data distribution is that the semantics of the XML-services were brought in line with the RDF data model and now have dereferenceable URIs as primary identifiers. While implementing the address vocabulary, we stumbled on competing international semantic standards and difficult choices on how to extend them to fit the local context. Therefore it is crucial to have a governance structure for making and institutionalising pivotal decisions. This can be realised through a policy framework for technical as well as domain-specific topics, comparable to the OSLO programme in Flanders. This paper identifies significant benefits in adopting the principles of Linked Data regarding base registries, not only by providing interoperability towards external stakeholders but also by fostering a more open architecture within the administration. A good example is the ability to unequivocally link ‘addresses’ to other business objects, averting the creation of ‘shadow databases’. We expect the insights from the linked address registry reported in this article to speed-up the process in other administrations that face the same complexity of publishing linked base registries.

⁵⁹ <http://linkeddatafragments.org/>

⁶⁰ <http://linkeddatafragments.org/concept/>

Acknowledgements

This paper is related to the projects ‘OSLO and CRAB as Linked Open Data’ and the ‘Flemish Building Registry’, funded by Flanders Information Agency and the program Flanders Radically Digital. The research activities in this article were funded by IMEC Ghent University. The authors would like to thank the teams of the projects ‘OSLO and CRAB as Linked Open Data’ and the ‘Flemish Building Registry’. We also would like to thank Per de Place Bjørn from the Danish Agency for Digitisation, Paul Janssens from Geonovum, Ruben Verborgh and Pieter Colpaert from imec IDLab - Ghent University for their valuable insights.

