# Thirty years of digital libraries research at the University of Padua: The systems side

Maristella Agosti, Giorgio Maria Di Nunzio, Nicola Ferro and Gianmaria Silvello

University of Padua, Italy
{name.surname}@unipd.it

**Abstract.** For the thirty years of the Information Management Systems (IMS) research group of the University of Padua, we report the main and more recent contributions of the group to the field of Digital Library Systems. In particular, we briefly describe the systems designed and developed by members of the group in the context of research infrastructures, digital archives, digital linguistics and scientific data.

## 1 Introduction

Digital libraries have contributed to supporting the creation of innovative applications and services to access, share and search our cultural heritage. One of the most important contributions of digital libraries is to make available collections of digital resources from different cultural institutions such as *libraries*, *archives* and *museums*, to make them accessible in different languages and to provide advanced services over them. Digital libraries are heterogeneous systems with functionalities that range from data representation to data exchange and data management. Furthermore, digital libraries are meaningful parts of a global information network which includes scientific repositories, curated databases and commercial providers. All these aspects need to be taken into account and balanced to support final users with effective and interoperable information systems.

In the last thirty years the Information Management Systems (IMS) research group of the University of Padua has contributed to the design and development of diverse digital library systems contributing to the foundations of the field by providing an interoperability layer between the DELOS model and the 5S model (Section 2), to research infrastructures with the CULTURA environment (Section 3), to digital archives with the SIAR system (Section 4), to digital linguistics with the ASiT project (Section 5) and to the access and re-use of scientific data with LoD DIRECT (Section 6).

## 2 Foundations: The DELOS Model and the 5S: Interoperability

The evolution of *Digital Library (DL)* has been favour ed by the development of two foundational models of what DL are, namely the *Streams, Structures, Spaces, Scenarios, Societies (5S)* model [20] and the *DELOS Reference*

*Model* [15], which made it clear what kind of entities should be involved in a DL, what their functionalities should be and how *Digital Library System (DLS)* components should behave, and fostered the design and development of operational DLS complying with them.

However, these two models are quite abstract and, while still providing a unifying vision of what a DL is, they allow for very different choices when it comes to developing actual DLS. This has led to the growth of "ecosystems" where services and components may be able, at best, to interoperate together within the boundaries of DLS that have been inspired by just one of the two models for DL.

In [9] we addressed the need for interoperability among DLS at a high level of abstraction and we showed how this is achieved by a semantically-enabled representation of foundational DL models. The ultimate goal has been to promote and facilitate a better convergence and integration in the context of libraries, archives and museums by lowering the barriers between them.

We proposed a common ontology which encompasses all the concepts considered by the two foundational models and creates explicit connections between their constituent domains. In particular, the user, functionality and content domains allow us to enable a high-level interoperability between the actors and the information/digital objects of DL as well as their functions/services.

The DELOS Reference Model and the 5S Model are defined starting from two different viewpoints. Indeed, in DELOS the approach is top-down since it defines the entities and relationships involved in a DL; whereas the 5S model is largely bottom-up starting with key definitions and elucidation of digital library concepts from a minimalist approach. For this reason, some of the concepts modelled by the DELOS Reference Model are not explicitly modelled by the 5S model. The common ontology we defined is particularly effective since it enriches the 5S model with the concepts defined by the DELOS Reference Model, creating further bridges between them and their implementations.

In Figure 1 we present the *Resource Description Framework (RDF)* graph of the unifying data model relating the DELOS Reference Model to the 5S model by means of a mapping between their most relevant high-level concepts. The presented RDF graph is a visual overview of the ontology we developed.

