Department of
Information Engineering
and Computer Science DISI

UNIVERSITY
OF TRENTO - Italy

DISI - Via Sommarive 14 - 38123 Povo - Trento (Italy)
http://www.disi.unitn.it

A METADATA-ENABLED SCIENTIFIC DISCOURSE PLATFORM

Fausto Giunchiglia, Ronald Chenu-Abente,
Hao Xu and Uladzimir Kharkevich

December 2010

Technical Report # DISI-10-069

A Metadata-Enabled Scientific Discourse Platform

Fausto Giunchiglia, Ronald Chenu-Abente, Hao Xu, Uladzimir Kharkevich

[1] Dipartimento di Ingegneria e Scienza dell'Informazione – Università degli Studi di Trento
Via Sommarive, 14 I-38123 Povo, Italy.
{fausto, chenu, hao, kharkevi}@disi.unitn.it

Abstract. Scientific papers and scientific conferences are still, despite the emergence of several new dissemination technologies, the de-facto standard in which scientific knowledge is consumed and discussed. While there is no shortage of services and platforms that aid this process (e.g. scholarly search engines, websites, blogs, conference management programs), a widely accepted platform used to capture and enrich the interactions of research community has yet to appear. As such, we aim to create new ways for the members and interested people working in research communities to interact; before, during and after their conferences. Furthermore, to serve as a base to these interactions, we want not only to obtain, format and manage a body of legacy and new papers related to this community but also to aggregate several useful information and services to the environment of a discourse platform.

Keywords: Metadata, Scientific Discourse, Semantic Search, Lightweight Ontologies, Social Network, Semantic Platform.

1 Introduction

The scientific paper or article was introduced during the 17th century when the first academic journals appeared. Since then these papers, along with the scientific conferences or symposiums, have become the cornerstones of the scientific community and research (as described in [1] and [2]).

Currently, as detailed in [3] and not very unlike those early times, when an author wants to publish a scientific paper he has to submit a physical and or digital copy of it to an academic journal, where it goes through a process of Peer reviewing to determine if its publication is suitable (with similar process occurs in the case of submitting papers in conferences and workshops). Furthermore, the most common (and sometimes effective) way of discussing ideas with colleagues is through the age-old tradition of organizing scientific conferences or symposiums.

This model has remained mostly undisturbed up to now, even with the transition to the electronic era reducing the costs implied in the dissemination process and the advent of the Internet and the Web providing new ways of contact and interaction. As such, several studies (for example [4] and [5]) have been carried out to find the existing limitations on the creation, dissemination, evaluation and credit attribution of these scientific artifacts. Furthermore, new types of less formal artifacts like web pages, blogs, comments, bookmark sites among others have also been increasingly popular in scientific environments. While these emerging web-based new types of scientific artifacts are generally less well-regarded than the traditional ones, it is nonetheless, irrefutable that they also are being used to disseminate discuss and structure and, ultimately, advance scientific knowledge.

To bridge these two seemingly antagonistic approaches we propose the creation of discourse platform where members of a scientific community can have access to the conventional papers and information to aid and enable visiting conventional conferences but, at the same time, also have access to new web-based and semantic-enabled services to enrich their interactions and contributions.

In particular, we want to create new ways for the members and interested people working in the AI research community to interact; before, during and after their conferences. To serve as a base to these interactions, we want not only to obtain, format and manage a body of legacy and new papers related to this community but also to aggregate several (previously dispersed in several sites) useful information and services to the environment of a discourse platform. This discourse platform, which is now being tested and would become used in 2011, would allow the members the community to discuss their work and to share their content (e.g. presentations, videos, notes and pictures).

