Article

The Zaragoza's Knowledge Graph: Open Data to Harness the City Knowledge

Paola Espinoza-Arias [1,*] **, María Jesús Fernández-Ruiz** [2] **, Victor Morlán-Plo** [2] **,**
Rubén Notivol-Bezares [2] **and Oscar Corcho** [1]

[1] Ontology Engineering Group, ETSI Informáticos, Universidad Politécnica de Madrid, 28660 Madrid, Spain;
    ocorcho@fi.upm.es

[2] Ayuntamiento de Zaragoza, Vía Hispanidad 20, 50071 Zaragoza, Spain; mjferuiz@zaragoza.es (M.J.F.-R.);
    vmorlan@zaragoza.es (V.M.-P.); internet@zaragoza.es (R.N.-B.)

[*] Correspondence: pespinoza@fi.upm.es; Tel.: +34-910673091; Fax: +34-913524819

check for updates

Abstract: Public administrations handle large amounts of data in relation to their internal processes as well as to the services that they offer. Following public-sector information reuse regulations and worldwide open data publication trends, these administrations are increasingly publishing their data as open data. However, open data are often released without agreed data models and in non-reusable formats, reducing interoperability and efficiency in data reuse. These aspects hinder interoperability with other administrations and do not allow taking advantage of the associated knowledge in an efficient manner. This paper presents the continued work performed by the Zaragoza city council over more than 15 years in order to generate its knowledge graph, which constitutes the key piece of their data management system, whose main strengthen is the open-data-by-default policy. The main functionalities that have been developed for the internal and external exploitation of the city's open data are also presented. Finally, some city council experiences and lessons learned during this process are also explained.

1. Introduction

Public administrations are large data producers and collectors due to the need to provide support for their internal processes, as well as for services offered to citizens. These data are traditionally available in several formats, according to their own information models, and are managed in different, sometimes isolated, data sources. From the point of view of the organization, the importance of data is undeniable. A proper organizational data management improves decision-making, operational efficiency and service provision. From the side of citizens, the public administration's data are an essential asset that must be made available in order to enhance transparency and the accountability [1]. From the point of view of re-users, having these data available enables developing value-added services and applications, consequently stimulating innovation and economy [2]. Having all these benefits in mind, public administrations are increasingly publishing their data as Open Data, defined as data that can be freely used, re-used and redistributed by anyone—subject only, at most, to the requirement to attribute and sharealike [3]. An Open Government Data strategy ensures that data are made available across city departments and to third parties, contributes to citizen engagement, increases democracy and serves to drive economic growth and social improvement [4,5].

However, these data are mostly represented with information models defined without previous agreements or consensus processes with other institutions, which hinders the possibility of using and

sharing them within the same organization or with third parties. This lack of adoption of shared models is one of the major causes of inefficiency of information management within large organizations [6]. In this regard, ontologies, defined as formal specifications of shared conceptualizations [7], have been used to describe data without ambiguities. Hence, these may be good artifacts to facilitate reusability, interoperability and data quality assurance.

When all these data are organized according to ontologies (also known as vocabularies in the context of Open Government Data) and represented in some graph-like format, we can talk about a Knowledge Graph. Despite there being no consensus about a single definition for a Knowledge Graph [8–11], in this work it is understood as a graph of data with the intent to compose knowledge [12]. Knowledge Graphs have become a powerful resource in big industries, for example, Google's Knowledge Graph, Facebook's Search Engine, and Microsoft's Satori, among others, use Knowledge Graphs for their business purposes [10,13]. Likewise, an increasing amount of public administrations are creating and maintaining Knowledge Graphs as a way to capture, represent, publish and exploit their data. In this respect, several works about Knowledge Graphs for public administration data management have been reported in the literature [14–17].

Managing a Knowledge Graph in any organization (e.g., a public administration) represents an advantage because it allows sharing and reusing common parts of knowledge across the organization. This internal reuse promotes that the main re-user will be the own public administration and as a result an open-data-by-default management. In addition, a knowledge graph is a valuable source for scientific studies, economic development, active involvement of citizens, etc.

Having all of these aspects in mind, the Zaragoza city council has been working for more than 15 years on releasing its public data as Open Data. The city council has released its data in structured and open formats, described by agreed vocabularies (when possible) and available from several tools. In addition, the city council efforts to implement this approach have resulted in the development of a well-defined data management process that eases and boosts an open-data-by-default policy. One of the main assets of this policy is the Zaragoza's Knowledge Graph, which clusters all these data in order to exploit the knowledge of the city.

In this paper, we describe the work performed to generate the Zaragoza's Knowledge Graph. The remainder of this manuscript is structured as follows: In Section 2 several public administration initiatives for Open Data publishing are introduced. In Section 3 a brief history of the Zaragoza's Knowledge Graph is presented. Then, the technological approach for Knowledge Graph generation and maintenance is shown in Section 4 and an overview of the Knowledge Graph is provided in Section 5. Finally, conclusions about this work and foreseen future work are listed in Section 6.

2. Related Work

According to the 2017 Open Data Barometer (http://devodb.wpengine.com/?_year=2017&indicator=ODB), Canada has ranked first in the world for publishing Open Data, followed by the United Kingdom and Australia. In this respect, the United Kingdom government has been one of the pioneers in generating linked open government data through data.gov.uk (https://data.gov.uk), its repository of public sector datasets [18]. A related effort is the Local government Open Data portal (https://opendata.esd.org.uk), which encourages standardization and publishing of local United Kingdom councils' Open Data. This portal includes local data inventories, schemas, and datasets to promote consistency of publication. In addition, several well-defined standards are employed in order to provide a common way of referring to main concepts, for example, the Local Government Service List (https://standards.esd.org.uk/?uri=list%2Fservices), which provides a unique number and description for each local council service. Furthermore, London is considered a role model for many other cities worldwide on regarding how to share Open Data, through to its London Datastore portal (https://data.london.gov.uk).

Likewise, the Canadian government launched its Open Data Portal (https://open.canada.ca/en/open-data) as a way to publish data collected by the government in a standard compliant

manner, making data more accessible to everyone [19]. One of the first representative city Open Data sharing efforts in this country was the City of Toronto Open Data portal (https://open.toronto.ca). This portal has been recently updated in order to lightweight the steps required to publish open datasets and to improve the internal processes. However, the use of schemas for data representation and the implementation of important functionalities such as visualizations, and public Application Prohramming Interfaces (APIs), are still announced as a work in progress. At present, the City of Edmonton's Open Data Portal (https://data.edmonton.ca) is considered as the Canadian leader in Open Data publication. The portal includes several assets such as datasets, visualizations, maps, and APIs, what allows exploiting the knowledge of the city.

Similarly, the Australian Open Government Data portal (https://data.gov.au) provides access to anonymized public data published by federal, state and local government agencies. In addition, publicly-funded research data and datasets from private institutions that are in the public interest are also provided. Some representative efforts about local Open Data portals publishing datasets in this country are the City of Melbourne's Open Data Platform (https://data.melbourne.vic.gov.au) and the Brisbane Datastore (https://www.data.brisbane.qld.gov.au) [20].