## 3 Digital Library Research Infrastructures: The CULTURA Environment

The CULTURA environment is service oriented and is composed of a set of services which integrate to create a rich and engaging experience that supports users of different categories which range from academic and professional users to the general public. The services are conceived and developed to be applicable to a wide variety of cultural collections. The potential generality of the environment is demonstrated by the fact that CULTURA is supporting different use cases that are represented by the *Imaginum Patavinae Scientiae Archivum (IPSA)* and 1641 collections, which differ in morphology, language, modality and metadata. This

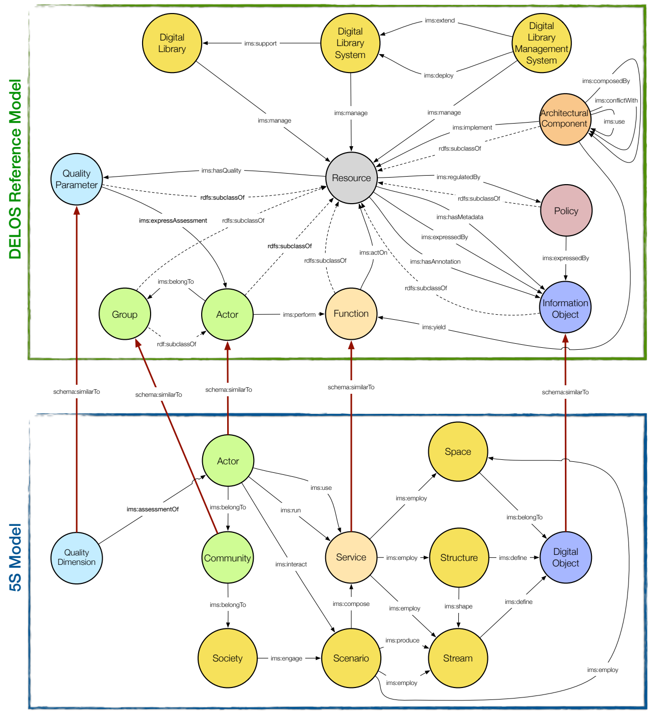

**Fig. 1.** Semantic mapping of the high level concepts in the 5S model and DELOS Reference Model and their relationships [9].

means that the environment and the supported services need to consider the peculiarities of different documents and different ways of making use of them by diverse categories of users. One of the supported services which must be conceived and made available, taking into specific account the peculiarities of the documents of different collections, is the annotation service [3].

Almost everybody is familiar with annotations and has his own intuitive idea about what they are, drawn from personal experience and the habit of dealing with some kind of annotation in everyday life, which ranges from jottings for the shopping to taking notes during a lecture or even adding a commentary to a text. This intuitiveness makes annotations especially appealing for both researchers and users: the former propose annotations as an easily understandable way of performing user tasks, while the latter feel annotations to be a familiar tool for carrying out their own tasks. Therefore, annotations have been adopted in a variety of different contexts, such as content enrichment, data curation, collaborative and learning applications, and social networks, as well as in various information management systems, such as the Web (semantic and not), digital libraries, and databases.

The role of annotations in digital humanities is well known and documented [2, 5, 7]. Subsequently, many different tools which allow for the annotation of digital humanities content have been developed. Unfortunately, tools designed specifically for an individual portal are typically only compatible with that system. More general solutions, which can be easily distributed across various sites, have been developed, but these systems often have limited functionality (only annotating a single content type, no sharing features etc.). FAST-CAT (Flexible Annotation Semantic Tool - Content Annotation Tool) is a generic annotation system that directly addresses this challenge by providing a convenient and powerful means of annotating digital content. Figure 2 shows an example of an annotation supported by the CULTURA environment. According to this model, an annotation is a compound multimedia object which is constituted by different signs of annotation. Each sign materializes part of the annotation itself; for example, we can have textual signs, which contain the textual content of the annotation, image signs, if the annotation is made up of images, and so on. In turn, each sign is characterized by one or more meanings of annotation, which specify the semantics of the sign; for example, we can have a sign whose meaning corresponds to the title field in the Dublin Core (DC) metadata schema, in the case of a metadata annotation, or we can have a sign carrying a question of the authors about a document whose meaning may be "question" or similar. An annotation has a scope which defines its visibility (public, shared, or private), and can be shared with different groups of users. Public annotations can be read by everyone and modified only by their owner; shared annotations can be modified by their owner and accessed by the specified list of groups with the given access permissions, e.g., read only or read/write; private annotations can be read and modified only by their owner.

