Research Article

SREQP: A Solar Radiation Extraction and Query Platform for the Production and Consumption of Linked Data from Weather Stations Sensors

José Luis Sánchez-Cervantes,[1] Mateusz Radzimski,[2] Cristian Aaron Rodriguez-Enriquez,[1] Giner Alor-Hernández,[1] Lisbeth Rodríguez-Mazahua,[1] Cuauhtémoc Sánchez-Ramírez,[1] and Alejandro Rodríguez-González[3]

[1]*Division of Research and Postgraduate Studies, Instituto Tecnológico de Orizaba, Avenida Oriente 9, 852 Colonia Emiliano Zapata, 94320 Orizaba, VER, Mexico*
[2]*Computer Science Department, Universidad Carlos III de Madrid, Avenida Universidad 30, Leganés, 28911 Madrid, Spain*
[3]*ETS de Ingenieros Informáticos, Universidad Politécnica de Madrid, Campus de Montegancedo, Boadilla del Monte, 28660 Madrid, Spain*

Correspondence should be addressed to Giner Alor-Hernández; galor@itorizaba.edu.mx

Nowadays, solar radiation information is provided from sensors installed in different geographic locations and platforms of meteorological agencies. However, common formats such as PDF files and HTML documents to provide solar radiation information do not offer semantics in their content, and they may pose problems to integrate and fuse data from multiple resources. One of the challenges of sensors Web is the unification of data from multiple sources, although this type of information facilitates interoperability with other sensor Web systems. This research proposes architecture SREQP (Solar Radiation Extraction and Query Platform) to extract solar radiation data from multiple external sources and merge them on a single and unique platform. SREQP makes use of Linked Data to generate a set of triples containing information about extracted data, which allows final users to query data through a SPARQL endpoint. The conceptual model was developed by using known vocabularies, such as SSN or WGS84. Moreover, an Analytic Hierarchy Process was carried out for the evaluation of SREQP in order to identify and evaluate the main features of Linked-Sensor-Data and the sensor Web systems. Results from the evaluation indicated that SREQP contained most of the features considered essential in Linked-Sensor-Data and sensor Web systems.

1. Introduction

Sensor Web technology is an active field of research. It consists in the unification of the potential of the Internet as a global network with the capacity of sensor networks to real-time monitor certain parameters, such as solar radiation, pressure, humidity, temperature, and wind in different environment [1]. Furthermore, sensor Web favors the analysis of data from heterogeneous sensors and the development of simulators, models, and tools to support decision-making [2]. Sensor Web has an important application field in solar radiation, since this phenomenon has gained continuous attention. Solar radiation is one of several factors that cause human skin diseases. Its effects on humans are primarily due to increases in biologically effective Ultraviolet Radiations (UVR), mostly Ultraviolet Radiations of B-type (UVB), producing cataracts, skin cancer, and possible effects on immune responses [3].

Nowadays, a wide variety of computer systems have been developed to provide information on solar radiation in order to take preventative measures for skin health care. However, this information tends to be imprecise for different motives. For instance, data source employed may not contain updated information and could gather insufficient or irrelevant data. Also, data acquisition might lack real-time mechanisms to measure changes in solar radiation.

Therefore, the use of Linked Data [4] allows for the semantic generation and publication of datasets of solar radiation data.

Authors Corcho and García-Castro [5] identified five challenges in the area of sensor Web that relate to the characteristics of data sources handled in typical sensor Web applications. First of all, sensor data can, in general, be obtained, processed, and managed. Sensors can be managed through high-level formalisms, such as declarative continuous queries over streams, thereby insulating clients and users from the infrastructural and syntactic heterogeneities of autonomously deployed sensor networks.

Another challenge is related to the adequate characterization and management of quality of sensor data (quality of service). Issues such as the unavailability of a piece of data over a period of time may be the result of many factors, such as an unavailable sensor, possible inexistence of events to trigger data generation during that time, or damaged communication with the sensor. On the other hand, the third challenge in sensor Web is associated with the integration and fusion of data coming from autonomously deployed sensor networks with varying qualities of service and different throughput rates and geographical scales. However, this concerns not only the integration of data coming from different sensor networks but also the combination of such data with data that persisted in other sources, such as static data or archived sensor data.

The fourth challenge of utmost importance is related to the previous one and concerns the identification and location of relevant sensor-based data sources with which data integration and fusion tasks can be performed. Finally, the fifth and last challenge refers to the need of rapid development of applications able to handle sensor data by considering the aforementioned characteristics and challenges. This includes dealing with data integrity and validation issues as well as the need for common interfaces and formats among applications, databases, and sensor networks, to mention but a few.

Linked Data aims to help revolutionize the world of data access, more precisely the reuse of data [6]. The approach proposed in this research targets the field of weather data and mainly contributes as an alternative to help overcome challenges identified by Corcho and García-Castro [5] concerning the characteristics of data sources handled in typical sensor Web applications. For instance, an alternative solution to overcome the first challenge may be a SPARQL endpoint in SREQP used by users to query solar radiation data from sensors, so that they are insulated from the infrastructural and syntactic heterogeneities of autonomously deployed sensor networks.

The third challenge could be overcome with SREQP, since it allows for the integration and fusion of data from multiple sources (sensors, platforms) and from autonomously deployed sensor networks. The fourth challenge could also be overcome with the SREQP proposed in this research, since SREQP uses the location property of the WGS84 vocabulary to provide information on the geographic location, such as city or province (latitude, longitude, and altitude), of the weather stations from which data will be unified. Also, SREQP reuses ontologies, vocabularies, and instances such as MUO to represent measurement values combined with UCUM instances to represent physical units of measurement. Several works have addressed knowledge representation with ontologies or semantic technologies in related environments such as ambient intelligence [7, 8], smart homes [9], and pervasive computing [10, 11] as an introductory area of interest related to sensors. Finally, data stored in SREQP can be reused to develop Web and mobile applications, which provides an alternative solution to the fifth challenge. Solutions to overcome the second issue are not proposed, since they concern the quality of service of information sources and include issues, such as unavailability of sensors, lack of data over a period of time, and lack of communication with the sensor.

The development of SREQP arises from the need to produce Linked Data with solar radiation data extracted from the Spain State Meteorological Agency (AEMET) to provide users with more accurate information for skin care in the different geographical areas of Spain. Such information on solar radiation can be provided by AEMET after the payment of access fees.

Linked Data brings several benefits to SREQP such as (1) data portability across current datasets that allows for the constant update of information from sensors Web, (2) platform-independent data and information access, and (3) the fact that dataset is distributed for the Web. These two last benefits enable reusing information from sensors located in different geographical locations without directly accessing them.

Similarly, Linked Data also allows for the application of collective intelligence and inference on the published dataset. This provides support to sensor Web, since it facilitates the semantic analysis of solar radiation data in order to generate semantic systems for weather prediction and semantic recommendation systems. Another important feature of SREQP and its semantic model is the use of common and well-adopted semantic vocabularies and taxonomies. The resulting dataset can be easily federated or integrated with external data. Furthermore, establishing "sameAs" relations between corresponding concepts of different datasets impacts on the scope of data-driven integration among other weather monitoring systems.

The remainder of the paper is organized as follows. Section 2 introduces the state of the art according to the objective of this work, while Section 3 describes the architecture and functionality of each component of SREQP. On the other hand, Section 4 describes the vocabulary and radiation information in Linked Data developed, and Section 5 concerns the description to access the RDF triplestore through SPARQL-based queries. Finally, discussions and conclusions and future work are presented in Section 6.

2. Related Work

Great amount of research has studied forms of obtaining information through sensor Web and its application in different fields. Some of these works have obtained outstanding results by using semantic technologies. Therefore, these initiatives reviewed for this research were classified into two

categories: (1) models and architectures for sensors Web data and (2) sensors Web systems.

2.1. Models and Architectures for Sensors Web Data. According to Pouchard et al. [12], a considerable amount of scientific data are not stored on the cloud or on the Web. They are rather stored in multi-institutional data centers that provide tools and add value through its quality assurance, validation, curation, dissemination, and analysis. These authors proposed a scenario of river-channel transportation that required biogeochemical experimental data and global climate-simulation model data from many sources in order to publish, share, and link scientific data and processes end-to-end loosely coupled workflows, which would allow them to share and reuse scientific data. Therefore, Pouchard and collaborators focused on the use of ontologies and formal machine-readable descriptions of the domain to facilitate the search and discovery of this data. The present research focuses on extraction, dissemination, and analysis of solar radiation data. Similarly, Barnaghi et al. [13] described a semantic modeling scheme, a naming convention, and a data distribution mechanism for sensor streams. They proposed solutions that addressed important challenges, such as the increase of sensor streams and the observation and measurement that data provided via these streams. This initiative of authors enabled dealing with large-scale sensor data emerging from the Internet of Things resources. Results showed that the proposed solutions could scale for large number of sensor streams with different types of data and various attributes. Sensor networks have been considered as a major source of information for Digital Earth, which demands highly dynamic information systems, new sources of information, and stronger capabilities for their integration. Authors Janowicz et al. [14] introduced a Linked Data model and a RESTful proxy for OGC's (Open Geospatial Consortium) Sensor Observation Service to improve integration and interlinkage of observation data for the Digital Earth. Similarly, authors Vilches-Blázquez et al. [15] proposed a process to generate geographical Linked Data from four Infrastructure for Spatial Information in the European Community (INSPIRE) (European Commission 2007) themes. The main lessons learnt were they that could be extrapolated to similar integration processes of geographical information. The main goal of their process was to combine different sources (heterogeneous, multidisciplinary, multitemporal, multiresolution, and multilingual) using Linked Data principles to solve current problems of information integration and direct geographical information toward the next decade scenario, that is, "Linked Digital Earth." Malewski et al. [16] described StarFL, a new modularised metadata language for sensor descriptions. This language followed a more restrictive approach and incorporated concepts from the recently published Semantic Sensor Network Ontology to overcome key issues that users experienced with SensorML. However, unlike Barnaghi et al. [13], Vilches-Blázquez et al. [15], Janowicz et al. [14], and Malewski et al. [16], the platform proposed in this research designs a conceptual data model inspired by the principles of Linked Data [17]. The goal is to generate a set of RDF triples containing information on