In addition, it is important to highlight that according to the Open Data Maturity Report of 2019 (https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf), Spain occupies the second position in Europe as a country with an advanced Open Data strategy and a strong commitment to Open Data as a strategic asset to drive digital transformation, only surpassed by Ireland. In this regard, some examples of Spanish cities working on Open Data publication are Alcobendas (https://datos.alcobendas.org), Cáceres (http://opendata.caceres.es), Rivas Vaciamadrid (https://datosabiertos.rivasciudad.es), and Málaga (https://datosabiertos.malaga.eu). These cities share their open datasets, in most of the cases, using vocabularies in order to represent their data, and providing an endpoint or an API to facilitate their exploitation. Finally, it is worth mentioning other relevant Open Data publishing efforts from cities in other countries, such as the Dublin Open Data Store (http://www.dublinked.ie), NYC Open Data (https://opendata.cityofnewyork.us), DataSF (https://datasf.org) and OpenData.Swiss (https://opendata.swiss/en), to name a few.

3. History of Zaragoza's Knowledge Graph

Since 2003, the Zaragoza city council has been developing an ambitious project in order to become a national and international reference in technological change and innovation. With this goal in mind, the city council launched several actions with the main focus on the use of new information and communication technologies to provide more efficient public services and guarantee all citizens, entities, companies and institutions the exercise of their full rights to information access.

In 2004, *IDEZar* (Spatial Data Infrastructure of Zaragoza: https://www.zaragoza.es/ciudad/idezar) was developed in order to provide several internal and external services publishing its geographical characteristics compliant to international standards like the INSPIRE Directive (https://inspire.ec.europa.eu) or those provided by the Open Geospatial Consortium (http://www.opengeospatial.org). Since its first days, *IDEZar* was semantically oriented, due to the fact that one of these services was focused on a semantic search engine for public services. It included visualizations, interoperable services, a thematic catalog, etc.

In the road to provide Open Data services, the city council developed the *Planifica tu Visita* (Plan your Visit: http://www.zaragoza.es/turruta/Turruta/index_Ruta) application. This service, released in 2007, consisted of the integration of several legacy databases and the *IDEZar* web services. This integration was achieved by means of a Tourism ontology (https://www.zaragoza.es/sede/portal/datos-abiertos/vocabularios/cruzar) and a data transformation process to the Resource Description Framework (RDF) format. This application was well received, worldwide recognized and selected as a W3C Best Practice (https://www.w3.org/2001/sw/sweo/public/UseCases/Zaragoza-2).

Around those days, the Zaragoza city council's Open Data policy was implemented in order to comply with the Spanish law about public information reuse [21]. The aim of this policy is that public data should be: unique, shared, accessible, and reusable by third parties. Thus in 2010, the city council started to publish several services via the *Datos de Zaragoza* (Open Data Zaragoza: http://datosabiertos. zaragoza.es) portal, promoting information reuse by citizens, enterprises, and other organizations in order to increase transparency. *Datos de Zaragoza* mainly aggregates the previous developments and experiences of the city council. Thus, having in mind that the Spatial Data Infrastructures (SDI) has been the precursors of data opening, the *IDEZar* service represents the core of the current Open Data portal [22]. Currently, one of the IDEZar key reference services is the *Callejero* (http://idezar.zaragoza.es/callejero), which allows extracting information based on city roads, searches and visualizations.

Nowadays, this Open Data policy makes the city operate an open-data-by-default policy for all the public sector information handled by the city council. It has been possible because of the implementation of a data management system for the Zaragoza's website (http://www.zaragoza.es/sedeelectronica). This website contains all city council information and services, for instance, the open data portal. The data management system has the following goals: using open-data-by-default, unifying data life cycle management tools and promoting code reuse developed by other administrations. In order to accomplish these goals an API has been implemented as a core piece of this system. The business logic developed for the API becomes the only data access point for read and write operations. Thus, it allows external or internal management tools to be integrated with the API. This approach ensures the sustainability of the Open Data platform, in contrast to other approaches where data publication is performed at the end of the data management life cycle, sometimes creating graveyards.

Following Zaragoza's approach, a Spanish project named *Ciudades Abiertas* (Open Cities: https://ciudadesabiertas.es) is currently under development. The main objective of this project is generating an interoperable and collaborative platform in order to promote Open Government, which refers to the relationship between citizens and public administrations. On the one hand, citizens will have permanent access to information on administrative processes in order to exercise adequate public control and to become part of government actions. On the other hand, public administrations will manage data in a unique, shared, open-by-default, georeferenced, and semantically annotated manner. This project represents a step forward in order to afford a common representation and generic infrastructure between four pilot Spanish city councils (Zaragoza, Madrid, A Coruña and Santiago de Compostela), to be extended to others in the future.

4. Technological Approach for Zaragoza's Knowledge Graph Generation

There is no one way to create, maintain, and exploit Knowledge Graphs. The process depends on the characteristics of the organization, on use cases, technical aspects, etc. [23]. The Zaragoza's Knowledge Graphs is not an exception, and therefore the city council has followed its own approach. As argued in Section 3, Zaragoza follows an open-data-by-default policy which aims to release data proactively and ensure that the main Open Data re-user is the city council itself. Thus, Zaragoza has implemented a data management system whose core tool is an API. The Zaragoza API has been defined as a homogeneous and unique data access point without data source restrictions. This means that if there are datasets without semantic representation, they can be also accessed in non-semantic formats, for example, CSV. However, in this section only those activities and tasks for semantically annotated open dataset publication are presented.

In Figure 1, the technological approach for Knowledge Graph generation is depicted. It involves two main parts: (a) RDF Generation; and (b) API Management. The first part aims to materialize data as RDF for all those data sources which have a vocabulary describing them. It allows having the Knowledge Graph with a SPARQL endpoint where complex queries over these data may be posed. The second part aims to provide data source (e.g., databases) access and management via API