References

- [1] Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), pp. 399-418. <https://doi.org/10.1016/j.giq.2015.07.006>
- [2] Archer, P., Dekkers, M., Goedertier, S., Hazard, N., Loutas, N. (2013). Business models for Linked Open Government Data: what lies beneath?, <https://www.w3.org/2013/share-psi/workshop/krems/papers/LinkedOpenGovernmentDataBusinessModel.pdf>
- [3] Archer, P., Goedertier, S., Loutas, N. (2012). D7.1.3 - Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC, ISA Programme. <https://joinup.ec.europa.eu/sites/default/files/document/2013-02/D7.1.3%20-%20Study%20on%20persistent%20URIs.pdf>
- [4] Audit Vlaanderen. (2016). Thema-audit gemeentelijke belastingen en retributies. p. 81. http://www.auditvlaanderen.be/sites/default/files/atoms/files/Global_rapport_thema_audit_gemeentelijke_belastingen_en_retributies.pdf
- [5] Baker, T. et al. (2011). Library linked data incubator group final report. report, W3C Incubator Group. P.25. <https://www.w3.org/2005/Incubator/ldd/XGR-ldd-20111025/BelgischStaatsblad.pdf>
- [6] Belgisch Staatsblad. (1977). Decreet tot bescherming van de namen van de openbare wegen en pleinen, B.S. 07.04.77, http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=nl&la=N&cn=1977012832&table_name=wet
- [7] Belgisch Staatsblad. (2007). Decree on the re-use of public sector information of 27/04/2007, B.S. 19.04.07. pp.56250-56256 http://ec.europa.eu/information_society/newsroom/cf/dae/document.tcfm?doc_id=1701
- [8] Belgisch Staatsblad. (2009). Decreet betreffende het Centraal Referentieadressenbestand van 8 mei 2009, B.S.01.07.2009, https://www.agiv.be/~media/agiv/producten/crab/documenten/crab-decreet%20in%20belgisch%20staatsblad_090701.pdf
- [9] Belgisch Staatsblad. (2011). Wet houdende instemming met het samenwerkingsakkoord van 2 april 2010 tussen de Federale Staat, het Vlaamse Gewest, het Waalse Gewest en het Brussels Hoofdstedelijk Gewest voor de coördinatie van een infrastructuur voor ruimtelijke informatie. B.S.26.04.2011, http://www.geopunt.be/~media/geopunt/voor-experts/documenten/afederale%20instemmingswet_tcm307-240907.pdf
- [10] Belgisch Staatsblad. (2012). Decreet houdende wijziging van het CRAB-decreet van 8 mei 2009, B.S.12.06.12. <http://reflex.raadvst-consetat.be/reflex/pdf/Mbbs/2012/06/12/121565.pdf>
- [11] Belgisch Staatsblad. (2015). Decreet tot wijziging van het decreet van 27 april 2007 betreffende het hergebruik van overheidsinformatie en het decreet van 18 juli 2008 betreffende het elektronische bestuurlijke gegevensverkeer, B.S. 30.06.15, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=12909
- [12] Bizer, C, Heath, T., Berners-Lee, T. (2009). Linked data-the story so far. *Semantic Services, Interoperability and Web Applications: Emerging Concepts*. pp.205-227. <https://eprints.soton.ac.uk/271285/1/bizer-heath-berners-lee-ijswis-linked-data.pdf>
- [13] Bronhouders Afnemers Overleg. (2016). Objectenhandboek BAG. pp.6-12. <http://www.kadaster.nl/documents/32706/37743/bag+objectenhandboek+2009+inclusief+besluiten+zorg-en+studentencomplexen/389eb2fe-1c0f-450e-a56e-cac1b88d4111>
- [14] Bronhouders Afnemers Overleg. (2013). Processenhandboek BAG, basisregistraties adressen en gebouwen. p.13. <http://www.kadaster.nl/documents/32706/37743/bag+processenhandboek+2013/cefe5fce-9330-4e4e-acae-f838e759e783, 2013>
- [15] Buyle, R., De Vocht, L., Van Compernelle, M., De Paepe, D., Verborgh, R., Vanlshout, Z., De Vidts, B., Mechant, P., Mannens, E. (2016). OSLO: Open Standards for Linked Organizations. Proceedings of the 2016 3rd International Conference on Electronic Governance and Open Society: Challenges in Eur-asia'. (November 21-22 2016). <http://dx.doi.org/10.1145/3014087.3014096>
