4th Imeko TC19 Symposium on Environmental Instrumentation and Measurements
Protecting Environment, Climate Changes and Pollution Control
June 3-4, 2013, Lecce, Italy

MONITORING ELECTROMAGNETIC POLLUTION: A LINKED-DATA-ORIENTED PERSPECTIVE

A. Esposito, L. Tarricone and M. Zappatore

Dept. of Engineering for Innovation, University of Salento, Italy,
{alessandra.esposito;luciano.tarricone;marco.zappatore}@unisalento.it

Abstract: This paper describes how to publish data coming from scattered and heterogeneous Electromagnetic (EM) field sensing systems in compliance with novel knowledge management techniques. Ontological modeling and Linked Data Principles have been exploited to interlink and organize, from a semantic point of view, measurement data with additional information describing sensors, measurement sites and techniques, emitting sources and devices, etc. From the very first achieved results, the proposed approach exhibits many appealing features. For instance, the possibility to seamlessly access different legacy data sources or to create semantic relationships across related scientific areas. Therefore, traditional data publication techniques for EM pollution monitoring campaigns can now be thought-back and broadened by adding new features to reach more variegated categories of users rather than only technicians or managing authorities.

Keywords: Electromagnetic Monitoring, Linked Data, Ontologies, R2RML, Knowledge Management.

1. INTRODUCTION

Nowadays, it is extremely important to have immediate availability of heterogeneous data in scientific research contexts, as well as in everyday life. Researchers, local administrators and even ordinary citizens may indeed achieve tangible benefits in their activities if provided with structured and semantically well-organized knowledge. In the same manner, it should be advisable to have novel, quicker and more performing modalities to gather and search for information in comparison with traditional techniques. According to such premises, recent advances in the field of the so-called *"Web of Data"* promise to reach that knowledge pervasivity which exists, at the moment, only in theory.

As a consequence of that, our research activity has a twofold purpose. On the one hand, we aim to enhance the way into which data coming from Electromagnetic (EM) field monitoring campaigns are traditionally published. In order to do that, the most up-to-date knowledge management techniques (ontological modeling and Linked Data) have been exploited.

On the other hand, we aspire to make people capable of *"sharing and consuming structured data on the Web as easily they can share documents today"*, according to Heath and Bizer's Linked Data foundational idea [1].

Therefore, the proposed approach tries to make any kind of users well aware of data coming from heterogeneous EM field sensing devices scattered across a given geographical area, by providing additional information exhibiting a rigorous semantic structure and coming from diverse external sources.

2. EM POLLUTION MONITORING: A BRIEF OVERVIEW

Currently, many local authorities and public healthcare agencies perform, even sporadically, monitoring campaigns of EM fields in many geographical areas. Continuous monitoring stations are usually placed for certain periods of time in the proximity of high frequency (e.g., cellular base transceiver stations, Wi-Fi access points, radio-tv repeaters, etc.) and low frequency (e.g., aerial power lines, electrical transformer cabins, power plants, etc.) emitting sources. Monitoring activities are also performed around sensitive sites such as schools, hospitals and playgrounds as well as near densely populated areas. In addition, many other ad-hoc, sporadic measurement campaigns are often accomplished by technical operators on behalf of telecommunications managing authorities in order to locate new installation sites, to control in-situ emitting devices, to tune system emissions.

A typical EM field sensing scenario involves a set of monitoring platforms (each mounting one or more different field probes) located in specific sites in a given geographic area. Monitoring stations continuously detect electric or magnetic fields depending on the technical specifications of the sensors they have on board. Radio-electric and location parameters of monitored emitting devices determine the selection of proper field probes. Operational frequency range, deployment type (i.e., indoor or outdoor), presence of surrounding metal structure, co-location of multiple radiating sources are some of those factors.

Collected measurements are then sent to some data repositories, usually remote relational databases (RDBs), in order to store them for long period of times. These storage solutions usually have to be maintained unchanged.