operations. It allows having an only data access point where get and write operations over data sources may be executed. Briefly, the whole workflow works in the following manner when a data request from the Zaragoza's website (or from external apps or third-parties) is posed it is processed in order to execute a specific operation and retrieve the appropriated response. For example, if a write operation over a data source is requested it is processed at the API management part. Then, the operation is executed at the data source and, if the data source has a semantic representation, this change is automatically updated over the RDF. In the following subsections a detailed description of RDF Generation and API Management parts is presented.
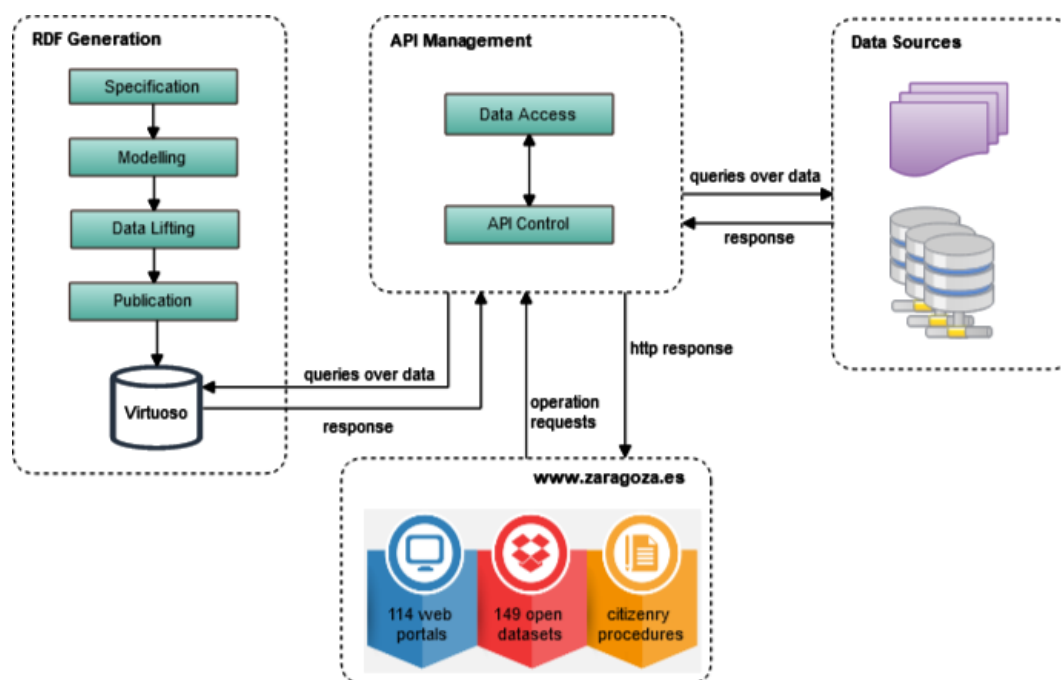

Figure 1. Technological approach for Zaragoza's Knowledge Graph generation.

4.1. RDF Generation

This first part, inspired by the activities from the guidelines presented on [24], aims to generate an RDF representation of the data. The activities performed in this part are: specification, modelling, data lifting, and publication.

4.1.1. Specification

The first activity refers to data identification and selection. In this activity, several variables are studied in order to analyze the feasibility of data publication. Some of these variables are information domains, source formats, data licenses, technical limitations of data sources, relevance for re-users (e.g., citizens, enterprises, city council, public administrations), and transparency needs, etc. Having in mind the above criteria, the datasets are selected for publication by the Open Data management team or discarded.

4.1.2. Modelling

In this activity, data are connected to vocabularies. These vocabularies have either been (a) directly reused if there is a suitable one which encompasses the data, or (b) developed if there is not any available vocabulary that is suitable for all entities. General vocabularies (e.g., Dublin Core [25], schema.org [26], among others) are used to annotate, for example, data properties which are common in most datasets, such as identifiers, and geospatial information, etc. Specific vocabularies are used to annotate data particularities (e.g., public procurement contract details [27]).

Specific vocabularies have been designed following the best practices described in [28] and the Spanish Technical Interoperability Standard [29]. In order to avoid their creation from scratch, these vocabularies have been developed reusing existing ones as much as possible. These specific vocabularies have been made available on-line as human-readable documentation (HTML) and in a machine-readable format (Ontology Web Language—OWL). Both versions of the vocabulary are reachable from the vocabulary Uniform Resource Identifier (URI) by means of content negotiation mechanisms following good practices in ontology publication. It is worth mentioning that the city council has participated in standardization groups, projects, etc., which have resulted in some of these vocabularies being promoted further and now being part of larger efforts. For example, Zaragoza was a member of the group who defined the AENOR norm UNE 178301:2015 (https://www.aenor. com/normas-y-libros/buscador-de-normas/une?c=N0054318). This norm, entitled "Smart Cities. Open Data", aims to provide a set of indicators to measure the maturity level of Open Data projects for smart cities; and it also provides a set of common vocabularies in order to ease the development and deployment of these Open Data projects. All vocabularies used to represent the published datasets are presented in Section 5.

4.1.3. Data Lifting

This activity aims to represent data in RDF. The data lifting activity involves two tasks: transformation and linking. In the first task, the datasets selected during the first activity are transformed according to the vocabularies specified in the modeling activity. In the second task, external data sources are used in order to enrich some datasets. The two main external data sources used so far are DBpedia (http://dbpedia.org) and Geonames (http://www.geonames.org). For instance, the *Callejero* dataset has links to several DBpedia resources. This last task has been implemented with scripts which execute API calls to the specific resources of the external sources.

4.1.4. Publication

This activity consists on data storage and publication. In order to store and make available the resulting RDF, the Virtuoso triplestore (http://virtuoso.openlinksw.com) has been used. Virtuoso allows storing RDF triples in order to be accessed by SPARQL queries. These queries may be executed at the Zaragoza's SPARQL endpoint (https://www.zaragoza.es/sede/portal/datos-abiertos/servicio/ sparql). This technology allows taking advantage of data, because it enables more complex queries than those allowed by an API, for example, unified queries on multiple datasets.

4.2. API Management

This part aims to implement all necessary procedures for API administration. It involves two activities: API Control and Data Access. In order to deal with both activities a programming layer has been developed in Java with SPRING MVC and HIBERNATE as main frameworks. More precisely, the REST API is based on JSR 311 (JavaTM API for RESTful Web Services: https://jcp.org/en/jsr/ detail?id=311) It is an ad hoc implementation compliant with the Spanish Technical Interoperability Standard [29]. For instance, URI identifiers follow this standard and are usually composed of a base URI: https://www.zaragoza.es/sede/servicio/{sector}/{domain}/{class}/{ID}.

4.2.1. API Control

In this activity, all the requests are processed to execute a specific operation and retrieve the corresponding response. As depicted in Figure 1, several operations to the API may be requested from the Zaragoza's website (or from external apps of third parties). When a new request is received it is automatically redirected to its specific operation, which is defined in the controller layer, developed for this purpose. Likewise, this specific operation uses the corresponding operation against data sources, which is configured in the data access layer as it will be explained in Section 4.2.2. In addition

to the operations, in this controller layer the response formats and the possible error responses are also defined.

The API allows five operations: POST, GET, PUT, PATCH, and DELETE. Furthermore, it is possible to perform optional operations such as HEAD and OPTIONS. Regarding the formats, the API supports: JSON, GeoJSON, XML, CSV, JSON-LD, RDF, Turtle, and N3. Furthermore, the API allows two methods for requesting data in a specific format: (1) providing the *Accept* parameter embedded in the request header and performing content negotiation (suggested method), for example *Accept:application/rdf+xml*, and (2) specifying the format extension in the request, for example https://www.zaragoza.es/sede/servicio/monumento.rdf.