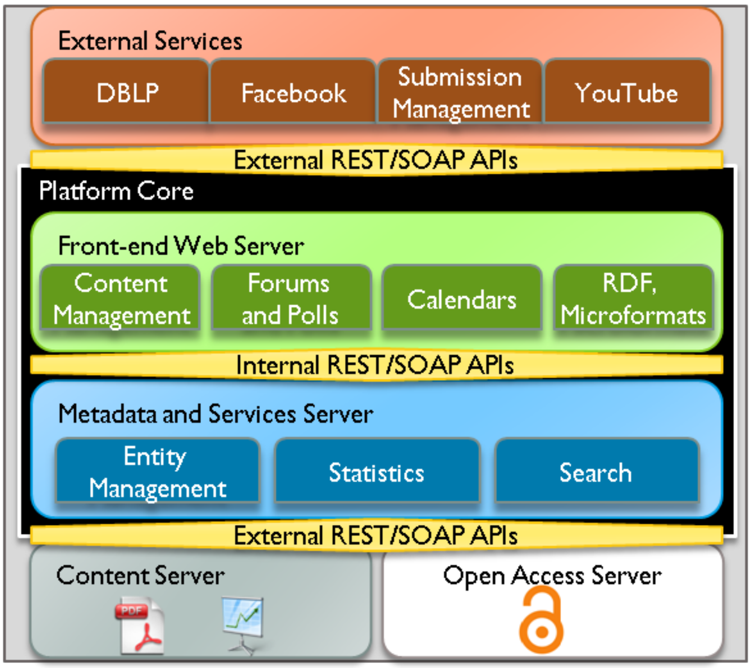
This paper first discusses the architectural details for the discourse platform in Section 2. We then focus on explaining the main metadata structures that make possible to capture all the information needed and enable services, to later offer concrete working examples of the platform's services in Section 4. The related works of the platform are cited and disused in Section 5. Finally, throughout the whole paper but especially in Section 6, we focus in the opportunities and advantages of this semantic-enabled discourse platform and discuss the future work and applications.

2 Architectural Analysis

As explained in the introduction, the main objective in this paper is to bridge the conventional paper/conference-based scientific discourse with the new web-enabled one. This section will present the general architecture and some details of the most important subsystems of a discourse platform that aims to comply with that objective.

2.1 General Platform Architecture

The diagram at Figure 1 shows the general architecture of the scientific discourse platform.



More specifically, in Figure 1, the following macro-elements may be identified:

- *Platform Core*: the central dark box denotes the platform's core elements. These are
  - *Front-end Web Server*: currently implemented using Drupal[1], this component is not only in charge of generating the web pages that would ultimately give the users access to all the features of the platform but also of interacting directly with other available web services like YouTube or Google Scholar.
  - *Back-end Metadata and Services Server*: currently implemented by a custom metadata and semantic managing application developed by the Knowdive[2] group, the backend offers both conventional and semantic-based structure and services. This component may also interact directly with more data-based external services.

  Note that both of these components interact by using an internal REST/SOAP API.
- *External Services*: several platforms and services; like scientific paper metadata (DBLP[3]), currently existing social networks, submission management tools for conferences, among others; currently offer interesting and useful features for

[1] www.drupal.org
[2] http://dit.unitn.it/~knowdive/
[3] http://dblp.uni-trier.de/

finding, accessing and discussing scientific material. As opposed to trying to re-implement everything, we would like to take these tried-and-true approaches and aggregate them coherently and enrich them into our discourse platform.

- *Content Server*: the actual content (papers, presentations, etc.) is left in their original content servers while the metadata is stored into our back-end metadata server. This allows our discourse platform to avoid any copyright or licensing complications that may arise when handling and copying the actual content to be discussed and still be able to refer to the actual content by the use of URLs.
- *Open Access Server*: to comply with the Open Access Metadata Harvesting[4] Protocol, we have created a separate Open Access server that contains the subset of all the available metadata necessary to comply with this standard. By complying with this standard all of the papers managed in our platform would be indexed by the major scientific search engines.

To specify more details about the Metadata and Services Server from Figure 1, the next subsection will give a quick introduction on how the metadata is handled in the platform.