This briefly sketched scenario includes a great variety of data sources, which differ with respect to many aspects such as time, geographical provenance, employed measuring devices and techniques, targeted radiating sources, normative references. Moreover, data storage locations usually adopt proprietary schemas that make data exchange and integration very difficult: for instance, the outcomes of the same electric field measurement procedure can be stored in many different ways depending on how the storing RDBs have been modeled.

Publishing EM sensor data can be troublesome as well, since it is mainly affected by two limitations.

The first one refers to a relative scarce availability of measurement data: they can be usually consulted under restricted access policies (e.g., by operators from public healthcare organizations or telecommunication companies), or limited to specific geographical areas or no longer accessible at all. This is the case, for instance, of the data referring to the on-going national EM measurement campaign [2].

The second drawback regards the way measurement data are presented to their potential users, which is quite often a static one. Allowed users are simply provided with measurement values and few additional knowledge about locations and sensors.

In this paper, it will be shown that by exploiting Linked Data main features and ontological modeling techniques, other kinds of information can be presented to final users.

3. LINKED DATA AND ONTOLOGIES

The foundational idea of Linked Data is to share structured data on global scale, imagining the Web as a huge, interlinked, distributed database where each piece of knowledge can be stored and cost-efficiently retrieved or transferred in a standardized manner. By doing so, not only the information naturally produced within each information system but also external data would be available to end-users, overcoming traditional data search limitations and architectural gaps. In other words, publishing information as linked data simply consists in a series of mapping techniques applicable to already existing data without modifying the way they are stored in their original repositories.

A set of practices facilitating the correct publication of structured data on the Web is encompassed under the definition of Linked Data. Such recipes were firstly formalized by Tim Berners-Lee [3] in 2006 and can be summarized as follows:

1. to name each knowledge entity with URIs;
2. to use HTTP URIs that allow entities look up;
3. to adopt well-known standards such as RDF [4] to convey information about entities and SPARQL [5] to query them;
4. to include links to other URIs, enabling information discovery;
5. to be open source (optionally).

Figure 1 shows an RDF triple, the typical linked data formal representation. By employing RDF triples, each piece of knowledge is codified by a subject, a predicate and an object, thus potentially rendering any kind of concept expressed by a natural language sentence. Each triple element is reachable on the Web thanks to its HTTP URI (*dereferencing* process). HTTP URIs are shortened by using a *namespace prefix* associated to an *entity identifier*.

Data published according to the principles mentioned above can be included into the so-called linked data cloud [6], which already gathers some billions of entities coming from many different knowledge domains. Publishing knowledge entities and linking them to other linked data cloud elements allow users to follow a potentially infinite number of links across multiple information sources.

In the same way linked data provide a standardized method to publish structured data, ontologies offer a rigorous formalization enabling knowledge codification and promoting information sharing and reuse across different research organizations [7]. From astronomy [8] to healthcare [9], ontologies find a great variety of applications.

Figure 2 depicts our proposed vision applied to the EM field monitoring context: in a traditional scenario (Fig.2, upper part), sectorial users access sectorial data and then manually link and exchange them according to their needs. In the linked data approach (Fig.2, lower part), on the contrary, every kind of data is available at the same time and therefore presented to all categories of users. Starting from a set of distributed data sources collecting heterogeneous data, ontologies offer a unifying knowledge modelling, whilst information is published as linked data, thus enabling the interlinking between diverse data sources.

A wide range of benefits can be achieved by adopting the methodology just described. For instance, better integration between public healthcare monitoring entities and population as well as between wireless communication systems providers and public authorities; deeper and wider analysis of EM emissions; more efficient survey of monitoring platforms and sites; integration of data concerning sensor observations, location of emitting devices, sensitive targets, sensor descriptions and measurement thresholds; data availability on a larger scale; fast and reliable on-line retrieval of measurement data.