It is worth mentioning that the availability of the output result in some of these formats depends on the data source. For example, if a data source does not have a semantic representation it cannot be retrieved as Turtle, instead a CSV result may be possible. In addition, depending on the dataset, the API will accept some of these above-listed formats as required and others as optional. Required parameters are JSON (application/json) and XML (application/xml), and the remaining parameters are optional. Likewise, the API supports several options such as filtering, selection by field, sorting, and pagination of the results. Additionally to these operations, the Zaragoza API also allows evaluating more complex expressions using the FIQL query language (https://fiql-parser.readthedocs.io/en/stable/usage.html). In Table 1 an example of the use of all API operations over the *Monumento* resource is provided. In this case, the *Monumento* resource URI corresponds to https://www.zaragoza.es/sede/servicio/monumento.

Table 1. API operations and their actions over the *Monumento* resources.

Resource	POST	GET | OPTIONS	PUT	PATCH	DELETE
.../monumento/	create a new monument	list of monuments, with OPTIONS only HTTP headers	massive update of monuments	partial/massive update of monuments	delete monuments by query
.../monumento/{id}	error	show the monuments with {id}, with OPTIONS only HTTP headers	if the monument with {id} exists then update it; otherwise, return an error	if the monument with {id} exists then partially update it; otherwise, return an error	delete monument with {id}

In addition, Listing 1 shows a code excerpt example of the GET operation to retrieve a list of *Monumentos* (Monuments). At the beginning of this code, the result entity and the requested formats are specified with `@ResponseClass` and `@ResponseMapping` annotations, respectively. Then, the `apiListar` (Listing API) method is executed and the output will be a `ResponseEntity` annotated as `@ResponseBody`, which will be composed by the HTTP response and the objects with the requested content. The `dao` (data access object) corresponds to the layer where queries over data are defined, in this case a `searchAndCount` function of *Monuments*. Data access object operations are explained in next Section 4.2.2. In summary, when a request corresponds to a GET operation, the control layer executes the operation and calls the corresponding data access method to deal with data sources.

Regarding API write operations, when a POST, PUT, PATCH, or DELETE operation is executed over the data source, an automatic update over the Virtuoso triplestore is triggered. It is implemented with the Virtuoso Jena Provider functionality (http://vos.openlinksw.com/owiki/wiki/VOS/VirtJenaProvider). When the write operation over the data source was executed correctly, then specific actions are implemented in Virtuoso. It will depend on the type of operation, for example, a POST involves a simple data insertion of the triples associated with the new resource into the triplestore.

Listing 1. GET method code excerpt in Java.

```
1
2  @Cacheable(Cache.DURACION_30MIN)
3  @ResponseClass(value = Monumento.class , entity = SearchResult.class)
4  @RequestMapping(method = RequestMethod.GET, produces = {MimeTypes.GEOJSON, MimeTypes.GEORSS,
5  MimeTypes.JSON, MimeTypes.XML, MimeTypes.CSV, MimeTypes.JSONLD, MimeTypes.RDF, MimeTypes.TURTLE,
6  MimeTypes.RDF_N3})
7
8  public @ResponseBody ResponseEntity apiListar(@Fiql SearchFiql search) throws SearchParseException {
9      return ResponseEntity.ok(dao.searchAndCount(search.getConditions(Monumento.class)));
10 }
```

Finally, in order to deal with user restrictions, the API write access is managed by two parameters inserted in the request header. The first parameter is the client identification, allowing to check who performs the request. The second parameter is a hash code generated with the data of the request. If the hash regenerated in the server is the same as the one sent in the request and the client has privileges for executing the operation, the execution is authorized. Moreover, the API currently supports both HTTP and HTTPS protocols, and may be used with or without API keys (the latter recommended for third party developers), with no usage limits in any of the cases, although these may be applied in the future.

4.2.2. Data Access

In this activity, all the mappings and access operations against the data sources are manually configured. It involves two tasks: Entity Annotation and Data Access Objects Definition. In the first task, the datasets selected during the specification activity have been semantically annotated according to the vocabularies specified in the modelling activity. Both specification and modelling activities corresponds to those previously explained in the RDF generation part (Section 4.1). In order to define all the attributes that will be mapped with the data source, a Java entity class is generated for each entity. This annotation is performed on the Java entity class and class attributes with @Rdf and @RdfMultiple annotations, defining simple and multiple annotations, respectively. For example, Listing 2 shows a code excerpt of the *Monumento* (Monument) entity class and the annotations performed on it. At the beginning of the code excerpt, all concepts from several vocabularies are specified in order to represent the *Monumento* entity. Then, inside of the Monumento class definition, id and address attributes are also annotated according to properties from the specified vocabularies. Likewise, a column name for each attribute is defined, in order to configure the mapping against the specific column from the data source.

Once the Entity Annotation task has been configured, the Data Access Object Definition task is carried out. This second task allows accessing data sources and it is implemented by a data access object layer, where several operations are defined. It may be understood as a data source access interface that reuses the resulting mappings from the previous task. It works in the following manner: when a request from the API control activity executes a specific operation, it is also processed in the Data Access Object activity in order to perform the operation against the data source. Thus, if the operation is a GET, the workflow involves dataset extraction from the data source and entity annotation, both executed in this data access layer. Then, in the API control layer, the data retrieved is converted to an API response with the requested semantic requested format. Likewise, if the operation involves a write procedure (POST, PUT, or PATCH) the data source is updated and the triple store is also synchronized with those changes. Then, in the API control layer, the response of the executed operation is retrieved and it is converted to an API response.

Listing 2. *Monumento* entity annotation code excerpt in Java.

```java
1  //Entity annotations for the Monumento class
2  @RdfMultiple({
3      @Rdf(contexto = Context.ESTURISMO, propiedad = "LugarInteresTuristico", prefijo = "monumento"),
4      @Rdf(contexto = Context.SCHEMA, propiedad = "TouristicAttraction", prefijo = "touristic_attraction"),
5      @Rdf(contexto = Context.GEONAMES, propiedad = "Feature", prefijo = "feature"),
6      @Rdf(contexto = Context.CRUZAR, propiedad = "Monumento", prefijo = "monumento")})
7
8  public class Monumento extends EntidadBase {
9
10 //Annotations for id attribute
11 @Id
12 @Rdf(contexto = Context.DCT, propiedad = "identifier")
13 @Column(name = "ID_MONUMENTO", nullable = false, unique = true)@Digits(integer = 22, fraction = 0)
14 private Integer id;
15
16 //Annotations for address attribute
17 @RdfMultiple({
18     @Rdf(contexto = Context.SCHEMA, propiedad = "streetAddress"),
19     @Rdf(contexto = Context.VCARD, propiedad = "street-adr"),
20     @Rdf(contexto = Context.LOCN, propiedad = "fullAddress")})
21 @Column(name = "SITUACION") @Size(max = 20)
22 private String address;
23 .....
24 }
```

5. The Zaragoza's Knowledge Graph

In this section, an overview of content available in the Zaragoza's Knowledge Graph is presented. As previously stated, this Knowledge Graph aims to provide Open Data with semantic descriptions based on well-described vocabularies. Thus, first a brief description of the vocabularies used to provide a semantic representation of these data is presented. Then, several options for making it easier to access the knowledge stored in the graph are described. Finally, some details about the evaluation of the Knowledge Graph are provided. Despite the Knowledge Graph containing several domains, in this section all examples are provided for the *Monumentos* (Monuments) dataset.