solar radiation data extracted from external sensors (e.g., pyranometers). Furthermore, other authors have reused and analyzed solar radiation data stored in the set of triples. For instance, the work of Compton et al. [18] provided the following two contributions to demonstrate the usefulness of SSN ontology: (1) a description of the SSN ontology produced by the W3C Semantic Sensor Network Incubator Group (the SSN-XG) (http://www.w3.org/2005/Incubator/ssn/) which is based on Web Ontology Language 2 (OWL 2) to describe sensors and observations and (2) an example of the use of the SNN ontology and the discussion of projects and applications in which it has been used. Authors concluded that SSN ontology could describe sensors sensing the measurement capabilities of other sensors, resulting observations, and deployments in which sensors are used. The ontology covers large parts of the SensorML and O&M standards, omitting calibrations, process descriptions, and data types. On the other hand, the initiative proposed in the present research proposes an internal taxonomy to represent different types of observations. This taxonomy extends SSN (Semantic Sensor Network) ontology and AWS (ontology for meteorological sensors) ontologies by providing concrete subclasses and detailing aspects of sensors, such as platforms where sensors are collocated, type of sensors, observations, and measures. Also, authors Zaslavsky et al. [19] discussed several aspects of sensors, such as the architecture of the emerging Internet of Things (IoT), applications of large-scale sensor network, federating sensor networks, sensors data and capturing techniques of related context, and challenges in cloud-based management, as well as the storage, archiving, and processing of sensors data. Authors concluded that data streams coming from these devices would challenge traditional approaches for data management and contribute to the emerging paradigm of Big Data. Moreover, connecting sensing devices present in the physical world to detect and measure various physical phenomena (e.g., temperature, humidity, and pollution) and presenting them as Web resources to the end-users possess a heterogeneous nature; thus, most sensor Web studies focus on providing domain-specific solutions. For instance, Khan and Kim [20] introduced an improved SOA-based sensor Web architecture. This architecture provided an easy approach to integrate sensor services providers with information services providers and enable users to access it as a single, integrated, and searchable service. Another new architecture was designed by Babovic and Milutinovic [21] as an infrastructural platform to enable the integration of semantic-based sensor networks. The key idea behind this proposed design was to utilize a flexible distributed repository called column store to keep semantically modeled sensor data and provide a scalable platform capable of supporting huge amounts of sensor data and large numbers of users. Unlike Babovic and Milutinovic [21] and Khan and Kim [20] the present research paper proposes an architecture to extract solar radiation information from different external sources (sensors, databases, and platforms) and merge it on a single and unique platform based on the principles of Linked Data [17]. Table 1 shows a brief comparative summary of the discussed models and architectures for sensors Web data.

TABLE 1: Comparison of analyzed models and architectures for sensor Web data.

Research	Model or architecture as alternative solution	SREQP
A great amount of scientific data are not stored in the cloud or on the Web, but in multi-institutional data centers [12].	Scenario focused on the use of ontologies and formal machine-readable descriptions of domain to facilitate search and discovery of this scientific data.	Focused on the extraction, dissemination on Linked Data format, and analysis of solar radiation data from pyranometers.
Authors addressed important challenges such as the increase of sensor streams (Internet of Things) and the observation and measurement that data provided via these streams [13].	Semantic modeling scheme, a naming convention, and a data distribution mechanism for sensor streams.	
Authors considered that sensor networks are a major source of information for the vision of Digital Earth, which requires more dynamic information systems and stronger capabilities for their integration [14].	Linked Data model and RESTful proxy for the OGC's Sensor Observation Service to improve integration and interlinkage of data observation for the Digital Earth.	The design of a conceptual data model inspired by the Linked Data principles [17] with the aim of generating a set of RDF triples containing information about solar radiation data extracted from multiple sources, such as sensors, platforms, and databases.
Authors considered that it was necessary to overcome current problems of information integration and direct geographical information to the next decade scenario, that is, "Linked Digital Earth" [15].	Process based on Linked Data principles to combine different sources (heterogeneous, multidisciplinary, multitemporal, multiresolution, and multilingual) to enable the integration of geographical information.	
A language needed to be developed to help overcome key issues that users experience with SensorML language [16].	StarFL language presents a more restrictive approach and incorporates concepts from the recently published Semantic Sensor Network Ontology.	
Authors addressed the benefits of using SSN ontology produced by the W3C Semantic Sensor Network Incubator Group [18].	The SSN ontology can describe sensors sensing the measurement capabilities of other sensors, resulting observations, and deployments in which sensors are used.	An internal taxonomy that extends SSN and AWS ontologies by providing concrete subclasses and detailing aspects, observations, and measures of sensors.
Authors analyzed challenges in the Internet of Things architecture and those concerning application of large-scale sensor networks, federating sensor networks, sensors data and related context capturing techniques, as well as sensors cloud-based management [19].	Data streams coming from connected devices to the Internet of Things will challenge the traditional approaches to data management and will contribute to the emerging paradigm of Big Data.	
Sensors in the physical world should be connected into a network to detect and measure various physical phenomena (e.g., temperature, humidity, and pollution). Then, they should be introduced as Web resources to the end-users [20].	A SOA-based sensor Web architecture providing an easy approach to integrate sensor providers' services with information provider services and enable the users to access it as a single, integrated, and searchable service.	It is architecture to extract solar radiation information from different external sources and merge it on a single and unique platform inspired by the Linked Data principles [17]. Furthermore, this initiative provides access to solar radiation data stored in a Linked Data Repository through a SPARQL endpoint to query all data stored in the platform.
Authors believed that an infrastructural platform should be designed to enable the integration of semantic-based sensor networks [21].	A new architecture design that enabled the integration of semantic-based sensor networks. This architecture provided a scalable platform capable of supporting huge amounts of sensor data and large numbers of users.	

2.2. Sensors Web Systems. Authors Crowley et al. [22] proposed a framework to integrate and link heterogeneous data from various sources and transform them into Linked Data. This framework allows for the reuse and integration of produced data with other data resources, which enables spatial business intelligence for various domain-specific applications. Similarly, [23] discussed the use of FlexFT, which is a generic component-based framework for the construction of adaptive fault tolerant systems that can integrate and reuse technologies and deploy them across heterogeneous devices. FlexFT provided a standardized and interoperable interface for sensor observations by relying upon the "Sensor Web" paradigm established by the Open Geospatial Consortium (OGC). Beder et al. [23] implemented a Java prototype to the framework and the authors evaluated the potential benefits through case studies and performance measurements. By implementing and deploying these case studies in standard PCs as well as in sensor nodes, authors showed that FlexFT could cope high heterogeneity with minimal resource overheads. Authors Corno and Razzak [24] proposed "LO(D)D" a distributed framework that enabled systematic publishing of environment data that was continuously updated. Such updates could be issued at specific time intervals or bound to some environment specific event. Similarly, the framework targeted smart environments with networks of devices and sensors that interacted with one another and with their respective environments to gather, generate, and publish data. On the other hand, Stocker et al. [25] discussed a generic software framework for the organization and interpretation of sensor data. They demonstrated its application on data of a large-scale sensor network to monitor atmospheric phenomena. Results indicated that software support for the organization and interpretation of sensor data are valuable to scientists in scientific computing workflows. Regueiro et al. [26] analyzed the design, implementation, and evaluation of a framework that enabled the virtual integration of heterogeneous observation data sources through a Sensor Observation Service (SOS) standardized interface. The framework is being currently validated by the OGC compliant technology to publish the meteorological and oceanographic observation data generated by two public agencies of the regional government of Galicia (northwest of Spain). Another sensor Web system is the Sense2Web, a platform to publish Linked-Sensor-Data for the sensor network community. The platform was presented by Barnaghi et al. [27] and its main focus is to define an approach to enrich the sensor descriptions data. It enables publishing sensor description data of users as RDF triples, associating them with any other existing RDF sensor description data, linking them with the existing resources on publicly available Linked Data repositories, and making them available to consumers using SPARQL endpoint. In addition, SEIPF (Semantic Energy Information Publishing Framework) [28] is a semantic energy information publishing framework that provided the ability to query energy consumption information from residential gateways in a machine understandable format in order to achieve consumption coordination and intelligent negotiation. The framework is based on the client-server model and when queried by the client, it provides information regarding the energy consumption of the residency based on an Energy Profile ontology. The framework proposed in the present research also incorporates the lesson learned from the development of SEIPF. It enables moving from a client-server model to a publisher-subscriber pattern and allows for the separation of static information about the environment (using PID) and dynamic updates (using channels). Also, authors Yu and Liu [29] designed, developed, and implemented a system to achieve better data interoperability and integration by republishing real-world data into linked geosensor. The contributions of authors included (1) best practices of reusing and matching the W3C Semantic Sensor Network (SSN) ontology and other popular ontologies for heterogeneous data modeling in the water resources application domain; (2) a newly developed spatial analysis tool to create links; and (3) a set of RESTful OGC Sensor Observation Services (SOS) as Linked Data APIs. Results indicated that a linked sensor Web could be built and used within the integrated water resource application domain. Finally, the use of sensor Web was reported as an important component to obtain data directly from data sources. Authors Corcho and García-Castro [5] addressed some of the existing challenges in the area of the sensor Web related to the characteristics of data sources handled in typical sensor Web applications. Authors also discussed additional challenges on the creation of applications based on these data sources. Similarly, Atemezing et al. [30] proposed a system named AEMETLinkedData for the practical transformation of meteorological data into Linked Data. This was achieved by triplicating sensor data and extending SSN ontology to cover various meteorological observations. Data was stored in an RDF repository and visualized on the map by using sensor coordinate system. Moreover, the study planned to incorporate GeoLinkedData for further data integration.

This categorization has been a useful resource to interpret some open problems in this area. Table 2 provides a brief comparative summary of the aforementioned sensor Web systems.

Certain differences between these initiatives and the proposed SREQP architecture are the facts that (1) SREQP allows for the extraction and consumption of solar radiation data from external data sources, (2) it is based on Linked Data principles, and hence (3) it encourages the Use of Standards for data publishing and consumption, such as ontologies, taxonomies, and vocabularies; similarly, (4) it leads to the generation of a platform that can be reused by other users for their own interests, including the analysis of solar radiation data through the use of SPARQL endpoint of the platform.