4. CASE STUDY DESCRIPTION AND SETUP

A typical EM monitoring scenario contains several monitoring station, each providing a measurement every *n* minutes, according to adopted regulatory norms (e.g., [10] and [11]). In order to simulate such scenario, we developed a set of software modules that produce timed measurements

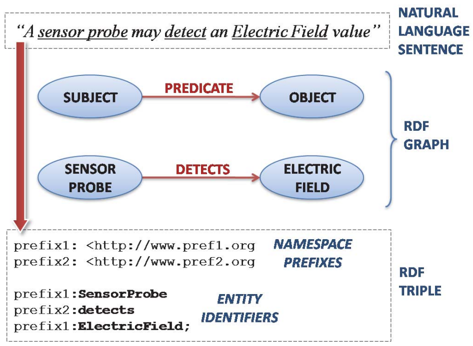

Fig.1 Natural language sentence to RDF triple mapping.

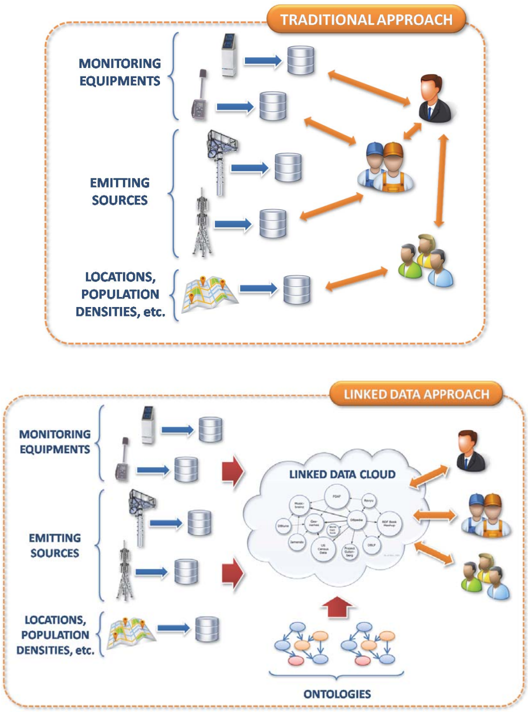

Fig.2 EM monitoring: traditional vs linked data approach

according to log-normal distributions, thus replicating the real behavior of sensor probes whose outcomes are generally below danger thresholds for the majority of observation time. This allows us to have a considerable number of data sources. Figure 3 depicts a set of 50 sample measurements from a simulated NARDA PMM EP-330 electric field probe, whose operational frequency range is 100 kHz – 3 GHz.

Then, we modeled and deployed a medium-sized (around 20 tables) relational DB in order to store both dynamic data (i.e., simulated measurements) and static data (e.g., information about monitoring sites, monitoring stations and platforms, field probes, operational periods, etc.). This RDB mimics a typical *legacy data source* that remains unchanged, as quite often happens in reality. Finally, both kinds of data have been properly mapped and then published as linked data.

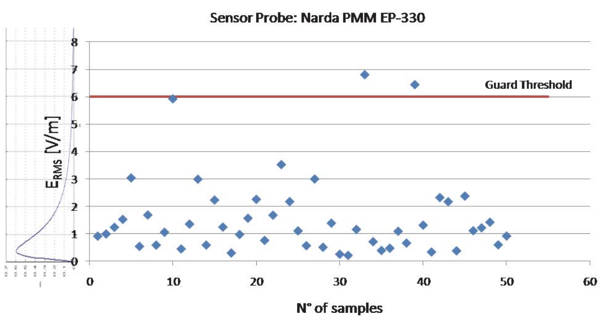

Fig.3 Simulated electric field probe. Observed quantity: *rms* value of electric field [V/m].

Before linked data publication may take place, a considerable number of preparatory steps has to be performed. Heath and Bizer [1] provide a useful publication checklist which we followed in our research work.