5.1. Semantic Representation

In order to annotate data from Zaragoza's datasets, several vocabularies have been used. These vocabularies are the main assets for the Knowledge Graph. As previously mentioned in Section 4.1.2, these vocabularies may be general or specific. Summarizing, 9 general and 16 specific vocabularies have been used so far. All of them are published in a list available on the Vocabularies page (https://www.zaragoza.es/sede/portal/datos-abiertos/vocabularios) of the Zaragoza's web site. For example, the *Lugar* (Place) vocabulary (http://vocab.linkeddata.es/datosabiertos/def/turismo/lugar) is used to describe the Monument dataset. This vocabulary allows representing data of touristic places such as historic buildings, artistic assets, and touristic routes, among others. In order to provide further details, a list with datasets and specific vocabularies used to semantically describe them is provided in Table A1 from Appendix A.

It is worth mentioning that Zaragoza continues working on a continuous semantization process. Therefore, the city council has planned providing a semantic representation for the following datasets in the coming year: city census, traffic status, public rental bikes, public debt, local premises census, agreements, budget execution, noise pollution, city council agenda, public busses and job offers. The vocabularies for representing the above mentioned datasets are under development in the *Ciudades Abiertas* project. Moreover, the information related with public procurement will be updated with a new vocabulary representation which will result from the European project TheyBuyForYou, in which the city council is a member [30], and which will be based on the Open Contracting Data Standard.

5.2. Knowledge Consumption

In this subsection a description of the intuitive components for consuming the data is provided. In addition, at the end a brief analysis of the details about each consuming alternatives is also presented.

5.2.1. Open Data Catalog

The Open Data Catalog collects the datasets posted on the Zaragoza's web site. Currently, there are 149 open datasets published in this catalog. Datasets are organized by domain, format, update frequency, and key-words; in addition, a quick access to georreferenced datasets is also provided (up to date, there are 66 datasets with a geographic component). The idea is to offer any public item of information in as many formats as possible. In Figure 2 an excerpt from the *Monumentos* dataset search on the catalog page (https://www.zaragoza.es/sede/portal/datos-abiertos/servicio/catalogo) is shown. The result of this search is a link to the Monument dataset and a list of available data formats.

Figure 2. *Monumento* dataset search on an Open Data Catalog page.

In addition, in order to allow re-users to easily discover relevant data, this catalog provides metadata for each dataset. Metadata annotation has been performed with the DCAT (https://www.w3.org/TR/vocab-dcat) and GeoDCAT-AP (https://inspire.ec.europa.eu/good-practice/geodcat-ap) vocabularies. It is worth noting that this catalog is federated to the Spanish National Catalogue (https://datos.gob.es) which allows providing greater visibility for the public datasets published in Spanish Open Data portals.

5.2.2. SOLR Classified Data

In order to provide a quick and friendly data access interface to the Knowledge Graph, the city council has also deployed a SOLR platform (http://lucene.apache.org/solr) posting Open Data on the Zaragoza's web site. This tool enables data indexing, full and faceted text searching, filtering, etc. Further details about these functionalities are available in the SOLR documentation section (http://www.zaragoza.es/sede/portal/datos-abiertos/solr) of the web site. As an example of its functionality, in Figure 3 a faceted search page of Zaragoza's Monuments (https://www.zaragoza.es/ciudad/turismo/en/organiza-viaje/buscar_Monumento) is shown. By default this page makes use of the API service to retrieve a list of Monuments. In Listing 3 a code excerpt of this API response (https://www.zaragoza.es/buscador/select?q=*:*%20AND%20-tipocontenido_s:estatico%20AND%20category:Tr%5Cu00e1mites&rows=20&wt=json) is presented. This Listing includes the total number of monuments, the number of records and the list of Monuments in JSON format. These data are the same appearing on the page depicted in Figure 3; `title`, `description`, and `lastUpdated` attributes are visible on the web interface. Due to space restrictions, only data of the first Monument is

presented both for the case of the Listing and for the Figure. In addition, if a user needs to search a specific Monument, they must introduce the Monument's name on the box and click on the *Consultar* (Search) button in order to retrieve the required record. Finally, further examples of SOLR queries are reachable on Zaragoza's GitHub repository (http://zaragoza-sedeelectronica.github.io/solr/queries).

Figure 3. Zaragoza's Monuments page based on a SOLR faceted search.

Listing 3. list of *Monumentos* retrieved in JSON format as a result of a SOLR request.

```
1  {
2      "totalCount": 175,
3      "start": 0,
4      "rows": 100,
5      "result": [
6      {
7          "id": 63,
8          "title": "Museo Provincial de Zaragoza: Secciones de Antiguedad  y Bellas Artes",
9          "description": "<p>Plaza de los Sitios , 6  Horario:De martes a sabado de 10 a 14.00 [...]",
10         "lastUpdated": "2019-12-10T00:00:00Z",
11         "sameAs": "http://www.zaragoza.es/api/recurso/turismo/monumento/63"
12     },
13     ...
14     ]
15 }
```

5.2.3. API User Interface

In order to provide a friendly access point for interested re-users, the API has been published as a user web service. It is implemented with SWAGGER (http://swagger.io), an open-source tool which allows describing, producing, consuming, and visualizing RESTful APIs. This user interface provides details about open datasets published in the API, operations, parameters and allowed responses. Depending on the user authentication, the authorized API operations will be reachable on the interface page. For example, in Figure 4 the Zaragoza API user interface (http://www.zaragoza.es/docs-api) of the Monument dataset is shown. In this example, as an anonymous user only GET operations over the Monument dataset are available. In this case, the first operation allows retrieving a list of all monuments and the second allows retrieving data from a specific Monument. Further examples of the API usage are available on the Zaragoza's GitHub repository (http://zaragoza-sedeelectronica.github.io/rest/queries).
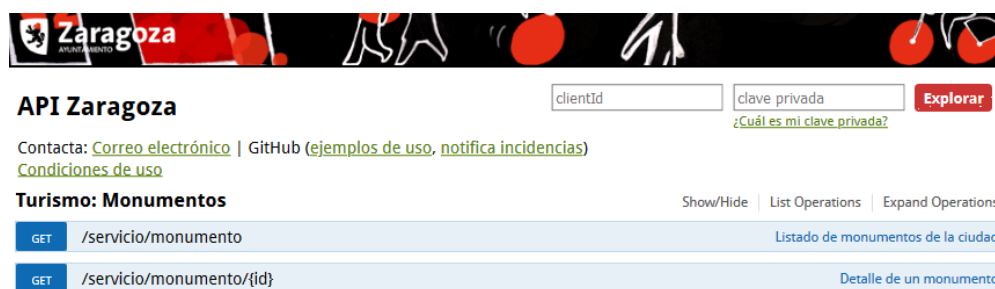
Figure 4. Zaragoza API user interface of the *Monumento* dataset.