2.2 Metadata Architecture

To represent all the entities, the files and media that the back-end metadata and services server handles we use an upcoming entity-based model, to provide a uniform representation of objects in both the real and the virtual world. In this model, an entity En, is described by its metadata and defined as:

$$En =< id; type; Attr; Rel; S > \quad (1)$$

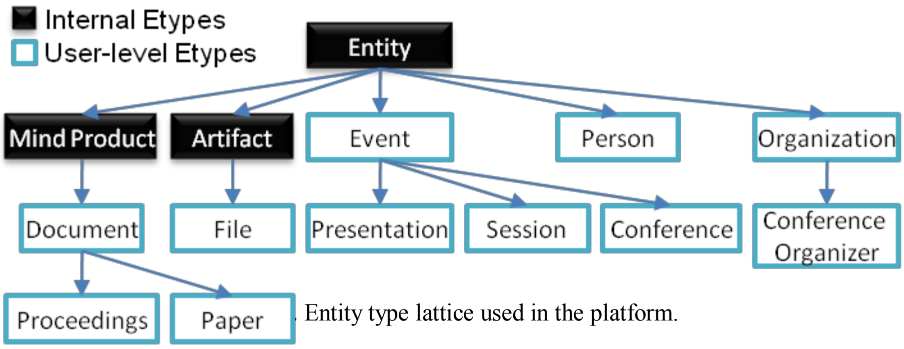
Where:

- *id*: is a unique identifier (e.g., an URI).
- *type*: is the type of entity, that is, the category to which it belongs to (e.g., the entity John" is of type Person).
- *Attr*: is a set of attributes composed of pairs *attr =< attr_name; attr_value >* describing the properties (e.g. John's date of birth is 02/01/88) of that particular entity.
- *Rel*: is a set of relational attributes composed of pairs *rel =< rel_name; rel_value >* describing the entity's relations (e.g., John is friendOf Paul") with other entities.
- *S*: is a set of services that can be leveraged on that specific entity; for example, a service "send email" can be enabled on the Person entity.

An interesting aspect of this model is that the entity types that are encoded on the *type* property of the entity can be arranged into an entity type lattice. In this way, the specific type to which an entity belongs to is used to infer their attributes and services. Furthermore, the hierarchy that is defined in the lattice (like the one defined for the platform in Figure 2) allows the definition of derived etypes through the inheritance and extension of the metadata and services from the parent types.

---

[4] www.openaccess.org



. Entity type lattice used in the platform.

As an example, consider the entity type 'Paper' (shown in Figure 2); this type would inherit metadata and services from the types Entity, Mind Product and Document. The concrete inheritance and extension that takes place will be shown in the next section.

Finally, to exemplify a semantic-enabled service offered by the platform, the next subsection will discuss semantic search, and how it is implemented in the platform.

2.3 Semantic Search

The Semantic entity-centric search engine from the platform is implemented by integrating a faceted search approach (e.g. [6]) with a semantic search approach (e.g. [7]), where the former approach is used to specify constraints on various entity attributes and the later approach is used to implement a semantic matching of entity attributes to specified constraints. For instance, it allows the user to search for papers with the topic 'semantic search' and authors from 'Italian universities' and to find papers about the 'concept search' approach with authors from 'Trento university' (that is assuming that the underlying knowledge base contains all of the necessary concepts and relations to allow for such inference).

Note that in the previous example, faceted search is used to specify two constraints on the paper entity, namely topic:'semantic search' and author.affiliation:'Italian universities'. A semantic search approach is used to perform the matching between these individual constraints and entity attributes, i.e., it allows us to compute that the phrase 'concept search' has more specific meaning than phrase 'semantic search' and that Trento is a city which is located in Italy.

More generally, a query is represented as an arbitrary boolean combination of constraints where each constraint is a triplet <*attr_name (an), relation (r), attr_value (av)*>. This is interpreted as an atomic query for an entity which has the attribute name equal or more specific than *an* and also the attribute value which is in a relation *r* with the attribute value *av*. In the example above, *an* ='topic', *r* = 'more specific', and *av* = 'semantic search'.

Semantic search on individual attribute names and values is implemented by using the Concept Search approach [8]. Concept Search is an information retrieval approach which extends syntactic search with semantics in order to address the problems related to the ambiguity of natural language (e.g. the problems of polysemy and synonymy) by substituting words, when possible, with concepts.