1. Identification of existing data sources.
2. Selection of linked data pattern to rigorously model knowledge sources.
3. Selection of URI syntax for proprietary vocabulary in order to make entities de-referenceable.
4. Selection and adoption of external, widely deployed sources: RDF vocabularies and OWL ontologies are needed to describe and annotate each entity.
5. Creation of outgoing links to other datasets to widen knowledge.
6. Metadata specification (provenance, licensing).
7. Access method definition, in order to specify how users can retrieve linked data, once published (e.g., SPARQL endpoints, RDF dataset dumps).

5. HTTP URIS AND LINKED DATA PATTERNS

In order to be profitably published as linked data, source data have to be structured in a precise and rigorous way. We adopted the well-known design pattern technique proposed by Davis and Dodds [12] to model them. Indeed, both in ontological engineering and in linked data publishing, design patterns are gaining a constantly growing interest since they offer well-defined, shared and agreeable modeling structures. The adopted pattern refers to quite each step of the linked data realization chain: 1) identification, 2) modeling and 3) publication of resources.

HTTP URI type. The most important feature linked data must comply with is represented by the adoption of de-referenceable HTTP URIs (i.e., any HTTP client looking up an URI must be capable of retrieving the description of its corresponding resource). Each linked data provider must select the most appropriate URI definition strategy for its purposes. Such strategies, called *303 URIs* and *hash URIs* [13], make URIs de-referenceable for software applications and quite understandable by humans as well. Without going into details, the former are more indicated to describe relatively small RDF datasets (up to a thousand triples), whilst the latter are usually adopted to describe resources in larger RDF datasets. Since it is plausible to have a huge number of observations coming daily from a relevant number of sensors distributed across a geographical area in a real-world scenario, we adopted *303 URIs*. As a consequence, a typical sensor observation is expressed as in the following: *[namespace]/[Observation identifier]*

HTTP URI type identifier pattern. As multiple measurements are sent from different sensors at the same time, we adopted hierarchical HTTP URIs identifier patterns [12] conforming to the following:

[namespace]/[platform_site]/[sensor]/[timestamp].

Each measurement is therefore identified by the hierarchy *"hosting platform, detecting sensor, timestamp"*. More in detail, the measurement sensed by sensor no.2

belonging to monitoring platform no.4 on April 29th, 2013 at 11:30:15 a.m. will be identified as:

[namespace]/site4/sensor2/20130429T113015.

where the timestamp is expressed in a fixed XSD-compliant format. The main advantage is that hierarchical URIs are more understandable by human users and have a substantially constant length.

Modelling pattern. In linked data modeling, it is mandatory to reuse available resources, in order to promote knowledge sharing and reuse rather than creating it from scratch. Therefore, existing vocabularies and/or ontologies must be searched for and referenced. Secondly, it is not necessary to completely model an entire knowledge area: a model can be iteratively populated, adding new contents in each step. Moreover, linked data do not need to be aligned with top-level entities: RDF triples can be focused around just few relevant notions, without considering more general entities they may derive from (since it is more important to create outgoing links towards external RDF datasets).

We adopted the so-called ***Sensor-Stimulus-Observation*** design pattern [14] to describe sensor data since it has been chosen by the W3C Incubator Group to model the Semantic Sensor Network (SSN-XG) Ontology [15] in compliance with two important aspects: a) minimal ontological commitment, b) absence of top-level ontologies. For these reasons, it seems very well-suited for linked data publishing. SSN-XG ontology has been adopted as well (see Section 6). *Stimulii* represent detectable changes in the physical world, acting as triggers for sensors. *Sensors* are seen as any physical object capable of performing observations about incoming stimulii. Each *Observation* consists in a transformation from a stimulus into a representation and should be described with additional parameters such as (at least) time and location.