5.2.4. SPARQL Endpoint

As above-mentioned in the Publication activity from Section 4, the city council manages a Virtuoso triplestore that provides a SPARQL endpoint (https://www.zaragoza.es/sede/portal/datos-abiertos/servicio/sparql). It allows to execute custom queries in order to retrieve specific data. For example, in Figure 5 an example of a Monuments query is presented. In this query, the requested attributes of Monuments are URI, title (name), description, and the coordinates of its location. In addition, an example of the location of the retrieved Monuments is presented over the map. The source code of this visualization example is available on the Zaragoza's GitHub repository (https://github.com/zaragoza-sedeelectronica/zaragoza-sedeelectronica.github.io/tree/master/sparql/ejemplos/monumentos.html). In addition, other SPARQL query examples are also provided in another directory of this repository (http://zaragoza-sedeelectronica.github.io/sparql/queries/).

```
PREFIX monument: <http://vocab.linkeddata.es/datosabiertos/def/turismo/lugar#>
SELECT DISTINCT ?uri ?title ?description min(?latitud) as ?latitud
min(?longitud) as ?longitud WHERE {
?uri a monument:LugarInteresTuristico.
  OPTIONAL{ ?uri rdfs:label   ?title}.
  OPTIONAL{ ?uri rdfs:comment  ?description}.
  ?uri geo:geometry ?geo.
    ?geo geo:lat ?latitud.
    ?geo geo:long ?longitud.
} group by ?uri ?title ?description
```
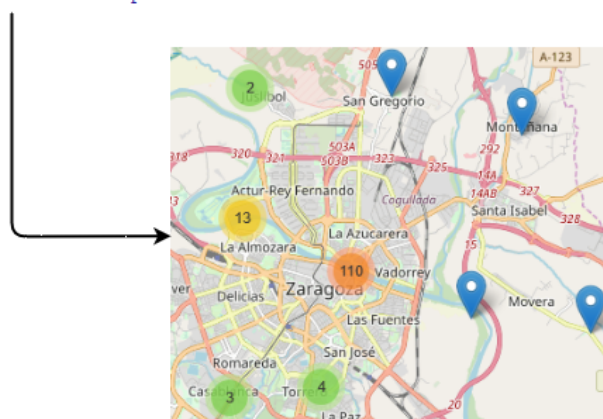

Figure 5. Example of *Monumento* SPARQL query and visualization of its results.

5.2.5. Consumption Technologies Comparison

The Zaragoza city council provides consumption functionalities, which are not limited only to simple data downloads. In order to enhance consumption, the Knowledge Graph's content can be accessed and analyzed via the previously mentioned technologies: SOLR, Rest API, and SPARQL Endpoint. With the aim to provide an overview about the main characteristics of these technologies,

a summary is presented in Table 2. The information of this table suggests that if faceted or by-relevance searching abilities are needed the best option would be SOLR. Likewise, if simple queries or writeoperations are requested the most appropriate option would be the API. Finally, if complex queries involving several datasets are demanded, the better option would be to use the SPARQL query interface. That is, the Knowledge Graph is made available using different technologies so as to facilitate its consumption for different use cases.

Table 2. Technologies comparison.

	Rest API	**SOLR**	**SPARQL**
Learning curve	low	medium	high
Number of formats	high	medium	high
Unified queries over multiple datasets	no	no	yes
Query language	FIQL	SOLR	SPARQL
Supported HTTP methods	GET, POST, PUT, DELETE	GET	GET
Query by distance	yes	yes	yes
Faceted search	no	yes	no
Full text search	no	yes	yes
Searching by relevance	no	yes	no

5.3. Knowledge Graph Evaluation

The Zaragoza's Knowledge Graph is a large information source. In Table 3, a summary with the main statistics about this graph, such as number of classes, and properties, is presented. However, besides the quantity it is also important to consider the quality of the encapsulated knowledge. In this respect, the quality of a Knowledge Graph may be measured by several dimensions [11,31]. The most common are: (a) freshness, which refers to the up-to-date information with respect to the data sources; (b) usage; which is related to the advantage the user gains from the use of information; (c) coverage, which refers to the completeness of the information; and (d) correctness, which implies the accuracy of the information.

Table 3. Statistics about Zaragoza's Knowledge Graph.

Name	**Number**
Concepts	312
Properties	594
Instances	5621017
Triples	28654968
Graphs	19

Concerning freshness, the Zaragoza's Knowledge Graph provides reliable and correct data when requested. This is possible due to the API's capability of serving data on-the-fly according to each data request. Thus, information is given on-time and compliant with data sources and vocabularies. In addition, when a change is performed on the data source, an automatic update of the Virtuoso triplestore is executed, guarantying on-time data synchronization.

Regarding the usage of the Knowledge Graph, the city council has developed several services which are consuming the API in order to offer the exploration of the published datasets, mainly including data visualization over maps. As stated above, IDEZar is a key piece for Open Data provision, more precisely those georeferenced datasets that are visualized in order to provide services for citizens, such as issues on the street and bicycle lanes. Further details about visualizations are available at Zaragoza's geoportal (https://www.zaragoza.es/sede/portal/idezar/visualizacion). In addition, other non-georeferenced visualizations are provided on the Zaragoza's transparency portal (https://www.zaragoza.es/sede/servicio/transparencia), for example, urban equipment, procedures and

services, complaints and suggestions, and indicator. All these services help to make data more understandable for citizens.

Moreover, concerning usage, third parties, particularly developers, are also able to consume data from the Zaragoza API. The provision of data stimulates users to create and use new services that are deployed by others. In this respect, several apps based on open datasets have been developed. Currently, the city council has registered 45 apps, which are available at the apps searching service (http://www.zaragoza.es/sede/servicio/aplicacion). Finally, several statistics about the API usage, for example, number of accesses by format and number of accesses by method, etc. are available at the data-use section from the Zaragoza's Open Data portal (https://www.zaragoza.es/sede/servicio/datos-uso).

Regarding coverage, the Knowledge Graph should contain the information of everything about each domain of the city council. In this respect, there are several techniques to analyze the completeness of a graph. In [11] the most relevant techniques, such as using a partial gold standard, using the same Knowledge Graph as a silver standard, or a retrospective evaluation are explained. However, these techniques perform a comparison against a target Knowledge Graph, which is costly to implement. Despite that, the most relevant data sources from the city council are available in the graph, and in the future it may be necessary to add missing knowledge to the graph. In this case several techniques may be applied, for example, data completion.