- [16] Colas, C., Goedertier, S., Kourtidis, S., Loutas, N., Rubino, F. (2013). Core Location Pilot - Interconnecting Belgian National and Regional Address. pp.9-46. https://joinup.ec.europa.eu/svn/core_location/pilot/D5.2.1_Core_Location_Pilot-Interconnecting_Belgian_National_and_Regional_Address_Data.pdf
- [17] Danish Enterprise and Construction Authority. (2010). Country Report Denmark: Addresses. pp. 1-11. http://danmarksadresser.dk/file/389699/0_CountryReport_2010_DK_EN.pdf
- [18] Davidson, P. (2011). Designing URI Sets for Location. Chief Technology Officer Council. https://data.gov.uk/sites/default/files/Designing_URI_Sets_for_Location-V1.0_10.pdf
- [19] Davidson, P. (2009). Designing URI Sets for the UK public sector. UK Chief Technology Officer Council. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/60975/designing-URI-sets-uk-public-sector.pdf
- [20] De Wolf L., Van Hemelryck, H (2016). The Large-scale Reference Database (LRD), INSPIRE Conference 2016, p.15. http://inspire.ec.europa.eu/events/conferences/inspire_2016/pdfs/2016_psessions/28%20WEDNESDAY_PSESSIONS_A_16.00-18.00_____pldwhvh160928av2_final_INSPIRE2016_GRB_met_comments.pdf
- [21] eGovernment Sub-group (2009). Visions and priorities for eGovernment in Europe, Orientations for a post 2010 eGovernment Action Plan. p.4. <https://www.aoc.cat/wp-content/uploads/2014/09/visions-priorities-egovernment-europe.pdf?x75273>
- [22] European Commission. (2010). European Interoperability Framework (EIF) for European public services: Annex 2. 2-21, https://ec.europa.eu/isa2/sites/isa/files/isa_annex_ii_eif_en.pdf
- [23] European Commission. (2011). Core vocabularies. pp. 11, 50. https://joinup.ec.europa.eu/catalogue/distribution/core_vocabularies-business_location_person_v100_specification_in.pdf
- [24] European Commission. (2013). A vision for public services. pp.1-6. http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=3179
- [25] European Commission. (2013). eGovernment Core Vocabularies: The SEMIC.EU approach. p13. <https://joinup.ec.europa.eu/sites/default/files/c1/23/4a/egovernme-nt-core-vocabularies.pdf>
- [26] European Commission (2014). Handbook for using the Core Vocabularies. pp 8-10. https://joinup.ec.europa.eu/site/core_vocabularies/Core_Vocabularie_s_user_handbook/Handbook-for-using-the-Core-Vocabularies_v0.50.pdf
- [27] European Commission (2017). European Interoperability Framework - Implementation Strategy, Annex 2. pp. 31-32. http://eur-lex.europa.eu/resource.html?uri=cellar:2c2f2554-0faf-11e7-8a35-01aa75ed71a1.0017.02/DOC_1&format=PDF
- [28] European Commission, ISA (2017). New European Interoperability Framework (EIF). p. 5. https://ec.europa.eu/isa2/sites/isa/files/eif_brochure_final.pdf
- [29] European Commission. (2013). Process and methodology for developing semantic agreements. p7. <https://joinup.ec.europa.eu/sites/default/files/Process%20and%20methodology%20for%20developing%20semantic%20agreements.pdf>
- [30] Fielding, R., Taylor, R. (2000). Architectural styles and the design of network-based software architectures. Doctoral dissertation: University of California, Irvine. pp. 151. <http://jpkc.fudan.edu.cn/picture/article/216/35/4b/22598d594e3d93239700ce79bce1/7ed3ec2a-03c2-49cb-8bf8-5a90ea42f523.pdf>
- [31] Folmer, E., Verdonk, Y. (2014). Linked data in beeld: Platform Linked Data Nederland. <https://www.geonovum.nl/sites/default/files/Linked%20Data%20in%20beeld-2014-Locked.pdf>
- [32] Garson D., Biggs, R. (1992). Analytic Mapping and Geographic Databases, SAGE university paper, p. 87.
- [33] Gottschalk, P. (2009). Maturity levels for interoperability in digital government. *Government Information Quarterly*, 26(1), 75-81. doi:10.1016/j.giq.2008.03.003
- [34] Homans, L. (2015). Beleidsbrief Bestuurszaken 2015-2016, pp. 25-27. <https://docs.vlaamsparlement.be/docs/stukken/2015-2016/g522-1.pdf>