6. ONTOLOGY AND VOCABULARY SELECTION

Both proprietary and external resources have been used. Proprietary resources encompass a modular ontology proposed by the authors, OntoCEM (Ontological Codification of ElectroMagnetism) [16], [17] as a semantic description and integration of different branches of Electromagnetism. This ontology has been used to describe important concepts such as antennas and monitoring stations as well as to model semantic relationships existing across different EM-related entities.

External resources, enlisted afterwards, have been chosen according to specific selection criteria:

- availability: to be open source and freely accessible;
- status: to be stable and as complete as possible;
- complexity: to have a minimal ontological commitment (few entities and few restrictions);
- consistency and validation: to be without semantic anomalies and to have been formally validated;
- domain independency: to have a small number of imported resources;
- modeling technique: to be property-based rather than class-based (since properties make easier the realization of outgoing semantic links between proprietary entities and external resources).

Geo-localization: places and entities are localized in terms of latitude, longitude and altitude (w.r.t. sea level) from the WGS84 reference datum specification thanks to the *Basic Geo WGS84 lat/long vocabulary* [18]. This RDF vocabulary has been selected due to its minimal size and since it is widely used for defining entity locations also in DBPedia and GeoNames.

Physical Places: places have been aliased by linking them to *DBPedia* [19] and *GeoNames* [20] corresponding elements, if present. Both of them are huge sources of structured information containing more than 1M RDF triples, therefore they act as *"a hub for links in the Web of linked data from other sources"* [21].

Temporal description of events: according to [22], due to a relative broad availability of time-related ontologies and vocabularies, a thorough analysis of available candidates has been performed with respect to selection criteria. The *Linking Open Descriptions of Events (LODE)* ontology [23] satisfied the majority of criteria, then has been selected. *OWLTime ontology* [24] and *Timeline ontology* [25] have been used to fill some semantic gaps in LODE. OWLTime provides properties to define XSD-compliant timestamp formats whilst Timeline allows to describe temporal durations and time lapses.

Numerical quantities and unit of measurements: in order to define actual numerical values for each observation and corresponding units of measurement, we employed the *NASA Quantity Unit Dimension (QUDT)* ontology**,** v1.1 [26]. Also in this case, other ontologies were available as well, such as *DOLCE* [27] and *Ontology of Units of Measure (OM)* [28] but they have been discarded. The former exhibits an insufficient level of detail to express units of measurement and the latter is still in development. Therefore, the potential lack of stability in OM may determine changes in ontological structure that affect the mapping with our linked data structure. In addition, whilst OM adopts a more intuitive way to define SI prefixes and presents a greater number of (sub)multiples units, QUDT covers many more scientific areas, making not necessary to search for additional units in specific use cases.

Monitoring sensors and platforms: the *Semantic Sensor Network XG (SSN-XG)* ontology [15] has been used. We adopted this ontology since it has been implemented after an in-depth analysis of 17 available ontologies dealing with sensors (i.e., sensor-centric) and sensor observations (i.e., observation-centric), by taking into considerations different modeling techniques and merging the most interesting features from each of them. As a consequence, SSN represents an up-to-date, state-of-the-art ontology concerning sensor networks.

Top-level ontology: two ontologies included in our conceptual model (i.e, LODE and SSN-XG) adopt the *Dolce Ultra Lite (DUL)* ontology [29] as upper level semantic layer. Therefore, despite it is not strictly requested to have higher level concepts directly related to linked data, we decided to use classes and properties from DUL.

7. MAPPING OF LEGACY DATA SOURCES

When structured data exist in the form of legacy RDBs, specific RDB-to-RDF wrappers are needed. We adopt the R2RML [30] mapping language since it allows customized mappings from RDBs to RDF datasets. The main advantage of this approach, if compared with direct mappings which simply translate schema structure in RDF datasets, is that R2RML allows the author to choose both personalized conceptual mapping structure and target vocabularies and finely tune them.

This extremely powerful technique starts maps each row of logical RDB tables into a set of RDF triples by using specific mapping rules called *triples maps*.