Finally, regarding correctness, identifying wrong information in the graph is a challenging task that the city council has not implemented yet. The methods for identifying errors in Knowledge Graphs can target various types of information, for example, type assertions, relations between individuals, literal values, and Knowledge Graph interlinks [11].

6. Conclusions & Future Work

Releasing Open Data represents a big effort for public administrations, involving much time and resources. In this regard, the Zaragoza city council set long-term goals in order to improve all the processes involved in Open Data publication. The fundamental and strategic basis for this improvement was the definition and adoption of the open-data-by-default policy. This policy ensures a unified data life cycle for all the city council services, which guarantees that they are the main reusers; therefore, the policy demonstrates sustainability over time.

Implementing this Open-Data-by-default policy involved new challenges, requirements, and technologies in order to collect and handle data from the city council. As a result, this policy involved the development of new data management techniques and new tools for data collection. These new developments have been an investment in order to improve the Open Data release to users and the data integration into the own city council systems for providing real-time and updated data. All of this ensures the management of unique, shared, and semantically annotated data.

In this last regard, using vocabularies to represent released datasets was also fundamental in order to ensure shared and common definitions describing the data. Representing datasets with semantic definitions opens up the opportunity of enrich these data with external knowledge resources and to exploit all the encoded knowledge. Therefore, it is crucial to define agreed-upon city-related data vocabularies and common pattern snippets in order to minimize ad-hoc commitments and to facilitate data interoperability. Despite the fact that the city council have planned providing a semantic representation for several datasets, further work is required. It is important to represent as many datasets as possible with vocabularies, in order to enjoy the benefits offered by the semantic. Thus, it will be necessary to promote new projects or collaborations with other city councils or organizations in order to specify the vocabularies that are needed. From past city council experiences, this kind of work has been a good way to develop or agree upon vocabularies with others, in order to increase the interoperability.

Finally, providing several data consumption functionalities enables the stimulation of users to explore and analyze data in different manners. Users may create and deliver new resources or apps

in order to provide new services and add value to the government data. Moreover, providing city council visualization services has resulted in a good way to present complex data in an understandable output to citizens. However, it is still a remaining task to evaluate the Knowledge Graph regarding its coverage and completeness. As future work, according to these evaluation results (for example, the detection of potentially erroneous statements or incomplete information), several methods for Knowledge Graph refinement may be implemented.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application Programming Interface
CSV	Comma-separated values
FIQL	Feed Item Query Language
HTML	Hypertext Markup Language
INSPIRE	INfrastructure for SPatial InfoRmation in Europe
OWL	Ontology Web Language
RDF	Resource Description Framework
REST	Representational State Transfer
SPARQL	Protocol and RDF Query Language
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium

Appendix A. Datasets and Vocabularies

In Table A1 the list of datasets and specific vocabularies to represent their data are shown.

Table A1. Datasets and vocabularies.

Dataset	Vocabulary
Ayudas y Subvenciones (Grants)	http://vocab.linkeddata.es/datosabiertos/def/sector-publico/subvencion
Equipamientos (Equipments)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal
Juntas administrativas (Administrative boards)	http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio
Puntos de interés (Points of interest)	http://www.opengis.net/ont/geosparql
Catálogo de datos (Data catalog)	http://www.w3.org/ns/dcat
Accidentes de tráfico (Traffic accidents)	http://vocab.linkeddata.es/datosabiertos/def/transporte/accidentalidad-trafico
Información polínica (Pollen information)	http://vocab.linkeddata.es/datosabiertos/def/medio-ambiente/calidad-aire

Table A1. *Cont.*

Dataset	Vocabulary
Callejero de Zaragoza (Zaragoza's streets)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/callejero
Estaciones de servicio (Fuel stations)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal
Aparcamientos públicos (Public parking lots)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal
Alojamientos (Accomodations)	http://vocab.linkeddata.es/datosabiertos/def/turismo/alojamiento
Monumentos (Monuments)	http://vocab.linkeddata.es/datosabiertos/def/turismo/lugar
Calidad del aire (Air quality)	http://vocab.linkeddata.es/datosabiertos/def/medio-ambiente/calidad-aire
Normativa municipal (City council regulations)	http://vocab.linkeddata.es/datosabiertos/def/sector-publico/normativa
Oferta de empleo (Job offer)	http://purl.org/ctic/empleo/oferta
Perfil de contratante (Contractor profile)	http://contsem.unizar.es/def/sector-publico/pproc
Barrios de Zaragoza (Zaragoza's neighborhoods)	http://vocab.linkeddata.es/datosabiertos/def/sector-publico/territorio
Vías públicas (Public roads)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/callejero
Zonas verdes (Parklands)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal
Edificios históricos (Historic buildings)	http://vocab.linkeddata.es/datosabiertos/def/turismo/lugar
Procedimientos, trámites y servicios (Procedures and services)	http://vocab.linkeddata.es/datosabiertos/def/sector-publico/servicio
Centros proveedores de servicios públicos (Public service providers)	http://vocab.linkeddata.es/datosabiertos/def/sector-publico/servicio
AparcaBicis (Bycicle parking)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal
Aparcamiento para personas con discapacidad (Parking for people with disabilities)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal
Organización Municipal (City council organization)	http://www.w3.org/ns/org
Zonas de estacionamiento regulado (Regulated parking zones)	http://www.opengis.net/ont/geosparql
Estaciones de taxi (Taxi stations)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal
Agenda de Zaragoza (Zaragoza's agenda)	http://vocab.linkeddata.es/datosabiertos/def/cultura-ocio/agenda
Restaurantes (Restaurants)	http://schema.org
Estacionamientos motos (Motorcycle parking)	http://vocab.linkeddata.es/datosabiertos/def/urbanismo-infraestructuras/equipamiento-municipal

References

1. Lourenço, R.P. An analysis of open government portals: A perspective of transparency for accountability. *Gov. Inf. Q.* **2015**, *32*, 323–332. [CrossRef]

2. Janssen, M.; Charalabidis, Y.; Zuiderwijk, A. Benefits, adoption barriers and myths of open data and open government. *Inf. Syst. Manag.* **2012**, *29*, 258–268. [CrossRef]

3. Knowledge, Open. Open Data Handbook. 2012. Available online: https://opendatahandbook.org/guide/en/what-is-open-data (accessed on 2 September 2019).

4. Safarov, I.; Meijer, A.; Grimmelikhuijsen, S. Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Inf. Polity* **2017**, *22*, 1–24. [CrossRef]

5. Ruijer, E.H.; Martinius, E. Researching the democratic impact of open government data: A systematic literature review. *Inf. Polity* **2017**, *22*, 233–250. [CrossRef]

6. Villazon-Terrazas, B.; Garcia-Santa, N.; Ren, Y.; Faraotti, A.; Wu, H.; Zhao, Y.; Vetere, G.; Pan, J.Z. Knowledge Graph Foundations. In *Exploiting Linked Data and Knowledge Graphs in Large Organisations*; Pan, J.Z., Vetere, G., Gomez-Perez, J.M., Wu, H., Eds.; Springer: Cham, Switzerland, 2017; pp. 17–55. [CrossRef]

7. Studer, R.; Benjamins, V.R.; Fensel, D. Knowledge engineering: Principles and methods. *Data Knowl. Eng.* **1998**, *25*, 161–198. [CrossRef]

8. Singhal, A. Introducing the Knowledge Graph: Things, Not Strings. 2012. Available online: https://www.blog.google/products/search/introducing-knowledge-graph-things-not (accessed on 15 September 2019).