3. Architecture and Components of SREQP

This section addresses the architecture of SREQP (Solar Radiation Extraction and Query Platform) and its main components. The platform was developed with the aim of extracting solar radiation information from external sources and merging it on a single and unique platform. SREQP is composed of several functional modules that carry out different tasks to perform the extraction and conversion processes that should be executed to prepare data and make

TABLE 2: Comparison of analyzed sensor Web systems.

| System type	Features
Framework to integrate and link heterogeneous data from various sources [22].	(i) It allows for the reuse, integration, and transformation of data in Linked Data format from several data sources. (ii) It enables spatial business intelligence for various domain-specific applications.
FlexFT Framework [23]	(i) It provides a standardized and interoperable interface for sensor observations by relying upon the "Sensor Web" paradigm established by the OGC. (ii) This framework has been implemented under a prototype developed in Java. (iii) Flex can cope with high heterogeneity of data with minimal resource overheads.
Distributed framework that provided systematic publishing of environment data [24].	(i) It targets smart environments having networks of devices and sensors that interact with one another. (ii) This framework allows for the publishing of environment data that is continuously updated. (iii) The environment data updated might be issued at specific time intervals or bound to some environment specific.
Generic software framework for the organization and interpretation of sensor data [25].	(i) It allows for the organization and interpretation of data of a large-scale sensor network to monitor atmospheric phenomena. (ii) This framework is useful to client software systems since it can be queried, integrated, reasoned, visualized, or annotated.
Framework that enables virtual integration of heterogeneous observation data sources [26].	(i) The virtual integration of heterogeneous observation data sources is enabled through a Sensor Observation Service standardized interface. (ii) The framework is being validated as the OGC compliant technology to publish meteorological and oceanographic observation data generated by two Spanish public agencies.
Sense2Web platform [27].	(i) It enriches the description of the sensors data. (ii) It allows users to publish their sensor description data as RDF triples and associate them with any other existing RDF sensor description data. (iii) It generates links to the existing resources on publicly available Linked Data repositories and makes them available to consumers by using SPARQL endpoint.
SEIPF Framework [28].	(i) This framework provides the ability to query energy consumption information from residential gateways in a machine understandable format to achieve consumption coordination and intelligent negotiation. (ii) It is based on the client-server model. When it is queried by the client, the framework provides information concerning the energy consumption of the household based on an Energy Profile ontology.
System to achieve better data interoperability and integration by republishing real-world data into linked geosensor data [29].	(i) This system relies on the best practices of reusing and matching the SSN ontology and other popular ontologies for heterogeneous data modeling in the water resources application domain. (ii) It provides a spatial analysis tool to create links. (iii) It provides a set of RESTful OGC Sensor Observation Service like Linked Data APIs.
AEMETLinkedData [30]	(i) The system transforms meteorological data into RDF triples. (ii) It reuses SSN to model sensor description and other upper ontologies (OWL Time, GeoBuddies). (iii) It uses geographical location to plot meteorological results on the map with Map4rdf* tool.
SREQP	(i) SREQP enables the solar radiation data extraction from different external sources, mainly sensors, platforms, and databases. (ii) It transforms and unifies solar radiation data extracted in order to generate and publish a Linked Data repository. (iii) The platform is reusable for users. All data stored in the repository can be obtained through SPARQL-based queries and using the SPARQL endpoint. (iv) The internal taxonomy (SOLRAD Taxonomy) of SREQP is also reusable. The taxonomy extends SSN and AWS ontologies in order to reuse recommended ontologies by W3C consortium and provide detailing aspects, observations, and measures of sensors.

*https://github.com/oeg-upm/map4rdf.
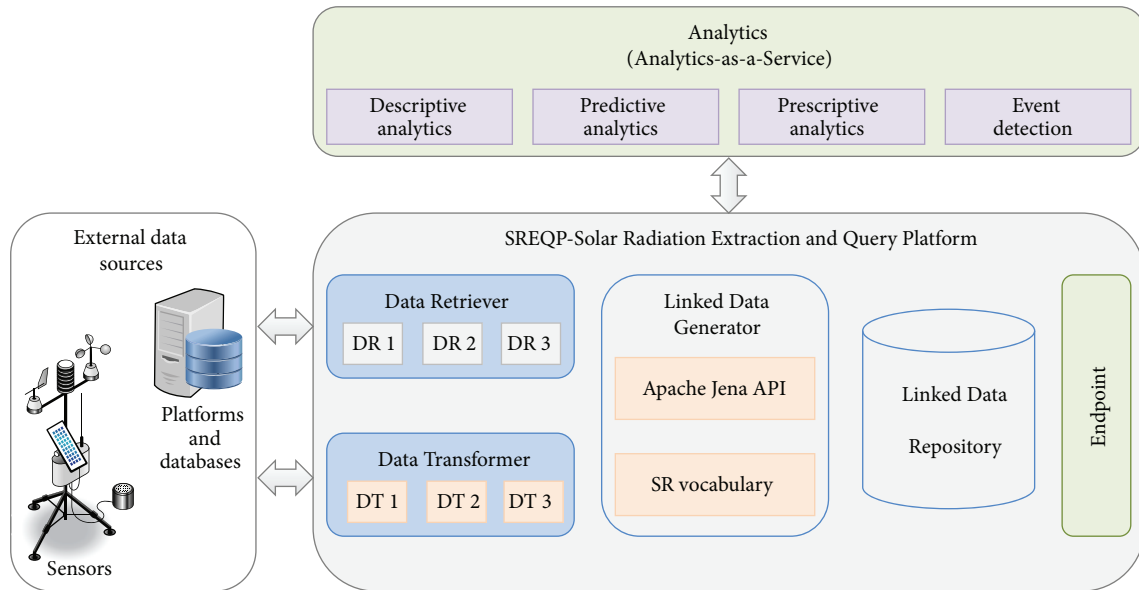
FIGURE 1: SREQP architecture.

them accessible. The current version of the platform has been developed to extract solar radiation data from legacy Spanish State Meteorological Agency (AEMET) (http://www.aemet.es/es/eltiempo/observacion/radiacion/ultravioleta?datos=tabla) data repositories. However, the system has been designed to process different types of external sources.

Figure 1 illustrates the architecture of SREQP.

The different modules of SREQP architecture are detailed below:

(i) External Data Sources (Sensors). This component includes the different external sources, such as sensors, platforms, and databases that provide information on solar radiation. Every platform can provide solar radiation data in different formats and all platforms can be accessed through different ways (Web services, ftp access, and database access, to mention but a few). The current data source is legacy AEMET data repositories.

(ii) Data Retriever. This module is the first component that must be executed to extract the information from external data sources (sensors, platforms, or databases). This component consists of a Java-based crawler integrated by several Data Retrievers (DR1 to DR*n*) (see Figure 1). Each Data Retriever is executed to track and download information from the data sources depending on the type of access or format for further transformation in Linked Data format. This module allows for the implementation of new Data Retrievers for every source.

(iii) Data Transformer. This module is executed to convert information obtained from each Data Retriever component (DR1 to DR*n*) into the format used by the platform, that is, Java Objects. Java Objects generated contain all the solar radiation information retrieved; such information would be converted into Linked

Data format through the execution of the Linked Data Generator component. Concrete Data Transformers are necessary for each Data Retriever.

(iv) Linked Data Generator. The execution of this component allows for the generation of RDF triples from information contained in the Java Objects generated by the Data Transformer component. To generate these RDF triples, the Linked Data Generator component uses the SOLRAD Taxonomy as data model (see Section 4) and methods from Apache Jena API [31] such as (1) jena.rdf.model to create and manipulate RDF graphs; (2) jena.datatypes that provide the core interfaces through which data types are described to Jena and; and (3) jena.rdf.arp which is the parsing subsystem in Jena to handle the RDF/XML syntax. Finally, RDF triples generated will be stored in a Linked Data Repository.

(v) Linked Data Repository. The execution of this component is closely related to the execution of the previous component, since RDF triples generated are serialized and stored on the Linked Data Repository Virtuoso Open Source [32]. Virtuoso Open Source was used since it seemed the most convenient platform to manage, access, and integrate Linked Data to support the RDF triples with solar radiation information generated. This decision was based on the results of benchmark tests for the execution of SPARQL-based queries which were performed by other authors, such as Bizer and Schultz [33] and Morsey et al. [34]. Authors compared various systems for managing data based on Linked Data and obtained results reported Virtuoso as the fastest platform.

(vi) Endpoint. The Endpoint module is a SPARQL endpoint provided by the Linked Data Repository, in this

case, Virtuoso Open Source. It allows external users to query all data stored (RDF triples) in the SREQP.

(vii) Analytics. The semantic sensor Web enables interoperability and advanced analytics for situation awareness and other advanced applications from heterogeneous sensors [35]. The analytics layer facilitates the generation of businesses objectives through data report in order to analyze trends. This creates predictive models to foresee future problems and opportunities and analyzes/optimizes business processes to enhance organizational performance [36]. From a taxonomical perspective, authors Delen and Demirkan [37] mentioned three main categories for analytics: descriptive, predictive, and prescriptive:

(1) Descriptive analytics: it is also called business reporting. It uses data to answer the questions "What happened and/or What is happening?" It includes simple standard/periodic business reporting, ad hoc/on-demand reporting as well as dynamic/interactive reporting (OLAP, slice/dice, drill-down/roll-up, etc.). The main output of descriptive analytics is the identification of business opportunities and problems.

(2) Predictive analytics: it uses data and mathematical techniques to discover explanatory and predictive patterns (trends, associations, affinities, etc.), which represent the inherent relationships among data inputs and outputs. Therefore, it responds to the questions "What will happen and/or Why will it happen?" Enablers of predictive analytics include data mining, text mining, Web/media mining, and statistical time-series forecasting. The main outcome of predictive modeling is an accurate projection of future events and the reasoning of why they occur.

(3) Prescriptive analytics: it uses data and mathematical algorithms to determine a set of high-value alternative courses of actions or decisions given a complex set of objectives, requirements, and constraints, with the goal of improving business performance. These algorithms may rely solely on data, expert knowledge, or a combination of both. Enablers of prescriptive analytics include optimization modeling, simulation modeling, multicriteria decision modeling, expert systems, and group support systems. The main outcome of prescriptive modeling is providing either the best course of action for a given situation or a rich set of information and expert opinions for decision-makers which could lead to the best possible course of action.