In this way, any existing RDB can be profitably mapped into an RDF dataset by using an R2RML processor fed by a user-defined set of triples maps. It is important to stress that relational data remain unchanged: no data is extracted nor duplicated, only logical RDF mapped datasets become accessible. Then, in order to expose and query mapping results, specific SPARQL endpoints (Section 4) must be attached to RDF datasets.

We adopted a free R2RML processor developed by Revelytix, namely Spyder [31], that also allows the creation of SPARQL 1.1 compliant endpoints. As a consequence, we developed a fully mapping which covers our legacy RDB. The following excerpt shows how a monitoring site can be rendered. Firstly, each row in `msite` table becomes an instance of the class `MonitoringSite` defined in OntoCEM ontology. Then, each column record becomes an annotation: for instance, `lat` values are rendered as object values for the homonym predicate taken from WGS84 ontology. Namespace prefixes are not reported for brevity. Figure 4 represents such mapping (in the same figure it is also presented an outgoing link to an equivalent entity in DBPedia).

```
@prefix rr: <R2RML>.
@prefix monit: <Monitoring module in OntoCEM>.
@prefix wgs84: <WGS84 ontology>.
@prefix em-lod: <EM-LOD dataset>.

<#TriplesMap_MonitoringSite>
rr:logicalTable [ rr:tableName "msite" ];

rr:subjectMap [rr:template "http://www.electroma
gnetics.unisalento.it/lod/em/1.0/msite-{id}";
rr:class monit:MonitoringSite ; ];
#latitude
rr:predicateObjectMap [ rr:predicate wgs84:lat ;
    rr:objectMap [ rr:column "lat" ] ; ].
```

8. QUERYING RDF DATASET

The main advantages of the proposed approach are clearly visible when the generated RDF dataset is queried through the SPARQL endpoint.

In Fig.4 a first RDF graph was presented, Fig.5 now illustrates a graph describing the complete semantic description and the resulting RDF triples of a single sensor measurement. As it can be seen, many ontologies are involved: OntoCEM for the description of specific EM-related entities; SSN-XG for measuring devices and their capabilities; LODE, OWLTime and Timeline for measurement timestamp and duration; QUDT for numerical values and unit of measurements. In the same way, any other available piece of knowledge has been semantically codified.

As for querying capabilities, we present an example of complex SPARQL query that overcomes limitations of traditional monitoring data presentation. Let's suppose a user wants to know the first 10 emitting sources which are located in a *max_dist=10* km range around a given measurement site, knowing its latitude *pLAT*, longitude *pLONG* and its angular geo-coordinates *phi* and *lambda*. In normal conditions, this requires to access multiple RDBs, manually compare data, search for antenna locations and so on. In a fully semantic and linked data scenario, on the contrary, it is sufficient to trigger a filtered SPARQL query specifying the range of interest to a SPARQL endpoint, provided that both measurement sites and antenna locations have been triplified. An example of such query is reported below.

```
PREFIX wgs84: <WGS84 ontology>
PREFIX xsd: <XMLSchema>
PREFIX ant: <Antenna module in OntoCEM>

SELECT  ?uri  (((pLAT-xsd:float(?lat))*(pLAT-
xsd:float(?lat)) + (pLONG-xsd:float(?lon))*(pLONG-
xsd:float(?lon))  *(phi-(lambda*xsd:float(?lat))))
AS ?dist_grad)
WHERE {    GRAPH  <EM-LOD> {
      ?uri geo:location ?loc; a  ant:Antenna.
      ?loc geo:long ?lon; geo:lat ?lat.
      FILTER(     (pLAT-xsd:float(?lat))*(pLAT-
xsd:float(?lat)) + (pLONG-xsd:float(?lon))*(pLONG-
xsd:float(?lon)) *(phi-(lambda*xsd:float(?lat))) <
max_dist ) } } LIMIT 10
```

9. CONCLUSION AND FURTHER DEVELOPMENTS

In this paper, a novel semantic and linked-data-based approach to publish EM field monitoring data has been presented. By exploiting ontological modeling and linked data best practices, it has been demonstrated how this scientific area can benefit from the adoption of such knowledge management techniques. Indeed, the more monitoring and monitored devices are mapped as RDF triples and interlinked with other RDF datasets, the wider becomes the knowledge accessible by end-users. Following semantic links across RDF datasets allows data consumers (both human and software agents) to discover new information and to find quick answers to their needs.