## 4 Digital Archives: SIAR

The main characteristics of archives are their structure and the objects they manage and preserve. An archive is a complex organization composed by several parts. The foremost component regards the descriptive part of an archive which is conceptually modelled by the *International Standard for Archival Description (General) (ISAD(G))* standard defining the hierarchical organization of archival descriptions and how to model the relationships between them.

We point out two main aspects that we have to consider when modelling an archive: *hierarchy* and *context*. The first aspect means that we have to be able to represent and maintain the hierarchical structure of an archive and its descriptions; the second aspect means that we have to retain the relationships between the archival descriptions and to exploit them to reconstruct the context of a document in relationship with its creation and preservation environment. In order to express hierarchy and context we need a model which allows us to represent the structure of an archive. Furthermore, we also need to represent the content of an archive which is described and managed by means of archival descriptions – that in a digital environment are represented by archival metadata.

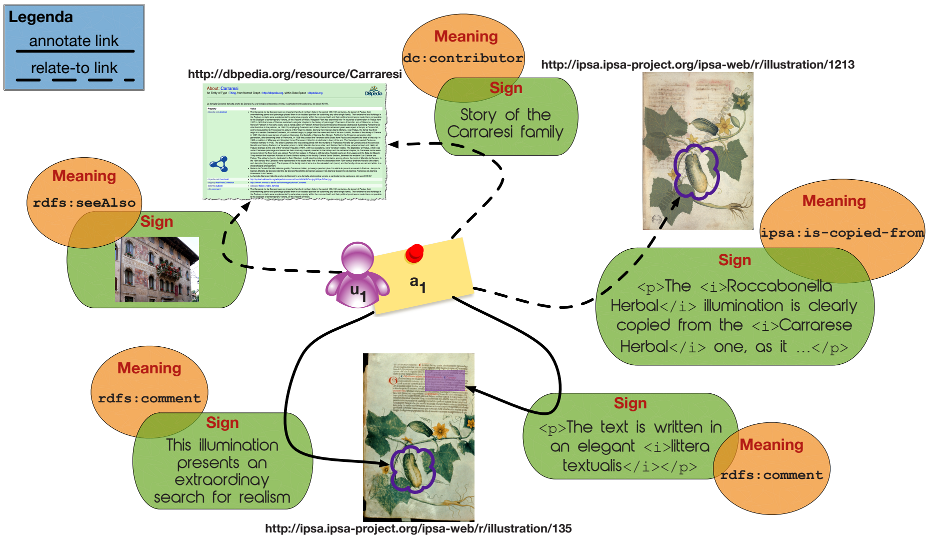
Legenda
annotate link
relate-to link
http://dbpedia.org/resource/Carraresi
Meaning
dc:contributor
Sign
Story of the Carraresi family
http://ipsa.ipsa-project.org/ipsa-web/r/illustration/1213
Meaning
rdfs:seeAlso
Sign
Meaning
ipsa:is-copied-from
Sign
<p>The <i>Roccabonella Herbal</i> illumination is clearly copied from the <i>Carrarese Herbal</i> one, as it ...</p>
u1
a1
Meaning
rdfs:comment
Sign
This illumination presents an extraordinay search for realism
Sign
<p>The text is written in an elegant <i>littera textualis</i></p>
Meaning
rdfs:comment
http://ipsa.ipsa-project.org/ipsa-web/r/illustration/135

**Fig. 2.** Example of annotation.

SIAR (*Sistema Informativo Archivistico Regionale*) was a project supported by the Italian Veneto Region, the aim of which was to design and develop a digital archive system. The main goal of the SIAR project was to develop a system for managing and sharing archival metadata in a distributed environment and to allow archivists to describe archival material in a collaborative fashion [8].

The context of the work is defined by a group of archivists working in the territory and centrally coordinated by a management and control office of the Veneto Region. The main task of the archivists is to describe the archives of pertinence and produce four main elements: an archival tree organizing the archival descriptive metadata, the descriptions of the preserver, the description of the producer and the finding aids.

The architecture of the system consists of three layers – data, application, and interface logic layers – in order to achieve a better modularity and to properly describe the behaviour of the service by isolating specific functionalities at the proper layer.

The SIAR system is exposed as a RESTful Web Service which allows us to develop different applications and plug-ins over it in an open, collaborative and scalable way which ensure sustainability over time.