To address problems related to complex concepts and related concepts, syntactic matching of words and phrases is extended to semantic matching of complex concepts, where complex concepts are extracted by analyzing meaning of natural language phrases and semantic matching is based on computing subsumption relations between complex concepts. The main idea behind concept search approach is to reuse highly optimized retrieval models and data structures of syntactic search and preserve their efficiency while allowing for improved results when high-quality semantic information is provided. For instance, the semantic matching ([9] and [10]) of complex concepts, i.e. the core building block in the concept search approach, is implemented by using the inverted index technology.

Semantic Search, as explained in this subsection, is the first semantic-based feature to be enabled in the current discourse platform

3 Entity-Centric Document Metadata

Papers, and more generally scientific documents are currently the center of scientific discourse and progress [3]. As such, this section will further specify the metadata and types that are used within the proposed discourse platform to represent and offer its services. Using the entity type lattice from Figure 2 as a guide, this section will describe the entity types that are used for representing scientific papers within discourse platform.

3.1 Mind Products

Since the entity type "entity" contains attributes that are too abstract, for the purposes of this explanation, the description will start from the entity type "Mind product". Within this context, mind product is referred to as to any piece of intellectual work created by the human mind. Mind product is a fairly general entity type and its specification is given in Table 1.

Table 1. Mind Product entity type specification.

Attribute	Datatype	Description
Author	<Entity>[ ]	The set of entities that participated in the creation of the artifact (e.g. a person, a software company).
Representation / Representation of	<Artifact>[ ]	The set of entities of type Artifact that represent a concrete representation (instance) of this mind product. e.g. A particular pdf file representing a document or a video clip for a given presentation.

The "representation" relational attribute is particularly interesting as, it serves as the connection between the Document entity, which contains all the metadata of the document as a mind product and the actual file that contains the data of that document.

3.2 Documents

Documents refer to artifacts created to transfer information or support claims. Examples of entities covered by this entity type include papers, email, presentations and even videos. More specifically, Table 2 contains the attributes that belong to the entity type document.

Table 2. Document entity type specification.

Attribute	Datatype	Description
Title	Semantic String	A short text describing the document's subject. E.g. Applied Mathematics, A Midsummer Night's Dream.
Language	String	Language of the resource. e.g. English, Italian.
Coverage	Semantic String	The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant. e.g. 16-19 century, Italy.
Editor	<Entity>[ ]	A set of the persons or organizations that edited (i.e. introduced changes or aggregated) this document.
Source / Source of	<Mind Product>[ ]	Related resource/s from which this document is derived or based. Source of is the inverse relation.
Reference / Reference of	<Mind Product>[ ]	A related (but external to the document) resource that is referenced, cited, or otherwise pointed to by the document. Reference of is the inverse relation,
Version / Version of	<Docu-ment>[ ]	The set of documents that were versioned (i.e. are newer) from this document instance. Version of is the inverse relation
Split / Split of	<Docu-ment>[ ]	The set of entities that start a new branch for the current document instance.Split of is the inverse relation.
Merge / Merge of	<Docu ment>[ ]	The set of documents from which the current document instance was branched from. Merge of is the inverse relation .

Of special interest in Table 2 are the relational attributes like Version and Reference. These attributes are used to link documents into semantic graphs that can later be navigated to help the user find the information he is looking for.

Also note that some of the datatypes from Table 2 are identified as “Semantic Strings” (e.g. Title, Coverage). These semantic strings, besides having the normal text string, also have complex concepts attached to it. So for example, if a semantic string contains the regular string: “I was drinking good java while programming in good Java code” the first occurrence of the word ‘java’ would have the concept of ‘coffee’ attached to it, while the second occurrence would have the concept of ‘Programming language’ attached to it (more in depth related information can be found in works about Lightweight Ontologies like [11]). Finally, it is assumed that this word to concept conversion process (WSD) is done through direct human interaction; either by the authors themselves or by other users of the platform. The project Insemtives[5] (Incentives for Semantics) is an example of project focused in encouraging and motivating the creation of semantic annotation of this kind.