Many other research aspects are at the moment under examination. For instance, the integration with third-party RDF datasets describing placements and technical specifications of cellular base stations. Indeed, they can be easily mapped with dedicated modules from OntoCEM ontology [16], thus broadening reachable knowledge. Specific free software tools for semi-automatic discovery of outgoing links are under evaluation as well. Finally, more performing R2RML processors with SPARQL endpoints enabling the submission of SPARQL federated queries are currently under test.

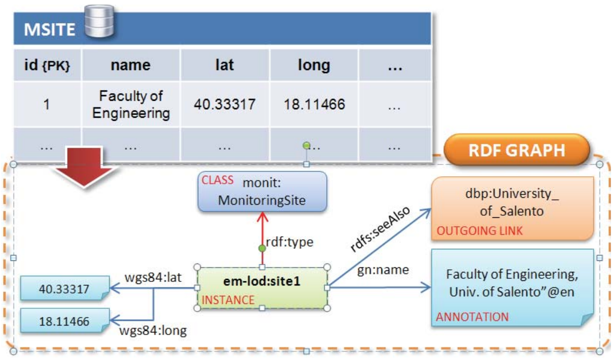

Fig.4 R2RML partial mapping of `msite` table.

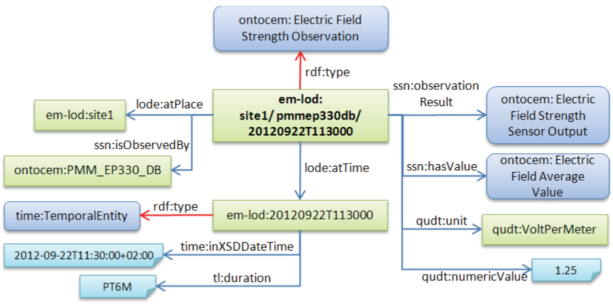

Fig.5 Electric field measurement: semantic mapping.

10. REFERENCES

[1] T. Heath and C. Bizer, Linked Data: Evolving the Web into a Global Data Space (1st Ed.). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, 2011

[2] Fondazione Ugo Bordoni, "Monitoraggio dei Campi Elettromagnetici". [Online] http://www.monitoraggio.fub.it (last accessed: April 2013).

[3] T. Berners-Lee, "Linked Data – Design Issues", 2006. [Online] http://www.w3.org/DesignIssues/LinkedData.html (last accessed: April 2013).

[4] G. Klyne and J.J. Carroll, "Resource Description Framework (RDF): Concepts and Abstract Syntax – W3C Recommendation", 2004.

[5] E. Prud'hommeaux and A. Seaborne, "SPARQL Query Language for RDF – W3C Recommendation", 2008.

[6] Linked Open Data Cloud, 2011 [Online] http://www.lod-cloud.net (last accessed: April 2013).

[7] G.T. Gruber, "Towards Principles for the Design of Ontologies Used for Knowledge Sharing", Int. Journal of Human-Computer Studies, vol. 43, issue 5-6, pp. 907-928, 1995.

[8] S. Lesteven et al., "Ontologies for Astronomy", Library and Information Services in Astronomy V: Common Challenges, Uncommon Solutions. ASP Conf. Series, vol.377, p.193, 2006.

[9] OpenClinical Ontologies, 2011 [Online] http://www.openclinical.org/ontologies.html (last accessed: April 2013).