The architecture of SIAR is designed at a high level of abstraction in terms of abstract *Application Program Interface (API)* using an object-oriented approach. In this way, we can model the behaviour and the functioning of SIAR without worrying about the actual implementation of each component. Different alternative implementations of each component can be provided, still keeping a coherent view of the whole architecture of the SIAR system.

We achieve this abstraction level by means of a set of interfaces, which define the behaviour of each component of SIAR in abstract terms. Then, a set of abstract classes partially implement the interfaces in order to define the actual behaviour common to all of the implementations of each component. Finally, the actual implementation is left to the concrete classes, inherited from the abstract ones, that fit SIAR into a given architecture. Furthermore, we apply the abstract factory design pattern, which uses a factory class that provides concrete implementations of a component, compliant with its interface, in order to guarantee a consistent way of managing the different implementations of each component.

Finally, the presentation logic and part of the business logic are implemented via a Liferay Web application, which manages the interaction with the user, controls the flow of the application and translates it into proper *Asynchronous JavaScript Technology and XML (AJAX)* calls to the SIAR RESTful Web Service.

At the core of the system there is the *NEsted SeTs for Object hieRarchies (NESTOR)* model [18, 19], which is composed of two set data models called *Nested Set Model* (NS-M) and *Inverse Nested Set Model* (INS-M); these two set data models allow us to model hierarchically structured resources by means of an organization of nested sets that is particularly well-suited to archives. The set data models are independent from the tree but they are strongly related to it. Together with the archivists we discussed these data models, pointing out that if we apply them to the archives we are able to maintain the hierarchical structure and the context as well as we can do with the tree data structure, but at the same time they granted us new possibilities of overcoming some of the issues that were highlighted in the ideation phase.

The SIAR system is currently used by the archivists of the Veneto Region for describing and accessing the publicly available archival material of the territory and it is available at the following URL: http://siar.regione.veneto.it/

## 5 Digital Linguistics: ASiT

Language Resources (LRs) are very important in the development of applications for overcoming language barriers, documenting endangered languages, and for supporting research of several fields. Given the impact of LRs, the methodological and technological boundaries existing in linguistic projects need to be overcome in order to find common grounds where linguistic material can be shared and re-used over a long period of time. Consequently, a possibly standardized methodology for designing linguistic databases is necessary to develop linguistic resources that fully meet the desiderata of The FLaReNet Strategic Agenda which presented a set of recommendations for the development and progress of LRs in Europe [24]

One of the basic problems we have to deal with when setting up a database with linguistic data is related to the qualitatively and quantitatively different types of data that have to be classified and retrieved. A linguistic database with

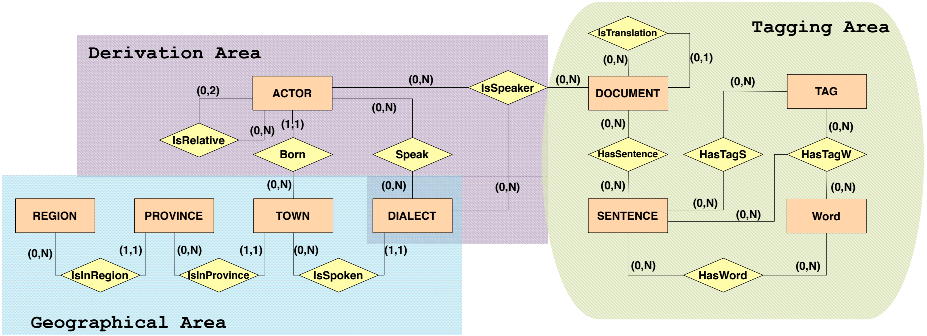

**Fig. 3.** Entity-Relationship Diagram for the ASIt Digital Library.

the function of the old linguistic atlases (and hopefully many more) contains in addition to the obvious linguistic data also many other kinds of data among those information about geographic locations, the type of inquiries adopted to gather the data, the speakers who have delivered the data, all of them being relevant to the linguistic analysis and therefore to be made accessible to the user.