3.3 Papers

Finally, the paper entity type shown in Table 3, is used to capture the information specifically about scientific papers.

Table 3. Paper entity type representation.

Attribute	Datatype	Description
Abstract	Semantic String	A text that is related to the main topics/concepts of the paper.
Keywords	Semantic String	A text containing a set of words that are related to the main topics of the paper
DBLP identifier	String	Digital Bibliography and Library Project
Citation / Citation of	<Entity>[ ]	A set of documents that are cited on the paper. Citation of is the inverse relation
Submitted	<Entity>[ ]	Organization(s) to where this paper was submitted to (for approval/review, etc)
Accepted	<Entity>	Organization(s) that accepted/approved this paper.
Publisher	<Entity>[ ]	Person or Organization responsible for making the resource available.
Date of publication	<Moment>	Date of formal issuance by a publisher.

As seen in Table 3, most of the attributes from the paper etype are very specific to the case of scientific documents.

It is worth noting that discussing the subtleties of the implemented entity type management system are beyond the scope of this paper, so topics like the concrete differences between this system and regular object-oriented inheritance systems, multiple entity type inheritance, and semantic enabled attribute names will all be left for future works more focused on this aspect.

[5] http://www.insemtives.org/

4 Current Implementation and Functionalities

This section will highlight some of the concrete implementation details and higher level functionalities built on top of the previously discussed concepts, models and specifications.

4.1 Entity Markup on Papers

For any scientific write-up, incorporating references to account for the sources and related work is an important task. Fortunately, LaTex[6] and its auxiliary tool BibTex[7] offer a variety of options to make the handling of these bibliographies a much easier task. BibTex, in particular, allows people to store all the information of referenced material into a .bib file (which can be considered an external, flat-file database of references). The author can later link this database to any LaTeX file and proceed with further citing and formatting.

One of the features of the platform proposed in this paper is to offer a more powerful and flexible extension to the current BibTex features. This enriched BibTex is not only able to manage bibliographical papers but also other referenced entities such as authors, institutions and events. This feature can potentially improve:

- *Writing*: allowing authors to import entity-related information preloaded on the platform's database. For example, with the \citeauthor{author key} command the processor will automatically import metadata like affiliation, email and address. Fine-tuning of the appearance can be done by using the BibTex style file (.bst) and templates for widely-used formats (e.g. LNCS, IEEE, ACM) will also be provided.
- *Reading*: the reader of the pdf file will have all entity references appear as links that when clicked, would open the metadata-rich profile of that entity in the platform.

Figure 3 shows an example of a LaTex editing environment and the generated PDF.

[6] www.latex-project.org/
[7] www.bibtex.org/

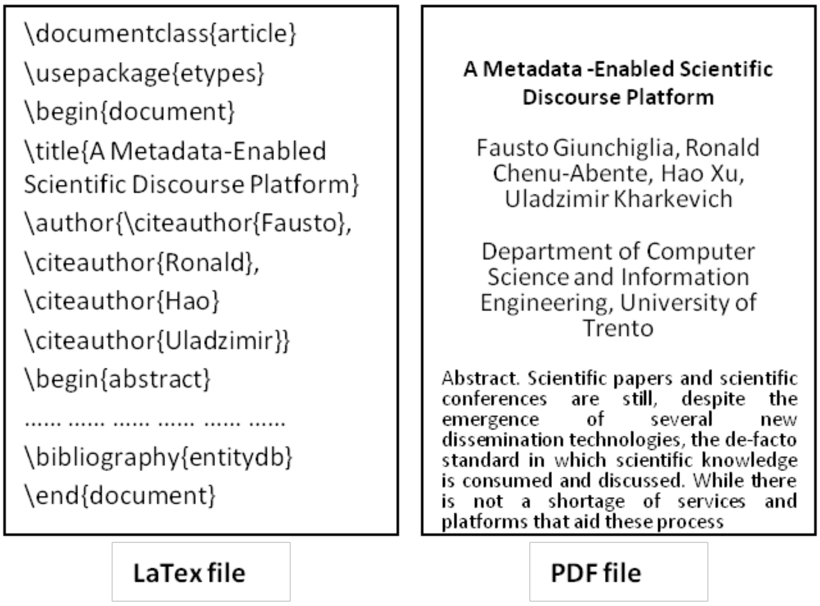


Figure 3. Entity Citation in LaTeX.