- [35] Horst, N. L., Bjerre, S., Lind, M., Hvingel, L. (2014). The Basic Data Programme—A Danish Infrastructure Model for Public Data. *Geoforum Perspektiv*, pp. 36-38. <https://journals.aau.dk/index.php/gfp/article/viewFile/637/518>
- [36] INSPIRE. (2003). Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:02003L0098-20130717>
- [37] INSPIRE. (2007). Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community. Published in the official Journal on the 25th April, <http://inspire.ec.europa.eu/documents/directive-20072ec-european-parliament-and-council-14-march-2007-establishing>
- [38] INSPIRE (2013). Data Specification on Addresses. pp. VII, 6-7. http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_AD_v3.0.pdf
- [39] ISA programme by PwC EU Services EESV. (2013). Case study: How Linked Data is transforming eGovernment. pp 1-5. <http://ec.europa.eu/isa/documents/case-study-data-integration-v2.pdf>
- [40] Jønsen, B. Basisdata: forståelsesramme og analysemodel til kategorisering af basisdata: Afrapportering fra Udvalg til nytænkning vedrørende basisdata. Servicefællesskabet for Geodata,(2004). http://vbn.aau.dk/files/1815466/BHJ_Basisdata.pdf
- [41] Kubicek, H., Cimander, R. (2009). Three dimensions of organizational interoperability. *European Journal of ePractice*, 6, pp. 1-12. http://www.ifib-consult.de/publikationsdateien/Kubicek_Cimander_ePractice_Journal_vol_6.pdf
- [42] Lind, M. (2013). The Danish Address Program – Socio-economic Benefits from improved Addressing. pp. 4-13. <http://odimpact.org/case-denmarks-open-address-data-set.html>
- [43] Loutas, N., Brulé, D., Wyns, B., Zajac, A. EU FISCAL DATA AS LINKED OPEN DATA. Publications Office of the EU. pp. 7-8 (2017). https://joinup.ec.europa.eu/sites/default/files/document/2017-07/eu_buget_as_lod_v1.0_0.pdf
- [44] Mostaert, K., Buyle, R., Vanderstraete, T. (2016). Interoperability as a key enabler for citizen-centric location based services, INSPIRE. http://inspire.ec.europa.eu/events/conferences/inspire_2016/schedule/submissions/394.html
- [45] Overbeek, H., Brentjens, T. (2013). Draft URI Strategy for the NL Public Sector. Ministry of the Interior of the Netherlands, Geonovum, https://www.w3.org/2013/04/odw/odw13_submission_14.pdf
- [46] Overbeek, H., van den Brink, L. (2013). Towards a national URI-Strategy for Linked Data of the Dutch public sector. http://www.pilod.nl/doc/D1-2013-09-19_Towards_a_NL_URI_Strategy.pdf
- [47] Peeters, K. (2011). Ministerieel besluit tot vaststelling van de CRAB-specificaties. B.S. 12.04.11, <https://www.agiv.be/~media/agiv/producten/crab/documenten/ministerieel%20besluit%20tot%20vaststelling%20van%20de%20crab-specificaties.pdf>
- [48] Rocheleau, B. (2003). Politics, accountability, and governmental information systems. *Public information technology: Policy and management issues*. pp. 20-52.
- [49] Turnbull, J. (2014). The Docker Book: Containerization is the new virtualization. James Turnbull. https://dockerbook.com/TheDockerBook_sample.pdf
- [50] Vanlshout, Z., Buyle R. Open Standards for Linked Organisations. (2016). INSPIRE conference. p.10. 2016, http://inspire.ec.europa.eu/events/conferences/inspire_2016/pdfs/2016_psessions/inspire_Linked-final.pdf
- [51] Verborgh, R., Vander Sande, M., Colpaert, P., Coppens, S., Mannens, E., Van de Walle, R. (2014). Web-Scale Querying through Linked Data Fragments. LDOW https://www.researchgate.net/profile/Ruben_Verborgh/publication/264274086_Web-Scale_Querying_through_Linked_Data_Fragments/links/53f498b10cf2f2ceacc6e918d.pdf
- [52] Vlaamse Overheid. (2009). Decreet betreffende de Geografische Data-Infrastructuur Vlaanderen, Belgisch Staatsblad B.S. 28.04.09. http://www.ejustice.just.fgov.be/mopdf/2009/04/28_2.pdf#Page25
- [53] Vlaamse Overheid. (2008). Decreet betreffende het elektronische bestuurlijke gegevensverkeer, Belgisch Staatsblad B.S. 29.10.08, http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=nl&la=N&table_name=wet&cn=2008071823
- [54] Vlaams Parlement. (2009). Ontwerp van decreet betreffende het Centraal Referentieadressenbestand, Stuk 2067, p.5. <https://www.agiv.be/~media/agiv/producten/crab/documenten/crab-voorontwerp-decreetg2067-1.pdf>
- [55] Wood, D., Zaidman, M., Ruth, L., Hausenblas, M. (2014). Linked Data - Structured Data On The Web, Manning, p.4. <https://livebook.manning.com/#!/book/linked-data/chapter-1/99>
- [56] Zeginis, D., Hasnain, A., Loutas, N., Deus, S., Fox, R., Tarabanis, K. (2012). Collaborative development of a common semantic model for interlinking Cancer Chemoprevention linked data sources. *The Semantic Web Journal*. p. 14 <http://www.semantic-web-journal.net/sites/default/files/swj263.pdf>