The ASIt (Syntactic Atlas of Italy) linguistic project builds on a long standing tradition of collecting and analyzing linguistic corpora, which has given rise to different studies and projects over the years [1,10–12]. Research on the syntax of Italian is of great interest to several important lines of research in linguistics: it allows comparison between closely related varieties (the dialects), hence the formation of hypotheses about the nature of cross-linguistic parametrization; it allows contact phenomena between Romance and Germanic varieties to be singled out, in those areas where Germanic dialects are spoken; it allows syntactic phenomena of Romance and Germanic dialects to be found, described and analyzed to a great level of detail [10]. The conceptual model of ASIt is depicted in Figure 3.

The ASIt Digital Library System was originally intended to support the first line of research, i.e. comparison between closely related Italian varieties [1]. The corpus to be automatically handled was firstly envisioned and secondly mapped on a conceptual schema in order to be general enough to handle diversified geolinguistic projects with tagging on different linguistic units. This was a crucial methodological investment to support the other lines of research, specifically those involving the relationship between Romance and Germanic varieties and investigated in a multidisciplinary and collaborative project, “Cimbrian as a test case for synchronic and diachronic language variation” [10].

Exposing linguistic data as Linked Open Data enhances the interoperability between existing linguistic datasets and allows for their integration with other resources that use a Resource Description Framework (RDF) approach such as lexical-semantic resources already available as Linked Data, e.g. a general knowl-

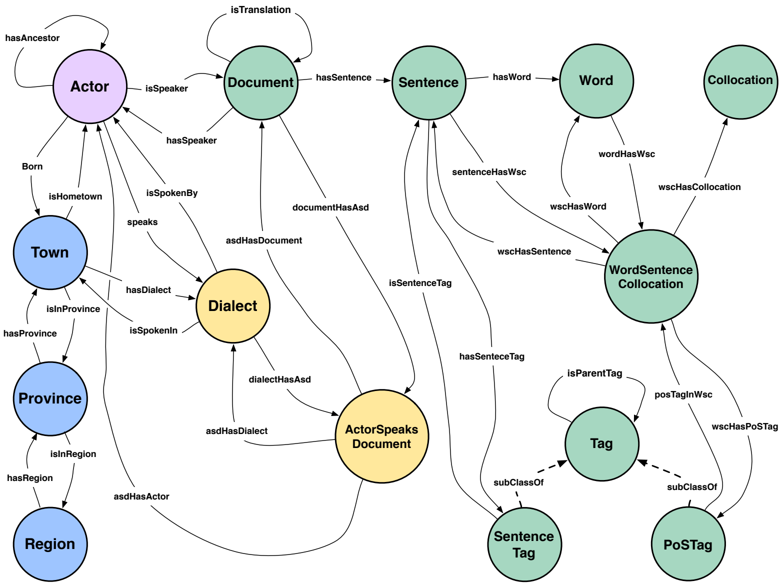

**Fig. 4.** Diagram representing the RDF/S defined for the ASIt enterprise.

edge base like DBpedia[1], or linguistic resources like WordNet[2] or Wiktionary[3]. In order to make ASIt re-usable and interoperable, we defined the ASIt Linguistic Linked Dataset based on the conceptual schema of the curated database [14]. In Figure 4 we report the main classes and properties defining the RDF schema.

The generalizability of the ASIt approach that is materialized in the developed conceptual schema has been shown in a recent test case [11]: the DFG-Projekt PO 1642/1-1.[4] The objective of this project is the synchronic and diachronic analysis of the syntax of Italian and Portuguese relative clauses. Since the project aimed at investigating a set of phenomena related to different types of relative clauses, syntactic phenomena under investigation are captured through a new dedicated sentence level tag set tailored for this project. This database is the first attempt to investigate different types of relative clauses in a corpus of spoken colloquial language in a systematic way. The challenge consisted in adapting the tools of the ASIt project to the corpus data, i.e. adapting a design originally created to deal with a purely experimental setting to a much freer and less controlled set of data coming from a pre-existing corpus.