9. Ehrlinger, L.; Wöß, W. Towards a Definition of Knowledge Graphs. *Semantics (Posters Demos SuCCESS)* **2016**, *48*, 1–4.

10. Gomez-Perez, J.M.; Pan, J.Z.; Vetere, G.; Wu, H. Enterprise knowledge graph: An introduction. In *Exploiting Linked Data and Knowledge Graphs in Large Organisations*; Springer: Cham, Switzerland, 2017; pp. 1–14.

11. Paulheim, H. Knowledge graph refinement: A survey of approaches and evaluation methods. *Semant. Web* **2017**, *8*, 489–508. [CrossRef]

12. Bonatti, P.A.; Decker, S.; Polleres, A.; Presutti, V. Knowledge Graphs: New Directions for Knowledge Representation on the Semantic Web (Dagstuhl Seminar 18371). In Proceedings of the Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, London, UK, 3–6 June 2019.

13. Yan, J.; Wang, C.; Cheng, W.; Gao, M.; Zhou, A. A retrospective of knowledge graphs. *Front. Comput. Sci.* **2018**, *12*, 55–74. [CrossRef]

14. Corradi, A.; Foschini, L.; Ianniello, R. Linked data for open government: The case of Bologna. In Proceedings of the 2014 IEEE Symposium on Computers and Communications (ISCC), Madeira, Portugal, 23–26 June 2014; pp. 1–7.

15. Jaakola, A.; Kekkonen, H.; Lahti, T.; Manninen, A. Open data, open cities: Experiences from the Helsinki Metropolitan Area. Case Helsinki Region Infoshare www. hri. fi. *Stat. J. IAOS* **2015**, *31*, 117–122. [CrossRef]

16. Krishnamurthy, R.; Awazu, Y. Liberating data for public value: The case of Data. gov. *Int. J. Inf. Manag.* **2016**, *36*, 668–672. [CrossRef]

17. Simperl, E.; Corcho, O.; Grobelnik, M.; Roman, D.; Soylu, A.; Ruíz, M.J.F.; Gatti, S.; Taggart, C.; Klima, U.S.; Uliana, A.F.; et al. Towards a Knowledge Graph Based Platform for Public Procurement. In Proceedings of the Research Conference on Metadata and Semantics Research, Limassol, Cyprus, 23–26 October 2018; pp. 317–323.

18. Shadbolt, N.; O'Hara, K.; Berners-Lee, T.; Gibbins, N.; Glaser, H.; Hall, W. Linked open government data: Lessons from data. gov. uk. *IEEE Intell. Syst.* **2012**, *27*, 16–24. [CrossRef]

19. Baker, M. Canada's growing open data movement. *Feliciter* **2011**, *57*, 96–98.

20. Chatfield, A.T.; Reddick, C.G. A longitudinal cross-sector analysis of open data portal service capability: The case of Australian local governments. *Gov. Inf. Q.* **2017**, *34*, 231–243. [CrossRef]

21. Fernández-Ruiz, M.J.; Alonso, J.M.; ALvarez, M.; Morlán-Plo, V.; Perez, M.J.P.; Zarazaga-Soria, F.F.J. La Política de Datos Abiertos del Ayuntamiento de Zaragoza: datosabiertos. zaragoza. es. In Proceedings of the Actas de CAEPIA, San Cristóbal de la Laguna, Tenerife, Spain, 7–10 November 2011; pp. 7–10.

22. Fernández-Ruiz, M.J.; Zarazaga-Soria, F.J. *IDEZar: La Infraestructura de Datos Abiertos Espaciales del Ayuntamiento de Zaragoza (2004-)*; Servicio de Publicaciones; Universidad de Zaragoza: Zaragoza, Spain, 2017; pp. 10–26.

23. Denaux, R.; Ren, Y.; Villazon-Terrazas, B.; Alexopoulos, P.; Faraotti, A.; Wu, H. Knowledge Architecture for Organisations. In *Exploiting Linked Data and Knowledge Graphs in Large Organisations*; Pan, J.Z., Vetere, G., Gomez-Perez, J.M., Wu, H., Eds.; Springer: Cham, Switzerland, 2017; pp. 57–84. [CrossRef]

24. Villazón-Terrazas, B.; Vilches-Blázquez, L.M.; Corcho, O.; Gómez-Pérez, A. Methodological Guidelines for Publishing Government Linked Data. In *Linking Government Data*; Wood, D., Ed.; Springer: New York, NY, USA, 2011; pp. 27–49. [CrossRef]

25. Weibel, S.; Kunze, J.; Lagoze, C.; Wolf, M. Dublin core metadata for resource discovery. *Internet Eng. Task Force RFC* **1998**, *2413*, 132.

26. Guha, R.V.; Brickley, D.; Macbeth, S. Schema. org: Evolution of structured data on the web. *Commun. ACM* **2016**, *59*, 44–51. [CrossRef]

27. Muñoz-Soro, J.F.; Esteban, G.; Corcho, O.; Serón, F. PPROC, an ontology for transparency in public procurement. *Semant. Web* **2016**, *7*, 295–309. [CrossRef]

28. Berrueta, D.; Phipps, J.; Miles, A.; Baker, T.; Swick, R. Best Practice Recipes for Publishing RDF Vocabularies. Available online: https://www.w3.org/TR/swbp-vocab-pub (accessed on 1 October 2019).

29. Spain, M.o.F.P.A. Technical Interoperability Standard for the Reuse of Information Resources. 2013 Available online: https://administracionelectronica.gob.es/pae_Home/dam/jcr:a8d2c143-ce9a-4fc7-afe7-ef5d9ba7c4a1/ENGLISH_Interoperability_Agreement_for%20the%20Reuse%20of%20Information%20Resources.pdf (accessed on 1 October 2019).

30. Soylu, A.; Elvesæter, B.; Turk, P.; Roman, D.; Corcho, O.; Simperl, E.; Konstantinidis, G.; Lech, T.C. Towards an Ontology for Public Procurement based on the Open Contracting Data Standard. In Proceedings of the Conference on e-Business, e-Services and e-Society, Trondheim, Norway, 18–20 September 2019; pp. 230–237.

31. Zaveri, A.; Rula, A.; Maurino, A.; Pietrobon, R.; Lehmann, J.; Auer, S. Quality assessment for linked data: A survey. *Semant. Web* **2016**, *7*, 63–93. [CrossRef]