(4) Event detection: an event is an arbitrary classification of a space/time region. It might have actively participating agents, passive factors, products, and a location in space/time [38]. In the sensor networks context, event detection is one of the most important data services, since it is a form of reaching meaningful information out from the huge volume of data produced. It aims to find "right data" at the "right place" and ensures that data is sent at the "right time" [39].

(5) Analytics-as-a-Service: each category of analytics previously described can be provided as a service, that is, Analytics-as-a-Service (AaaS). This concept is often referred to as agile analytics and is fueled by the idea of turning utility computing and virtualization into a service model for data analytics [40]. Compared to Data- and Information-as-a-Service, Analytics-as-a-Service is a relatively newer concept. Management of complexity in models, development of service-based analytic models, and the standardization of interfaces among models are among the unique challenges that made Analytics-as-a-Service a late-emergent endeavor in information technology [37].

Although the aforementioned modules introduced the generic SREQP architecture, it is necessary to develop concrete modules to ensure that the platform obtains information from different data sources. Modules generated have several interfaces that define the methods and processes necessary to obtain information from the different data sources and convert such information.

This research considers AEMET (http://www.aemet.es/es/eltiempo/observacion/radiacion/ultravioleta?datos=tabla) Data Repository as the external data source. This repository is used as the case study to show the behavior of the system. AEMET is the State Meteorological Agency of Spain, functioning and sponsored by the Ministry of Agriculture, Food, and Environment. It represents the Spanish country among international meteorological institutions, such as the World Meteorological Organization (WMO), the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), and the European Centre for Medium-Range Weather Forecasts (ECMWF). The main task of AEMET is "to develop, implement, and provide meteorological services falling within the competences of the State, as well as to give support to other public and private activities which improve the safety and quality of life of the Spanish society." [41] The activities of AEMET include meteorological observations in Spain, the storage of these observations, weather monitoring and forecasting, and carrying out scientific research in numerical weather prediction models.

AEMET Data Repository contains information on solar radiation measures performed in different weather stations distributed throughout Spain (this information can be provided to users by AEMET after the payment of access fees). The different weather stations are composed of several weather sensors, including pyranometers, to measure information about solar radiation. AEMET publishes the information collected by its weather stations in the AEMET Data Repository, which is accessible through FTP protocol. The information is updated daily at 12:00 p.m. UTC. However, the main constraint of AEMET solar radiation publishing method is that information available is no older than a week.

It is impossible to obtain previous radiation information without charges. Thus, the Data Retriever module connects to AEMET Data Repository via FTP daily at 12:30 p.m. UTC to obtain last values collected from the different weather stations. The Data Transformer interface associated with AEMET is executed in order to process the CSV (Comma-Separated Values) file downloaded by the Data Retriever. Once the file is processed, the Linked Data Generator module is executed with all information provided by Data Transformer to generate the associated triples.

4. Representing Radiation Information as Linked Data

4.1. Semantic Model. Publishing datasets in form of Linked Data involves the adoption of Linked Data principles, which are [17]

  (1) using URIs as names for things,

  (2) using HTTP URIs so that people can look up those names,

  (3) that when someone looks up a URI, provide useful information, using the Semantic Web standards such as RDF and SPARQL,

  (4) including links to other URIs in order to discover more URIs.

Furthermore, publishing datasets implies the design of a conceptual data model to ensure further data reusability and interoperability. This step requires analysis of existing efforts, such as vocabularies and ontologies, in order to reuse them to model specific properties within the dataset. Using well-established vocabularies and ontologies should guarantee understandability, fewer problems with future data reuse, and higher uptake rate by data consumers.

AEMET dataset was modeled to represent the following "dimensions":

  (i) Sensor output—numerical value (measurement value) as registered by sensor during the measurement time.

  (ii) Sensor class—type of sensor performing measurement.

  (iii) Date and time of measurement—the time when measurement started and the total duration of the measurement.

  (iv) Observation type—physical property measured by sensor.

  (v) Location—geographical location of the weather station.

  (vi) Measurement units.

Note that in most cases these "dimensions" are related. This means that a particular physical property is measured with concrete device (or device class) using specified units of measurement.

AEMET provides typical observation-based dataset and it fits Stimulus-Sensor-Observation pattern, successfully adopted for modeling sensor data. This approach was followed in the creation of the Semantic Sensor Network (SSN) Ontology [42] within W3C SSN Incubator Group, which also aims at aligning previous efforts to represent sensor data. Therefore, the data model has been modeled reusing SSN ontology. In this case, only classes and properties relevant to aforementioned "dimensions" were considered. Although SSN covers a greater number of concepts related to sensor measurements, the model has been tailored to represent AEMET data in a practical way. It is important to mention that, in terms of modeling sensor-specific observations, our ontology subclassifies the SSN upper ontology in order to properly represent solar radiation domain, while keeping the taxonomy lean and focused on a convenient and practical knowledge representation.

On the other hand, SSN model only represents the Stimulus-Sensor-Observation aspect of the sensor data, and SREQP dataset still lacks means to represent other "dimensions." For instance, SSN model focuses on modeling sensor observations, while other domain concepts, such as time or geographical locations, are imported from other specialized ontologies. Therefore, it is necessary to reuse not one but various vocabularies, taxonomies, and ontologies in order to cover various aspects of sensor data. Thus, the dataset constructed within SREQP reuses the following vocabularies and ontologies: (i) OWL Time (http://www.w3.org/tr/owl-time/) to represent measurement start time and duration; (ii) MUO vocabulary (http://idi.fundacionctic.org/muo/) to represent values of measurement, combined with (iii) UCUM (http://idi.fundacionctic.org/muo/ucum-instances.html) to represent physical units of measurements; (iv) WGS84 (http://www.w3.org/2003/01/geo/) to represent basic geospatial data; and (v) AWS ontology (http://www.w3.org/2005/Incubator/ssn/ssnx/meteo/aws) to represent meteorological sensor's classes. Figure 2 depicts an example of single measurement in the SREQP model. Note that concepts without namespace explicitly stated belong to SOLRAD namespace.

An internal taxonomy (SOLRAD Taxonomy) has been constructed to represent different types of observations. The taxonomy was also constructed by using the ontology editor and framework to build intelligent systems Protégé [43, 44]. This taxonomy extends SSN and AWS ontologies by providing concrete subclasses and detailing such aspects as (i) platform where sensors are collocated; (ii) pyranometers (sensors) types; (iii) observations types; and (iv) observation results. The taxonomy structure is depicted in Figure 3.

4.2. Cross-Systems Data Integration and Data Fusion. Reusing common vocabularies, taxonomies, and ontologies makes the structure of the dataset understandable for people and machines not familiar with the dataset itself but at least with basic concepts. However, Linked Data does not merely involve the reduction of existing vocabularies; it is an effort to interconnect data. Therefore, when a concept in one dataset is equivalent to another included within a different dataset, it is possible to create a link between both concepts using owl:sameAs property. If this property exists, navigation from one dataset to another is possible, which facilitates data merging from both sources.
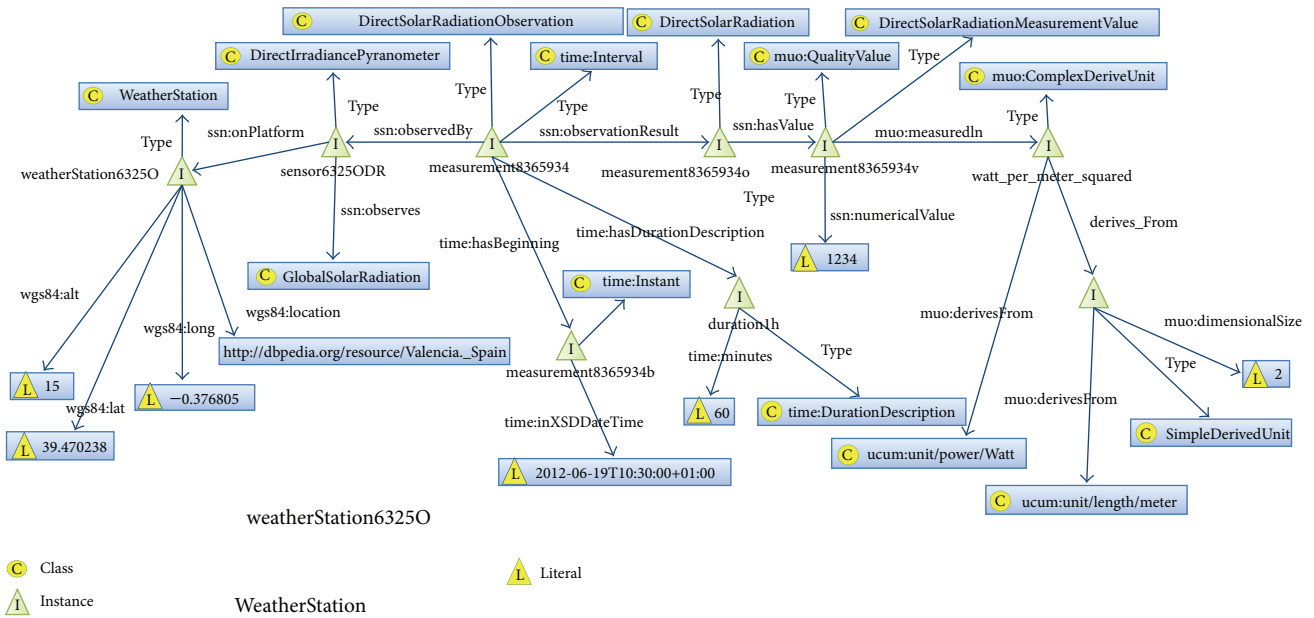
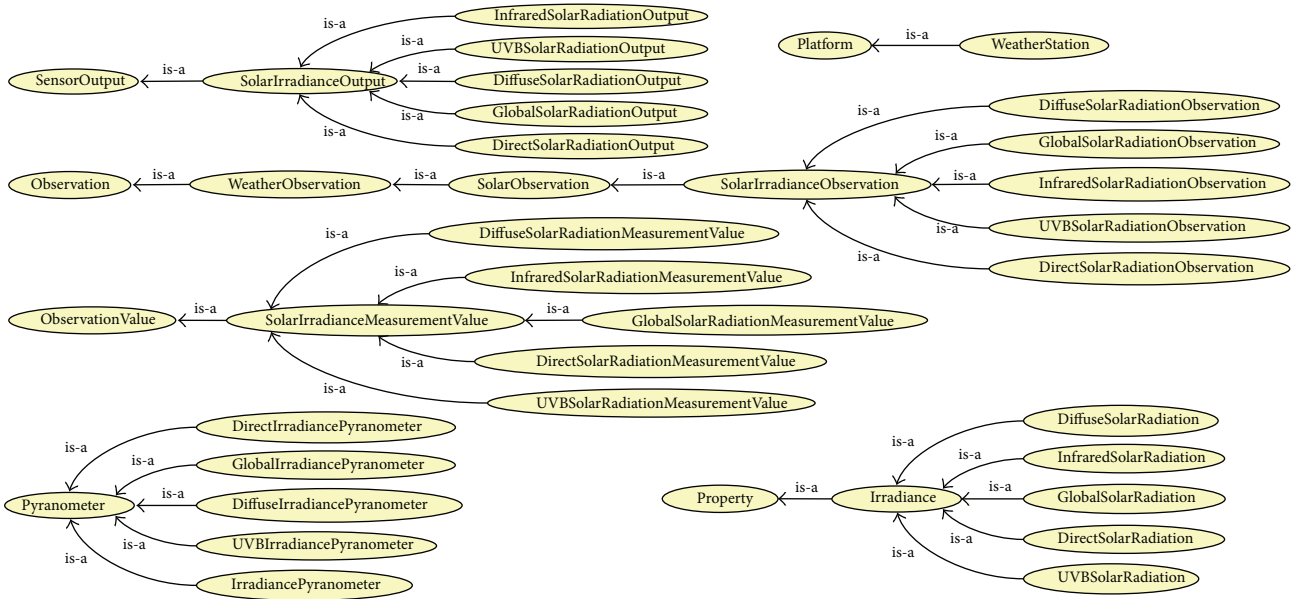Figure 2: Representation of a single measurement.