[1] http://wiki.dbpedia.org

[2] https://wordnet.princeton.edu

[3] https://www.wiktionary.org

[4] http://ims.dei.unipd.it/websites/portuguese-relclauses/index.html

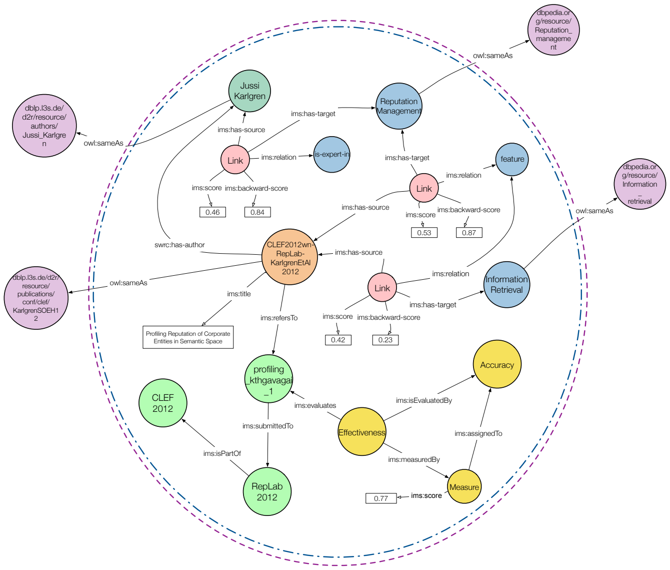

**Fig. 5.** An example of RDF graph showing how expertise topics and expert profiles are used for enriching IR experimental data.

## 6 Scientific Data: The LoD DIRECT System

The importance of research data is widely recognized across all scientific fields as this data constitutes a fundamental building block of science. Recently, a great deal of attention was dedicated to the nature of research data [13] and how to describe, share, cite, and re-use them in order to enable reproducibility in science and to ease the creation of advanced services based on them [16, 17, 23].

Nevertheless, in the field of *Information Retrieval (IR)*, where experimental evaluation based on shared data collections and experiments has always been central to the advancement of the field [21], the *Linked Open Data (LOD)* paradigm has not been adopted yet and no models or common ontologies for data sharing have been proposed. So despite the importance of data to IR, the field does not share any clear ways of exposing, enriching, and re-using experimental data as LOD with the research community.

We discuss an example of the outcomes of the semantic modelling and automatic enrichment processes applied to the use case of discovering, understanding and re-using the experimental data; the details of the full RDF model are reported in [22]. Figure 5 shows an RDF graph, which provides a visual representation of how the experimental data are enriched. In particular, we can see the relationship between a contribution and an author enriched by expertise topics, expert profiles and connections to the LOD cloud, as supported by the *Distributed Information Retrieval Evaluation Campaign Tool (DIRECT)* system which provides the conceptual model for representing and enriching the data [4, 6].

In this instance, the author (*Jussi Karlgren*) and the contribution (*KarlgrenEtAl-CLEF2012*) are data derived from the evaluation workflow, whereas all the other information are automatically determined by the enrichment process. The adopted methodology for expertise topics extraction determined two main topics, "reputation management" and "information retrieval", which are related to the *KarlgrenEtAl-CLEF2012* contribution. We can see that *KarlgrenEtAl-CLEF2012* is featured by "reputation management" with a score of 0.53 and by "information retrieval" with 0.42, meaning that both these topics are subjects of the contribution; the scores (normalized in the interval $[0, 1]$) give a measure of how much this contribution is about a specific topic and we can see that in this case it is concerned a bit more with reputation management than with information retrieval. Furthermore, the backward-score gives us additional information by measuring how much a contribution is authoritative with respect to a scientific topic. In Figure 5, we can see that *KarlgrenEtAl-CLEF2012* is authoritative for reputation management (backward-score of 0.87), whereas it is not a very important reference for information retrieval (backward-score of 0.23). Summing up, we can say that if we consider the relation between a contribution and an expertise topic, the score indicates the pertinence of the expertise topic within the contribution; whereas the backward score indicates the pertinence of the contribution within the expertise topic. The higher the backward score, the more pertinent is the contribution for the given topic.