In the example from Figure 3, the left part contains the actual LaTeX code as it would show up on the author's editor and the right part contains how it would look when compiled and rendered by a pdf viewer. Note that in the right part, the title, the authors and the institution would all be entities that when clicked would lead to the profile page of the entity in our proposal platform. A more detailed mockup is also available at this address[8] to better convey this feature.

4.2 Representing Discourse and Knowledge

This subsection will apply the previous specifications on how to capture and manage data/metadata to show the ways that the platform actually represents discourse and knowledge.

In particular, the following three are the ways that the platform currently captures discourse and knowledge:

- *Predefined Attributes and Relations*: attributes and relations (e.g. 'source', 'version' and 'split' from Table 2) defined in the platform's entity types are used to capture information and make explicit the relations existing between entities. This information can later be used to calculate and offer several interestingness, and knowledge-related metrics.

[8] http://www.dit.unitn.it/~chenu/papers/misc/Entity%20Markup%20mockup.pdf

- *User-defined Attributes and Relations*: as complement of the predefined attributes and relations, the platform will support the ability for the users to create their own semantics-rich attributes and relations for the available entities. This will allow the creation of custom relations (e.g. the 'I dislike' relation that would connect a person with a paper) and custom attributes (e.g. the 'suggested reference' that when applied to a paper would contain as value the links to other publications that the creator of the annotation thinks are important for it) to further capture the knowledge and discourse of the community.
- *Forums, Polls and other tools*: these more unstructured and conventional tools would serve as think-tank and discussion venue where that the users of the platform would use to agree on the attributes and relations to create (much in the same way that the discussion pages of Wikipedia are used).

4.3 Information display and Navigation

Another use of the infrastructure defined above is to display entity information and offer seamless entity navigation on a web interface. Take as an example of the previous, the series of screenshot that start from Figure 4 and that were taken from the first trial application of the platform loaded to contain information from the IJCAI conference series.

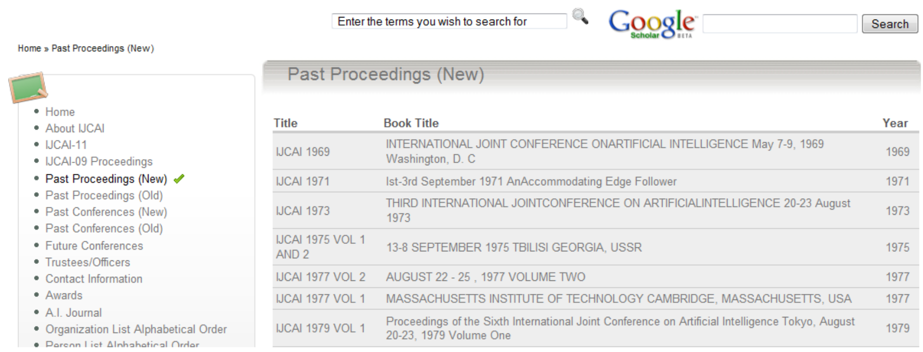
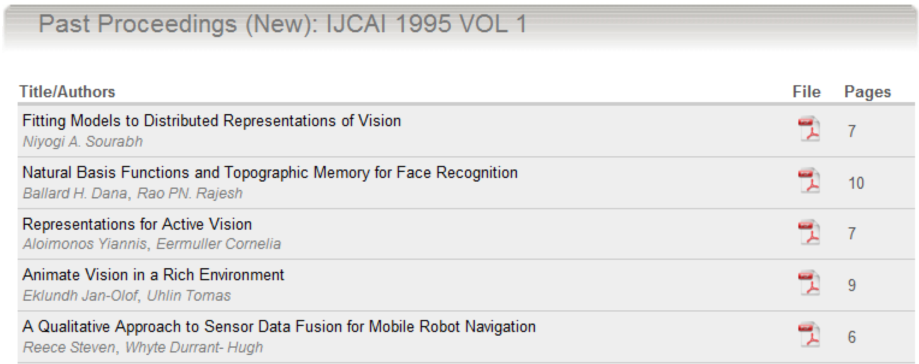


Figure 4. Example page listing the proceedings of the IJCAI conference series.