Figure 3: SOLRAD Taxonomy.

SREQP system provides data-driven integration capabilities based on semantic concept matching and by reusing common ontologies and taxonomies. It is based on three principal axes:

  (i) Spatial data integration.

  (ii) Temporal data integration.

  (iii) Integration based on the measured observation.

Spatial data integration is based on the SREQP model and can inform of locations of weather stations, such as city or province, using WGS84 location property. For instance, querying SREQP data with "wgs84:lat" and "wgs84:long" properties is a uniform approach to access sensors measures based on geospatial location. Moreover, WGS84 vocabulary is a standardized form of accessing other geospatial data in the Linked Open Data cloud and allows for further SPARQL query federation. This feature is especially powerful in combination with repositories, such as Virtuoso or GraphDB that support geospatial SPARQL queries through high-level built-in constructs, such as "nearby" or "within." This allows for a more flexible use of WGS84 properties. This also fosters

```
select ?quantityvalue
where {
  ?station wgs84:location "<http://dbpedia.org/resource/Barcelona>".
  ?sensor ssn:onPlatform ?station.
  ?sensor a solrad:DirectIrradiancePyranometer.
  ?measurement ssn:observedBy ?sensor.
  ?measurement ssn:observationResult ?result.
  ?result ssn:hasValue ?value.
  ?value ssn:hasQuantityValue ?quantityvalue.
  ?measurement time:hasBeginning ?beginning.
  ?beginning time:inXSDDateTime ?timedate
  FILTER (?timedate = "2015-06-19T13:00:00.00+01:00"^^xsd:#dateTime>)
}
```

LISTING 1: Sample SPARQL query.

data fusion across different datasets, either public—such as LinkedGeoData knowledgebase (http://linkedgeodata.org/)—or private that follow Linked Data approach.

Apart from raw geospatial locations of weather stations, geographical locations of sensors were also aligned with concepts from DBpedia [45]. As it is depicted in Figure 2, every sensor is provided with a link to DBpedia resource describing the location (city) of the sensor. This allows the platform to integrate greater amount of data in order to enrich the query capabilities, like population, postal code, basic weather statistics, DBpedia categories, and other useful metadata. Linking to DBpedia is also a principal motive for data integration in Linked Data. Every other dataset linked to the same concept can be directly used with SREQP data. The following subsection introduces a concrete example on how to link data from SREQP to DBpedia using SREQP system.

Also, temporal data integration is based on the use of OWL Time ontology. Similarly to the geospatial queries, the taxonomy to describe temporal aspects of measurements can be queried in a universal and widely adopted form, which provides a uniform access to measurement data. Time ontology provides means to query for observations that happened at concrete time or at specified time range interval.

Finally, as for the integration based on measurement types, apart from geospatial and temporal properties, the SREQP system characterizes measurement observation by (i) observation result type (in the sense of SSN ontology) and its corresponding (ii) measurement value and (iii) measurement value type (expressed in MUO concepts). In order to interlink datasets, the principal concept type must be correctly matched against the 3rd-party dataset. As of writing of this paper, there were no DBpedia concepts that could be linked against. Therefore, tools such as Silk [46] or LIMES [47] must be used to match these concepts.

4.3. Querying. SPARQL is a common query language to access data in RDF repositories (triplestores). Thus, SPARQL is to semantic repository as SQL is to relational databases. SREQP infrastructure provides two interfaces to access data (endpoints): HTTP endpoint for data navigation and SPARQL endpoint for data querying. The latter can be used by

GUI applications to retrieve relevant data and present them to end-user. For instance, Listing 1 depicts a sample query that returns the measure of direct solar radiation in Barcelona in June 19 at 1:00 p.m. UTC.

Using established links to DBpedia and federated query feature of SPARQL 1.1 (http://www.w3.org/TR/sparql11-federated-query/) enables querying both datasets (DBpedia and SREQP dataset) at the same time. This allows obtaining more possibilities for data querying by using DBpedia or YAGO (Yet Another Great Ontology) classifications (http://www.mpi-inf.mpg.de/departments/databases-and-information-systems/research/yago-naga/yago/). Listing 2 shows such example of federated query that returns sorted list of results for direct solar radiation at the specified hour for all the cities on the Spanish Mediterranean coastline.

5. Evaluation

Kitchenham et al. [48] proposed several, quantitative, qualitative, and hybrid evaluation methods to evaluate software and tools. The present research describes the term Feature Analysis as a qualitative evaluation. It also mentions that Feature Analysis was based on both identifying the requirements that users have for a particular task or activity and mapping these requirements to features that a method/tool aimed at supporting that task/activity should possess.

The Sensor Web Enablement (SWE) in the Open Geospatial Consortium (OGC) Inc. (http://www.opengeospatial.org) context refers to Web accessible sensor networks and archived sensor data that can be discovered, accessed, and where applicable controlled using open standard Application Programming Interfaces (APIs) [49]. The SWE provides a suite of standards that specify encodings to describe sensors and sensor observations and/or interface definitions for Web services. These standards are [49, 50] (1) Observations and Measurements (O&M) Schema; (2) Sensor Model Language (SensorML); (3) Transducer Markup Language (TransducerML or TML); (4) Sensor Observation Service (SOS); (5) Sensor Planning Service (SPS); (6) Sensor Alert Service (SAS); and Web Notification Services (WNS). Similarly, Sheth et al. [35] defined semantics of sensor Web within space,

```
SELECT DISTINCT ?city_name ?quantityvalue
FROM <http://nadir.uc3m.es/solrad/dataset.rdf>
WHERE {
    SERVICE <http://nadir.uc3m.es:8890/solrad/sparql> {
        ?station wgs84:location ?city.
        ?sensor ssn:onPlatform ?station.
        ?sensor a solrad:DirectIrradiancePyranometer.
        ?measurement ssn:observedBy ?sensor.
        ?measurement ssn:observationResult ?result.
        ?result ssn:hasValue ?value.
        ?value ssn:hasQuantityValue ?quantityvalue.
        ?measurement time:hasBeginning ?beginning.
        ?beginning time:inXSDDateTime ?timedate
        FILTER (?timedate = "2015-06-19T13:00:00.00+01:00"^^xsd:#dateTime>)
    }
    SERVICE <http://dbpedia.org/sparql> {
        ?city dcterms:subject category:Mediterranean_port_cities_and_towns_in_Spain;
          rdfs:label ?city_name.
        FILTER langMatches(lang(?city_name), "en")
    }
}
ORDER BY DESC (?quantityvalue)
```

LISTING 2: Example of federated query.

time, and theme attributes. Relying on this definition and the SWE standards, some representative features in the Linked-Sensors-Data and sensors Web systems have been selected for the evaluation of SREQP. These features are classified and briefly described below:

(i) Sensors as External Data Sources (SEDS). It is the use of sensors located in different geographical zones as external data sources. From this perspective, a sensor is defined from an engineering point of view as a device that converts a physical, chemical, or biological parameter into an electrical signal [51]. Common examples include sensors to measure temperature (i.e., a thermometer), wind speed (an anemometer) conductivity, or solar radiation (pyranometers). While a sensor is the most basic unit, a sensor system is an aggregation of sensors attached to a single platform [52]. A sensor or a sensor system may be abstracted as a sensor resource. A sensor network consists of a number of spatially distributed and communicating sensor resources [53].

(ii) Spatial Attributes (SAT). Locating specific information for sensors can include very specific geolocation information, such as latitude, longitude, and altitude and/or high-level information that describes the location in high-level terms and relates to other domain concepts (e.g., post codes). In order to provide sensor observation and measurement data in OGC SWE standard Sensor Observation Service (SOS) [49], descriptions are expected to include location attributes explained using GML (Geography Markup Language) (http://www.opengeospatial.org/standards/gml) elements.

(iii) Use of Standards (UST). It consists in the Use of Standards to describe sensors and sensor observations, such as the suite of standards provided by SWE [49, 50]. In context of Linked-Sensor-Data, the Use of Standards consists in the reuse of vocabularies and ontologies (e.g., SSN Climate, Forecast, and WGS84) related to space, time, and theme domains to publish information from sensors in Linked Data format.