This information is confirmed by the expert profile data; indeed, looking at the upper-left part of Figure 5, the author *Jussi Karlgren* is considered "an expert in" reputation management (backward-score of 0.84), even if it is not his main field of expertise (score of 0.46).

All of this automatically extracted information enriches the experimental data enabling for a higher degree of re-usability and understandability of the data themselves. In this use case, we can see that the expertise topics are connected via an `owl:sameAs` property to external resources belonging to the DBPedia[5] linked open dataset. These connections are automatically defined via the semantic grounding methodology described below and enable the experimental data to be easily discovered on the Web. In the same way, authors and contributions are connected to the DBLP[6] linked open dataset.

In Figure 5 we can see how the contribution (*KarlgrenEtAl-CLEF2012*) is related to the experiment (*profiling_kthgavagai_1*) on which it is based. This experiment was submitted to the *RepLab 2012* of the evaluation campaign *CLEF 2012*. It is worthwhile to highlight that each evaluation campaign in DIRECT is defined by the name of the campaign (CLEF) and the year it took place (e.g., 2012 in this instance); each evaluation campaign is composed of one or more tasks identified by a name (e.g., RepLab 2012) and the experiments are treated as submissions to the tasks. Each experiment is described by a contribution

[5] `http://www.dbpedia.org/`

[6] `http://dblp.l3s.de/`

which reports the main information about the research group which conducted the experiment, the system they adopted, developed and any other useful detail about the experiment.

We can see that most of the reported information is directly related to the contribution and they allow us to explicitly connect the research data with the scientific publications based on them. Furthermore, the experiment is evaluated from the "effectiveness" point of view by using the "accuracy" measurement which has 0.77 score. Retaining and exposing this information as LOD on the Web allow us to explicitly connect the results of the evaluation activities to the claims reported by the contributions.

The described RDF model has been realized by the DIRECT [4, 6] system which allows for accessing the experimental evaluation data enriched by the expert profiles created by means of the techniques that will be described in the next sections. This system is called LOD-DIRECT and it is available at the URL: `http://lod-direct.dei.unipd.it/`.

The data currently available include the contributions produced by the *Conference and Labs of the Evaluation Forum (CLEF)* evaluation activities, the authors of the contributions, information about CLEF tracks and tasks, provenance events and the above described measures. Furthermore, this data has been enriched with expert profiles and expertise topics which are available as linked data as well.

At the time of writing, LOD-DIRECT allows access to 2, 229 contributions, 2, 334 author profiles and 2, 120 expertise topics. Overall, 1, 659 experts have been individuated and on average there are 8 experts per expertise topics (an expert can have more than one expertise of course).

## References

1. Agosti, M., Benincà, P., Di Nunzio, G., Miotto, R., Pescarini, D.: A digital library effort to support the building of grammatical resources for Italian dialects. In: IRCDL. pp. 89–100 (2010)
2. Agosti, M., Bonfiglio-Dosio, G., Ferro, N.: A historical and contemporary study on annotations to derive key features for systems design. International Journal on Digital Libraries (IJDL) 8(1), 1–19 ( 2007)
3. Agosti, M., Conlan, O., Ferro, N., Hampson, C., Munnelly, G.: Interacting with digital cultural heritage collections via annotations: The CULTURA approach. In: Marinai, S., Marriot, K. (eds.) Proc. 13th ACM Symposium on Document Engineering (DocEng 2013). pp. 13–22. ACM Press (2013)
4. Agosti, M., Di Buccio, E., Ferro, N., Masiero, I., Peruzzo, S., Silvello, G.: DIRECTions: Design and specification of an IR evaluation infrastructure. In Proceedings of the Third International Conference of the CLEF Initiative (CLEF 2012). LNCS 7488, Springer (2012)
5. Agosti, M., Ferro, N.: A formal model of annotations of digital content. ACM Transactions on Information Systems (TOIS) 26(1), 3:1–3:57 (2008)
6. Agosti, M., Ferro, N.: Towards an evaluation infrastructure for DL performance evaluation. In Evaluation of Digital Libraries: An Insight into Useful Applications and Methods. pp. 93–120. Chandos Publishing, Oxford, UK (2009)