[10] Comitato Elettrotecnico Italiano (CEI), Norma CEI 211-6, "Guida per la misura e la valutazione dei campi elettrici e magnetici nell'intervallo di frequenza 0 Hz – 10 kHz, con riferimento all'esposizione umana", CT 106, F. 5908, 2001.

[11] Comitato Elettrotecnico Italiano (CEI), Norma CEI 211-7, "Guida per la misura e la valutazione dei campi elettrici e magnetici nell'intervallo di frequenza 10 kHz – 300 GHz, con riferimento all'esposizione umana", CT 106, F. 5909, 2001.

[12] I. Davis and L. Dodds, Linked Data Patterns – A Pattern Catalogue for Modelling, Publishing and Consuming Linked Data, 2012. [Online] http://patterns.dataincubator.org (last accessed: April 2013).

[13] R. Cyganiak, R. Delbru and G. Tummarello, "Semantic Web Crawling: a Sitemap Extension", 2007. [Online] http://sw.deri.org/2007/07/sitemapextension (last accessed: April 2013).

[14] K. Janowicz and M. Compton, "The Stimulus-Sensor-Observation Ontology Design Pattern and its Integration into the Semantic Sensor Network Ontology", 3rd Int. Workshop on Semantic Sensor Networks (SSN10), 2010.

[15] L. Lefort, C. Henson and K. Taylor, "Semantic Sensor Network XG Final Report", 2011..

[16] A. Esposito et al., "A Proposal for an Electromagnetic Ontology Framework", Int. Journal of Web and Grid Services, Vol.4, No.3, pp.284-302, 2008.

[17] Ontological Codification of ElectroMagnetism v.2.2 (OntoCEM), University of Salento, Electromagnetic Lab Lecce (EML[2]). [Online] http://www.electromagnetics.unisalento.it/ontologies/OntoCEM/2.2/ (progressively available).

[18] D. Brickley, "Basic Geo (WGS84 lat/log) Vocabulary", W3C Semantic Web Interest Group Report, 2006..

[19] DBPedia.org, 2012. [Online] http://dbpedia.org/about (last accessed: April 2012).

[20] GeoNames, the Geographical Database, 2012. [Online] http://www.geonames.org (last accessed: April 2013).

[21] S. Auer et al., "Dbpedia: a Nucleus for a Web of Open Data", 6th Int. Semantic Web Conference, pp. 11-15. Busan, Korea, 2007.

[22] W.R. Van Hage, et al., "Design and Use of the Simple Event Model (SEM)", Web Semantics: Science, Services and Agents on the World Wide Web, 9 (2011); 128-136, 2011.

[23] R. Troncy, and L. Hardman (Eds.), "LODE: An Ontology for Linking Open Descriptions of Events", 2011..

[24] J.R. Hobbs and F. Pan (Eds.), "Time Ontology in OWL", W3C Working Draft, 27 Sept. 2006.

[25] Y. Raimond and S. Abdallah (Eds.), "The Timeline Ontology", 2007.

[26] R. Hodgson and P.J. Keller, "QUDT – Quantities, Units Dimensions and Data Types in OWL and XML", 2011.

[27] A. Gangemi et al., , "Sweetening WordNet with DOLCE", AI Magazine, 24(3): Fall 2003, pp. 13-24, 2003.

[28] H. Rijgersberg et al., "Ontology of Units of Measure and Related Concepts", Semantic Web Journal, 2012.

[29] A. Gangemi, The DOLCE+DnS Ultralite Ontology. [Online] http://ontologydesignpatterns.org/wiki/Ontology:DOLCE+DnS_Ultralite (last accessed: April 2013).

[30] S. Das, S. Sundara and R. Cyganiak (Eds.), "R2RML: RDB to RDF Mapping Language", W3C Recommendation, 27 Sept. 2012.

[31] Revelytix Spyder homepage, 2012. [Online] http://www.revelytix.com/content/spyder (last accessed: April 2013).