(iv) Temporal Attributes (TMPAT). Temporal Attributes in sensor data and their observation and measurement data are those describing attributes such as time zone and measurement timestamp. Using common ontologies for temporal specifications enables linked data consumers to query and access temporal features of data using standard models and interfaces [27].

(v) Thematic Attributes (THMAT). Thematic data provide links between sensor data and the domain knowledge. Attributes such as sensor type, tags, type of observation measurement, features of interest, and other more specific attributes such as operational and deployment attributes describe sensors with domain knowledge [54].

(vi) Sensor-Specific Attributes and Linking to other URIs (SSALU). Sensor data does not only consist of spatial, temporal, and thematic features. As a sensing device it also has more specific attributes and features. In the context of Linked-Sensor-Data, these attributes and features can be related to other resources, since every sensor has a unique URI that refers to its descriptions and can be related to descriptions and attributes of other sensors and resources available on

the Web to browse and access more information. URIs also enable establishing a link between other RDF descriptions of the sensor data and the high-level concepts defined as their property values [27].

(vii) HTTP Access (HTTPA). The Linked-Sensor-Data can be available through HTTP (Hypertext Transfer Protocol) access by simply publishing the sensors descriptions as Web documents. The sensor observation and measurement data can also be available through HTTP interfaces via Sensor Observation Services [27]. The Linked Data paradigm suggests providing SPARQL endpoints to query and access the Linked-Sensor-Data, including the measures, descriptions, and observations, to mention but a few.

(viii) Application Programming Interface (API). The API is necessary to manage sensors and retrieve sensory data. An example of API of sensor Web is the standard SOS (Sensor Observation Services) [49], which is an intermediate layer between data interpretation ("application program" or "client program") and the real-time sensor (or sensory data repository). The sensor metadata can also be retrieved through SOS [55].

Several works in the literature described in the state of the art of this research have relied on these selected features. Among these sensors Web systems LSM Framework [22], FlexFT [23], LO(D)D [24], RSKSensor Data (Wavellite) [25], VISO Data [26], Sense2Web [27], SEIPF [28], and ULDHSWeb [29] are found.

As it can be inferred from the literature, one of the issues of Linked-Sensors-Data and sensor Web systems is the difficulty in evaluating them in terms of semantic publication of data from sensors located in different geographical areas. It may be challenging to evaluate, by means of a quantitative assessment, the legitimacy of the solution provided to extract solar radiation data from external sources and merge it on a single and unique platform. Therefore, a qualitative evaluation was thus favored in this research in order to measure the diverse but also basic aspects of SREQP. Hence, this paper proposes a formal evaluation process to validate the usability of SREQP. The features to evaluate are (a) user satisfaction, (b) simplicity of use, (c) comprehensibility, and (d) perceived usefulness.

5.1. Evaluation Design. A qualitative assessment was adopted to measure the main design aspects of SREQP. This assessment method is different from the quantitative method but aims to provide a style for the semantic publication and consumption of solar radiation data from external data sources. Therefore, the evaluation approach was based on a weighted matrix. Several experts were consulted to propose an efficient evaluation strategy for SREQP, since, due to its complexity, it demands an evaluation strategy different from those performed on other sensor Web systems. The evaluation of SREQP comprised three stages. The first focused on identifying the essential features that must be offered by a platform for the extraction and publishing of solar

radiation data in Linked Data format. The second stage aimed at creating the weighted matrix according to the selected features. Finally, the third step focused on classifying the most important features of a sensor Web system, specifically for Linked Data, and validating whether SREQP provides these features. The third step is also an exploratory search that permits complementing the evaluation of SREQP.

In order to perform the second stage of the evaluation, the Analytic Hierarchy Process (AHP) was carried out [56, 57]. The AHP is defined in Saaty [58] as the theory of relative measurement of intangible criteria. This approach uses paired comparisons, unlike the traditional measurement where some scale is applied to measure any element. Moreover, the elements are measured individually, not by comparison with others.

The AHP enables focusing judgment separately, on each of the several properties necessary to make a sound decision. All elements to be measured, as well as a pair of these elements, must be taken and compared over a single property without worrying about other properties or other elements. The AHP is also useful to evaluate processes [59], transfer and select technology [60], or select product features [61] or critical success factors of executive information systems [62], open source CRM tools [63], and intellectual capital management tools [64], perform IT staff behavior analysis [65], or make IT automation decisions [66]. However, for this research, AHP is used to both classify the most important features of Linked-Sensors-Data and the sensors Web systems and validate whether SREQP provides these features.

In order to perform AHP analysis, the next activities must be covered: (1) select an expert panel; (2) run a pairwise comparison between features; (3) normalize values; and (4) derive conclusions. A 5-point Likert scale is used to run the comparison where 1 stands for "equally important," 2 stands for "slightly important," 3 refers to "more important," 4 means "considerably more important," and 5 stands for "extremely important." One expert in atmospheric sciences (Meteorology, Climatology, and Aeronomy), one expert in electronics engineering (instrumentation and control), and one expert in Semantic Web (Linked Data) met for this evaluation and allowed for the generation of the pairwise comparisons matrix as it is shown in Table 3.

The priority of each item was calculated once the pairwise comparisons matrix was generated. This stage is known as AHP synthesis, which starts by adding values of each matrix column. Then, every item is divided by the total value of its column. The resulting matrix is called normalized pairwise comparisons matrix as Table 4 shows.

The general priority matrix can be then derived, which shows the percentages obtained for every element, as Table 5 shows. Subsequently, the six most important features of a sensors Web system and the Linked-Sensors-Data are obtained, as it is shown in Table 6. According to this table, the most important features that any sensor Web system and Linked-Sensor-Data should have are Sensors as External Data Sources (SEDS), Spatial Attributes (SAT), Sensor-Specific Attributes and Linking to other URIs (SSALU), HTTP Access (HTTPA), Temporal Attributes (TMPAT), and Use of Standards (UST).

Table 3: Essential features for Linked-Sensors-Data and sensors Web systems.

Feature/value	SEDS	SAT	UST	TMPAT	THMAT	SSALU	HTTPA	API
SEDS	**1.00**	4.00	1.50	1.50	1.50	1.00	3.50	2.00
SAT	0.25	**1.00**	2.50	1.00	2.50	2.00	1.00	1.00
UST	0.67	0.40	**1.00**	2.00	1.75	1.00	0.55	1.00
TMPAT	0.67	1.00	0.50	**1.00**	3.75	1.00	0.50	1.00
THMAT	0.67	0.40	0.57	0.27	**1.00**	0.50	1.00	1.00
SSALU	1.00	0.50	1.00	1.00	2.00	**1.00**	2.50	1.00
HTTPA	0.29	1.00	1.82	2.00	1.00	0.40	**1.00**	1.75
API	0.50	1.00	1.00	1.00	1.00	1.00	0.57	**1.00**
Total	5.04	9.30	9.89	9.77	14.50	7.90	10.62	9.75

Table 4: Normalized values.

Feature/value	SEDS	SAT	UST	TMPAT	THMAT	SSALU	HTTPA	API	Total
SEDS	0.20	0.43	0.15	0.15	0.10	0.13	0.33	0.21	1.70
SAT	0.05	0.11	0.25	0.10	0.17	0.25	0.09	0.10	1.13
UST	0.13	0.04	0.10	0.20	0.12	0.13	0.05	0.10	0.88
TMPAT	0.13	0.11	0.05	0.10	0.26	0.13	0.05	0.10	0.93
THMAT	0.13	0.04	0.06	0.03	0.07	0.06	0.09	0.10	0.59
SSALU	0.20	0.05	0.10	0.10	0.14	0.13	0.24	0.10	1.06
HTTPA	0.06	0.11	0.18	0.20	0.07	0.05	0.09	0.18	0.95
API	0.10	0.11	0.10	0.10	0.07	0.13	0.05	0.10	0.76
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	8.00

Table 5: Percentages of features values.

Feature	%
SEDS	21.23
SAT	14.18
UST	11.04
TMPAT	11.60
THMAT	7.37
SSALU	13.23
HTTPA	11.83
API	9.53
Total	**100**

Table 6: Percentages of most important features values.

Feature	%
SEDS	21.23
SAT	14.18
SSALU	13.23
HTTPA	11.83
TMPAT	11.60
UST	11.04

Eight well-known sensor Web systems and Linked-Sensor-Data initiatives were compared with SREQP: LSM Framework, FlexFT, LO(D)D, RSKSensor Data (Wavellite), VISO Data, Sense2Web, SEIPF, and ULDHSWeb. According to the panel of experts in atmospheric sciences (Meteorology, Climatology, and Aeronomy), electronics engineering (instrumentation and control), and Semantic Web (Linked Data), the sensors Web systems and Linked-Sensors-Data initiatives selected do possess the essential features that Sensors Web system and the Linked-Sensors-Data initiatives should have.

For instance, LSM Framework allows for the reuse, interlinkage, and integration of the produced data with other data resources, such as open data (e.g., census, transport, government, crime, and schools to mention but a few), enterprise information systems, social data, and sensor data using the Linked Data principles [17]. Similarly, Sense2Web is a platform to publish Linked-Sensor-Data for the sensor network community. Sense2Web allows users to publish their sensor descriptions data as RDF triples, associate them with any other existing RDF sensor description data, link them to the existing resources on publicly available Linked Data repositories, and make them available to consumers using SPARQL endpoint. On the other hand, ULDHSWeb is a system that also uses Linked Data principles to achieve better data interoperability and integration by republishing real-world data into linked geosensor data. SEIPF is a framework for semantic energy information publishing that provides the ability to query energy consumption information from residential gateways in a machine understandable format in order to achieve consumption coordination and intelligent negotiation. LO(D)D is a framework aimed at smart environments having networks of devices and sensors that interact with one another and with their respective environments to gather, generate, and (if possible) publish data. FlexFT is a generic

TABLE 7: Comparison of SREQP with other Linked-Sensors-Data and sensors Web systems.