7. Agosti, M., Ferro, N., Frommholz, I., Thiel, U.: Annotations in digital libraries and collaboratories – Facets, Models and Usage. In Proc. 8th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2004). pp. 244–255. LNCS 3232, Springer (2004)
8. Agosti, M., Ferro, N., Rigon, A., Silvello, G., Terenzoni, E., Tommasi, C.: SIAR: A user-Centric digital archive system. In Digital Libraries and Archives - Proc. 7th Italian Research Conference (IRCDL 2011). pp. 87–99. CCIS 249, Springer (2011)
9. Agosti, M., Ferro, N., Silvello, G.: Digital library interoperability at high level of abstraction. Future Generation Computer Systems 55, 129–146 (2016)
10. Agosti, M., Alber, B., Di Nunzio, G.M., Dussin, M., Rabanus, S., Tomaselli, A.: A curated database for linguistic research: The test case of Cimbrian varieties. In Proc. of the Eight International Conference on Language Resources and Evaluation (LREC'12). European Language Resources Association (ELRA), (2012)
11. Agosti, M., Di Buccio, E., Di Nunzio, G.M., Poletto, C., Rinke, E.: Designing A long lasting linguistic project: The case study of ASIT. In Proc. of the Tenth International Conference on Language Resources and Evaluation LREC 2016)
12. Benincà, P., Poletto, C.: The ASIS enterprise: A view on the construction of a syntactic atlas for the Northern Italian dialects. In Nordlyd 34: 35-52. Monographic issue on Scandinavian Dialects Syntax (2007)
13. Borgman, C.L.: Big Data, Little Data, No Data. MIT Press (2015)
14. Di Buccio, E., Di Nunzio, G.M., Silvello, G.: A curated and evolving linguistic linked dataset. Semantic Web 4(3), 265–270 (2013)
15. Candela, L., Castelli, D., Ferro, N., Ioannidis, Y., Koutrika, G., Meghini, C., Pagano, P., Ross, S., Soergel, D., Agosti, M., Dobreva, M., Katifori, V., Schuldt, H.: The DELOS digital library reference model. Foundations for Digital Libraries. ISTI-CNR at Gruppo ALI, Pisa, Italy (2007)
16. Ferro, N.: Reproducibility challenges in information retrieval evaluation. ACM Journal of Data and Information Quality (JDIQ) 8(2), 8:1–8:4 (2017)
17. Ferro, N., Fuhr, N., Järvelin, K., Kando, N., Lippold, M., Zobel, J.: Increasing reproducibility in IR: Findings from the Dagstuhl Seminar on "Reproducibility of Data-Oriented Experiments in e-Science". SIGIR Forum 50(1) (2016)
18. Ferro, N., Silvello, G.: NESTOR: A formal model for digital archives. Information Processing & Management 49(6), 1206–1240 ( 2013)
19. Ferro, N., Silvello, G.: Descendants, ancestors, children and parent: A set-based approach to efficiently address XPath primitives. Information Processing & Management 52(3), 399–429 (2016)
20. Gonçalves, M.A., Fox, E.A., Watson, L.T., Kipp, N.A.: Streams, Structures, Spaces, Scenarios, Societies (5S): A formal model for digital Libraries. ACM Transactions on Information Systems (TOIS) 22(2), 270–312 ( 2004)
21. Harman, D.K.: Information Retrieval Evaluation. Morgan & Claypool Publishers, USA (2011)
22. Silvello, G., Bordea, G., Ferro, N., Buitelaar, P., Bogers, T.: Semantic representation and enrichment of Information Retrieval experimental data. International Journal on Digital Libraries (IJDL) 18(2):145-172, 2017.
23. Silvello, G., Ferro, N.: "Data Citation is Coming". Introduction to the special issue on data citation. Bulletin of IEEE Technical Committee on Digital Libraries (IEEE-TCDL) 12(1), 1–5 (May 2016)
24. Soria, C., Calzolari, N., Monachini, M., Quochi, V., Bel, N., Choukri, K., Mariani, J., Odijk, J., Piperidis, S.: The language resource strategic agenda: The flarenet synthesis of community recommendations. Language Resources and Evaluation 48(4), 753–775 (2014)