	LSM Framework	FlexFT	LO(D)D	RSKSensor Data (Wavellite)	VISO Data	Sense2Web	SEIPF	ULDHSWeb	SREQP
Sensors as External Data Sources	1	1	1	1	1	1	1	1	1
Spatial Attributes	1	1	1	1	1	1	1	1	1
Use of Standards	1	1	1	1	1	1	1	1	1
Temporal Attributes		1	1	1	1	1	1	1	1
Thematic Attributes	1	1	1	1	1	1	1	1	1
Sensor-Specific Attributes and Linking to other URIs	1	0	0	0	1	1	1	1	1
HTTP Access	1	0	1	1	1	1	1	1	1
Application Programming Interface	0	0	0	0	0	0	0	1	0

LSM Framework: Linked Stream Middleware architecture. FlexFT: a generic framework for developing fault tolerant applications in the sensor Web. LO(D)D: Linked Open (Dynamic) Data for smart sensing and measuring environments. RSKSensor Data (Wavellite (http://www2.uef.fi/en/envi/projects/wavellite)): Representing Situational Knowledge acquired from Sensor data for atmospheric phenomena. VISO Data: Virtual Integration of Sensor Observation Data. Sense2Web: Linked-Sensor-Data platform. SEIPF: Semantic Energy Information Publishing Framework. ULDHSWeb: Using Linked Data in a Heterogeneous Sensor Web.

component-based framework that provides a standardized and interoperable interface for sensor observations by relying upon the "Sensor Web" paradigm established by the Open Geospatial Consortium (OGC). RSKSensor Data (Wavellite) is a generic software framework for the organization and interpretation of sensor data applied to a large-scale sensor network for the monitoring of atmospheric phenomena. VISO Data is a framework that enables the virtual integration of heterogeneous observation data sources through a Sensor Observation Service (SOS) standard interface. Currently, VISO Data is being validated by the OGC compliant technology to publish the meteorological and oceanographic observation data generated by two public agencies of the regional government of Galicia, Spain.

The aim of this evaluation is to demonstrate that SREQP possesses the essential features of a Linked-Sensor-Data and sensor systems (see Table 7). Authors wish to note that this evaluation was performed in order to quantify quality of the proposed design. Comparisons of Linked-Sensor-Data and sensor Web systems were merely performed for the purpose of illustrating the benefits of SREQP. Table 7 shows these comparisons performed. Compared initiatives were abbreviated for space optimization and simpler interpretation.

5.2. Results. Table 7 shows the compared elements contained in some of the Linked-Sensor-Data and sensor Web systems described in the state of the art of this research paper. It can be observed that SREQP and the compared initiatives contain most of the selected elements considered essential. However, the Application Programming Interface (API) is merely present in ULDHSWeb, since most of these initiatives enable interlinking, publishing, and integrating data obtained

from sensors Web. Moreover, in some cases, such as in LSM Framework, data are obtained from multiple sources as census, transport, government and schools, enterprise information systems, social data, and sensors data. Although only ULDHSWeb provides a set of RESTful OGC Sensor Observation Services like Linked Data APIs, other initiatives that involve the use of Linked Data (Linked-Sensor-Data, such as LSM Framework, LO(D)D, Sense2Web, SEIPF, and SREQP) usually provide a data vocabulary and access their data through a mechanism of HTTP Access, such as SPARQL endpoint. The SPARQL endpoint enables the processing of SPARQL-based queries to obtain and reuse data stored in Linked Data format. Finally, this data reuse facilitates data navigation, data integration with other information sources, the development of mobile and Web applications, the analysis of solar radiation data, recommendation systems, and semantic systems for predicting weather, to mention but a few.

6. Conclusions and Future Work

This work proposed SREQP that provides a new integration schema to obtain solar radiation data by means of semantic technologies. The architecture of SREQP includes the foundations to provide solar radiation data extracted from several different sources. It allows the final user to easily query information through a SPARQL endpoint. The platform has been designed in such a way that information can be extracted from external data sources in several forms, emphasizing ftp access and Web services as main protocols. The components involved comprise a set of modules that allow for the extraction and conversion of the original information into

RDF triples based on the developed model. The model was designed using other well-known vocabularies, such as SSN Climate, Forecast, and WGS84 for data representation of the several variables that are involved in the model. Furthermore, the model allows for the production and consumption of solar radiation data in Linked Data format, using as use case the solar radiation data obtained from various geographical areas of Spain provided by AEMET.

An expert panel was needed to evaluate the SREQP. This expert panel was composed of specialists in atmospheric sciences, electronics engineering, and Semantic Web. These experts suggested a set of features that Linked-Sensors-Data and sensors Web systems should include. Subsequently, an Analytic Hierarchy Process was carried out before the comparison of SREQP with initiatives that were previously described. The results obtained demonstrate that SREQP initiative possesses the essential features of a Linked-Sensors-Data and sensors Web systems. Results also allowed for the illustration of the benefits SREQP.

Future work will be focused on the extension of SREQP to retrieve and publish, in Linked Data format, information from other types of sensors, such as anemometers, thermometers, barometers, and hydrometers. The approach of SREQP is also expected to extend to the Internet of Things (IoT) in order to generate an important Linked Data source that would contain all possible information concerning electronic devices connected to the Internet. This knowledgebase of data extracted from sensor and electronic devices could be used to analyze this type of data and obtain statistical information, develop semantic recommender systems for energy saving, and develop prediction systems to avoid unnecessary energy expenditure. The approach of SREQP can also be extended to other types of information sources, such as social networks. However, priority will be addressed to sensors and devices that could provide real-time information about weather and environmental measurements as well as energy consumption.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] K. A. Delin and S. P. Jackson, "Sensor web: a new instrument concept," in *Proceedings of the Functional Integration of Opto-Electro-Mechanical Devices and Systems*, vol. 4284 of *Proceedings of SPIE*, pp. 1–9, May 2001.

[2] M. Botts and A. Robin, "Bringing the sensor web together," *Geosciences*, vol. 6, pp. 46–53, 2007.

[3] F. Urbach, "Ultraviolet radiation and skin cancer of humans," *Journal of Photochemistry and Photobiology B: Biology*, vol. 40, no. 1, pp. 3–7, 1997.

[4] T. Heath and C. Bizer, "Linked data: evolving the web into a global data space," *Synthesis Lectures on the Semantic Web: Theory and Technology*, vol. 1, no. 1, pp. 1–136, 2011.

[5] O. Corcho and R. García-Castro, "Five challenges for the semantic sensor web," *Semantic Web-Interoperability, Usability, Applicability*, vol. 1, no. 1-2, Article ID 2019450, pp. 121–125, 2010.

[6] R. Colomo-Palacios, J. L. Sánchez-Cervantes, G. Alor-Hernandez, and A. Rodríguez-González, "Linked data: perspectives for IT professionals," *International Journal of Human Capital and Information Technology Professionals*, vol. 3, no. 3, pp. 1–12, 2012.

[7] T. Stavropoulos, D. Vrakas, D. Vlachava, and N. Bassiliades, "BOnSAI: a smart building ontology for ambient intelligence," in *Proceedings of the 2nd International Conference on Web Intelligence (WIMS '12)*, pp. 263–274, Craiova, Romania, June 2012.

[8] M. J. Santofimia, F. Moya, F. J. Villanueva, D. Villa, and J. C. Lopez, "An agent-based approach towards automatic service composition in ambient intelligence," *Artificial Intelligence Review*, vol. 29, no. 3-4, pp. 265–276, 2008.

[9] N. Shah, K.-M. Chao, T. Zlamaniec, and A. Matei, "Ontology for home energy management domain," *Communications in Computer and Information Science*, vol. 167, no. 2, pp. 337–347, 2011.

[10] A. Ranganathan, R. E. Mcgrath, R. O. Y. H. Campbell, and M. D. Mickunas, "Use of ontologies in a pervasive computing environment," *The Knowledge Engineering Review*, vol. 18, no. 3, pp. 209–220, 2003.

[11] D. Preuveneers and Y. Berbers, "Automated context-driven composition of pervasive services to alleviate non-functional concerns," *International Journal of Computing and Information Sciences*, vol. 3, no. 2, pp. 19–28, 2005.

[12] L. C. Pouchard, M. L. Branstetter, R. B. Cook et al., "A Linked Science investigation: enhancing climate change data discovery with semantic technologies," *Earth Science Informatics*, vol. 6, no. 3, pp. 175–185, 2013.

[13] P. Barnaghi, W. Wang, L. Dong, and C. Wang, "A linked-data model for semantic sensor streams," in *Proceedings of the IEEE and Internet of Things (iThings/CPSCom), IEEE International Conference on and IEEE Cyber, Physical and Social Computing, Green Computing and Communications (GreenCom '13)*, pp. 468–475, Beijing, China, August 2013.

[14] K. Janowicz, A. Bröring, C. Stasch, S. Schade, T. Everding, and A. Llaves, "A RESTful proxy and data model for linked sensor data," *International Journal of Digital Earth*, vol. 6, no. 3, pp. 233–254, 2013.

[15] L. M. Vilches-Blázquez, B. Villazón-Terrazas, O. Corcho, and A. Gómez-Pérez, "Integrating geographical information in the Linked Digital Earth," *International Journal of Digital Earth*, vol. 7, no. 7, pp. 554–575, 2014.

[16] C. Malewski, I. Simonis, A. Terhorst, and A. Bröring, "StarFL—a modularised metadata language for sensor descriptions," *International Journal of Digital Earth*, vol. 7, no. 6, pp. 450–469, 2014.

[17] T. Berners-Lee, *Linked Data—Design Issues. Linked Data*, 2009, http://www.w3.org/DesignIssues/LinkedData.html.

[18] M. Compton, P. Barnaghi, L. Bermudez et al., "The SSN ontology of the W3C semantic sensor network incubator group," *Web*

Semantics: Science, Services and Agents on the World Wide Web, vol. 17, pp. 25–32, 2012.

[19] A. Zaslavsky, C. Perera, and D. Georgakopoulos, "Sensing as a service and big data," in *Proceedings of the International Conference on Advances in Cloud Computing (ACC '13)*, Bangalore, India, 2013.

[20] M. S. Khan and D. Kim, "Enhanced service-oriented open sensor web architecture with application server based mashup," *International Journal of Distributed Sensor Networks*, vol. 2014, Article ID 313981, 10 pages, 2014.

[21] Z. Babovic and V. Milutinovic, "Chapter 2—novel system architectures for semantic-based integration of sensor networks," in *Advances in Computers*, A. Hurson, Ed., vol. 90 of *Connected Computing Environment*, pp. 91–183, Elsevier, 2013.

[22] D. N. Crowley, M. Dabrowski, and J. G. Breslin, "Decision support using linked, social, and sensor data," in *Proceedings of the 8th Americas Conference on Information Systems (AMCIS '13)*, pp. 1–11, Chicago, Ill, USA, 2013.

[23] D. M. Beder, J. Ueyama, J. P. de Albuquerque, and M. L. Chaim, "FlexFT: a generic framework for developing fault-tolerant applications in the sensor web," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 385892, 12 pages, 2013.

[24] F. Corno and F. Razzak, "Publishing LO(D)D: linked open (dynamic) data for smart sensing and measuring environments," *Procedia Computer Science*, vol. 10, pp. 381–388, 2012.

[25] M. Stocker, E. Baranizadeh, H. Portin et al., "Representing situational knowledge acquired from sensor data for atmospheric phenomena," *Environmental Modelling and Software*, vol. 58, pp. 27–47, 2014.

[26] M. A. Regueiro, J. R. R. Viqueira, J. A. Taboada, J. M. Cotos, and M. A. Regueiro, "Virtual integration of sensor observation data," *Computers & Geosciences*, vol. 81, pp. 12–19, 2015.

[27] P. Barnaghi, M. Presser, and K. Moessner, "Publishing linked sensor data," in *CEUR Workshop Proceedings: Proceedings of the 3rd International Workshop on Semantic Sensor Networks (SSN), Organised in Conjunction with the International Semantic Web Conference*, vol. 668, University of Surrey, 2012, http://epubs.surrey.ac.uk/470673/.

[28] D. Bonino, F. Corno, and F. Razzak, "Enabling machine understandable exchange of energy consumption information in intelligent domotic environments," *Energy and Buildings*, vol. 43, no. 6, pp. 1392–1402, 2011.

[29] L. Yu and Y. Liu, "Using Linked Data in a heterogeneous Sensor Web: challenges, experiments and lessons learned," *International Journal of Digital Earth*, vol. 8, no. 1, pp. 17–37, 2013.

[30] G. Atemezing, O. Corcho, D. Garijo et al., "Transforming meteorological data into linked data," *Semantic Web*, vol. 4, no. 3, pp. 285–290, 2013.

[31] Apache-Jena, *Apache jena*, 2013, http://jena.apache.org.

[32] O. Erling and I. Mikhailov, "RDF support in the virtuoso DBMS," in *Networked Knowledge—Networked Media*, Studies in Computational Intelligence, pp. 7–24, Springer, Berlin, Germany, 2009.

[33] C. Bizer and A. Schultz, "The Berlin SPARQL benchmark," *International Journal on Semantic Web and Information Systems*, vol. 5, no. 2, pp. 1–24, 2009.

[34] M. Morsey, J. Lehmann, S. Auer, and A.-C. N. Ngomo, "DBpedia SPARQL benchmark—performance assessment with real queries on real data," in *The Semantic Web—ISWC 2011*, Lecture Notes in Computer Science, pp. 454–469, Springer, Berlin, Germany, 2011.

[35] A. Sheth, C. Henson, and S. S. Sahoo, "Semantic sensor web," *IEEE Internet Computing*, vol. 12, no. 4, pp. 78–83, 2008.

[36] I. Lustig, B. Dietrich, C. Johnson, and C. Dziekan, "The analytics journey," *Analytics Magazine*, pp. 11–18, 2010.

[37] D. Delen and H. Demirkan, "Data, information and analytics as services," *Decision Support Systems*, vol. 55, no. 1, pp. 359–363, 2013.

[38] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*, pp. 851–860, ACM, Ralegh, NC, USA, April 2010.

[39] S. Li, S. Son, and J. Stankovic, "Event detection services using data service middleware in distributed sensor networks," in *Information Processing in Sensor Networks SE-34*, F. Zhao and L. Guibas, Eds., vol. 2634, pp. 502–517, Springer, Berlin, Germany, 2003.

[40] O. Ratzesberger, "Analytics as a Service," 2011, http://www.xlmpp.com/articles/16-articles/39-analytics-as-a-service.

[41] AEMET, *About the State Meteorological Agency (AEMET)*, 2015, http://www.aemet.es/en/quienes_somos.

[42] C. Stasch, T. Foerster, C. Autermann, and E. Pebesma, "Spatiotemporal aggregation of European air quality observations in the Sensor Web," *Computers and Geosciences*, vol. 47, pp. 111–118, 2012.

[43] J. H. Gennari, M. A. Musen, R. W. Fergerson et al., "The evolution of Protégé: an environment for knowledge-based systems development," *International Journal of Human-Computer Studies*, vol. 58, no. 1, pp. 89–123, 2003.

[44] N. F. Noy, R. W. Fergerson, and M. A. Musen, "The knowledge model of Protege-2000: combining interoperability and flexibility," in *Knowledge Engineering and Knowledge Management Methods, Models, and Tools*, vol. 1937 of *Lecture Notes in Computer Science*, pp. 17–32, Springer, Berlin, Germany, 2000.

[45] C. Bizer, S. Auer, G. Kobilarov, J. Lehmann, and R. Cyganiak, "DBpedia—querying wikipedia like a database," in *Proceedings of the Developers Track 16th International Conference on World Wide Web (WWW '07)*, pp. 8–12, May 2007.

[46] J. Volz, C. Bizer, M. Gaedke, and G. Kobilarov, "Silk—a link discovery framework for the web of data," in *Proceedings of the 2nd Workshop on Linked Data on the Web (LDOW '09)*, Madrid, Spain, April 2009.

[47] A.-C. N. Ngomo and S. Auer, "LIMES—a time-efficient approach for large-scale link discovery on the web of data," in *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI '11)*, pp. 2312–2317, AAAI Press, July 2011.

[48] B. Kitchenham, S. Linkman, and D. Law, "DESMET, a methodology for evaluating software engineering methods and tools," *Computing Control Engineering Journal*, vol. 8, no. 3, pp. 120–126, 1997.

[49] M. Botts, G. Percivall, C. Reed, and J. Davidson, "Sensor web enablement: overview and high level architecture," in *GeoSensor Networks SE—10*, S. Nittel, A. Labrinidis, and A. Stefanidis, Eds., vol. 4540, pp. 175–190, Springer, Berlin, Germany, 2008.

[50] A. Bröring, J. Echterhoff, S. Jirka et al., "New generation sensor web enablement," *Sensors*, vol. 11, no. 3, pp. 2652–2699, 2011.

[51] L. Bermudez, E. Delory, O. T. Reilly, and J. del Rio Fernandez, "Ocean observing systems demystified," in *Proceedings of the*

MTS/IEEE Biloxi—Marine Technology for Our Future: Global and Local Challenges (OCEANS '09), pp. 1–7, October 2009.

[52] C. Stasch, K. Janowicz, A. Bröring, I. Reis, and W. Kuhn, "A stimulus-centric algebraic approach to sensors and observations," in *GeoSensor Networks*, N. Trigoni, A. Markham, and S. Nawaz, Eds., vol. 5659 of *Lecture Notes in Computer Science*, pp. 169–179, Springer, Berlin, Germany, 2009.

[53] T. L. van Zyl, I. Simonis, and G. McFerren, "The sensor web: systems of sensor systems," *International Journal of Digital Earth*, vol. 2, no. 1, pp. 16–30, 2009.

[54] W. Wei and P. Barnaghi, "Semantic annotation and reasoning for sensor data," in *Smart Sensing and Context SE-6*, P. Barnaghi, K. Moessner, M. Presser, and S. Meissner, Eds., vol. 5741 of *Lecture Notes in Computer Science*, pp. 66–76, Springer, Berlin, Germany, 2009.

[55] Z. Papp and H. Hakkesteegt, "Sensor web, sensor networks: new possibilities and new challenges," *Sensor Web Enablement*, vol. 45, pp. 21–40, 2008.

[56] T. L. Saaty, "How to make a decision: the analytic hierarchy process," *European Journal of Operational Research*, vol. 48, no. 1, pp. 9–26, 1990.

[57] T. Saaty, "What is the analytic hierarchy process?" in *Mathematical Models for Decision Support SE-5*, G. Mitra, H. Greenberg, F. Lootsma, M. Rijkaert, and H. Zimmermann, Eds., vol. 48, pp. 109–121, Springer, Berlin, Germany, 1988.

[58] T. L. Saaty, "Analytic hierarchy process," in *Encyclopedia of Biostatistics*, John Wiley & Sons, New York, NY, USA, 2005.

[59] L.-A. Vidal, F. Marle, and J.-C. Bocquet, "Using a Delphi process and the Analytic Hierarchy Process (AHP) to evaluate the complexity of projects," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5388–5405, 2011.

[60] S. Lee, W. Kim, Y. M. Kim, and K. J. Oh, "Using AHP to determine intangible priority factors for technology transfer adoption," *Expert Systems with Applications*, vol. 39, no. 7, pp. 6388–6395, 2012.

[61] N. Subramanian and R. Ramanathan, "A review of applications of Analytic Hierarchy Process in operations management," *International Journal of Production Economics*, vol. 138, no. 2, pp. 215–241, 2012.

[62] J. L. Salmeron and I. Herrero, "An AHP-based methodology to rank critical success factors of executive information systems," *Computer Standards and Interfaces*, vol. 28, no. 1, pp. 1–12, 2005.

[63] Y.-C. Lee, N.-H. Tang, and V. Sugumaran, "Open source CRM software selection using the analytic hierarchy process," *Information Systems Management*, vol. 31, no. 1, pp. 2–20, 2014.

[64] D. Carlucci and P. Kujansivu, "Using an AHP rating model to select a suitable approach to Intellectual Capital management: the case of a not-for-profit welfare service," *International Journal of Information Systems in the Service Sector*, vol. 6, no. 3, pp. 22–42, 2014.

[65] V. Stantchev, K. Petruch, and G. Tamm, "Assessing and governing IT-staff behavior by performance-based simulation," *Computers in Human Behavior*, vol. 29, no. 2, pp. 473–485, 2012.

[66] I. Iskin, T. U. Daim, S. Noble, and A. Baltz, "Approaching IT automation decisions using Analytic Hierarchy Process (AHP)," *International Journal of Information Technology Project Management*, vol. 5, no. 1, pp. 77–89, 2014.