Figure 4 shows a webpage (from our live demo of the platform) displaying proceedings from the IJCAI conference series. While is possible to use the controls at the top of the screenshot to perform semantic search as described in Subsection 2.3, in this example the user chooses to scroll down and click on the proceedings of year 1995 to navigate to it. The results are shown in Figure 5.


Past Proceedings (New): IJCAI 1995 VOL 1

Title/Authors	File	Pages
Fitting Models to Distributed Representations of Vision Niyogi A. Sourabh		7
Natural Basis Functions and Topographic Memory for Face Recognition Ballard H. Dana, Rao PN. Rajesh		10
Representations for Active Vision Aloimonos Yiannis, Eermuller Cornelia		7
Animate Vision in a Rich Environment Eklundh Jan-Olof, Uhlin Tomas		9
A Qualitative Approach to Sensor Data Fusion for Mobile Robot Navigation Reece Steven, Whyte Durrant- Hugh		6

Figure 5. Example page listing the papers of the selected proceedings.

Figure 5 displays all the papers that where accepted in the previously selected IJCAI 1995. Note that in this web page, the following elements are clickable:

- The title of the paper, this would take the user to the paper profile page where the metadata of the paper can be viewed and edited (provided the user has the permissions to do so),
- The pdf icon to the right, this would take the user to the actual data of the paper. In this case, this is the download of the pdf file from the IJCAI servers and,
- Each individual author of the papers, the result of this is shown on Figure 6.

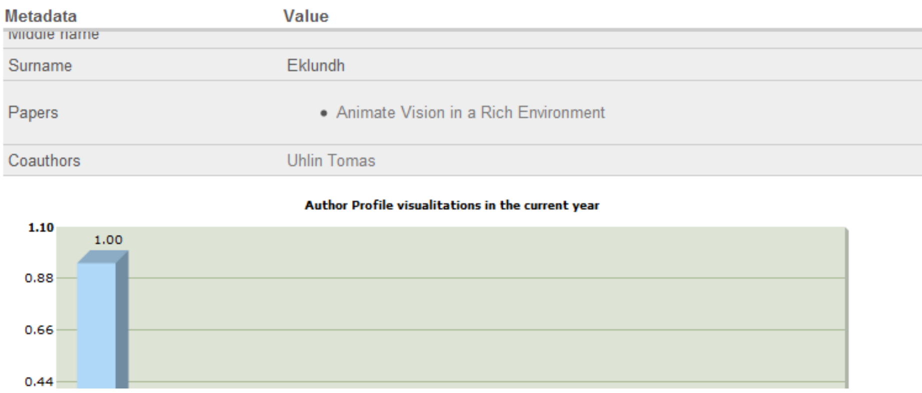


Figure 6. Example page showing an author profile.

Figure 6 shows a person entity profile, which displays with two attributes (middle name and last name); two relations (this person is author of the paper ‘Animate Vision in a Rich Environment’ and also a coauthor of ‘Uhlin Tomas’); and also a graph showing real-time statistics based on the information in the platform.

The predefined attributes and relations in the platform’s entity types are what power the display and navigation from this example. Nevertheless it is also possible, to offer these same services based on custom and user-defined attributes and relations.

5 Related Work

The main part of the conceptual groundwork for the discourse platform presented in this paper is based on the SKO (Scientific Knowledge Object) model introduced in [12] and having as foundation the theory of abstraction found in [13] and [14]. The creation of this still upcoming SKO model, has also taken into consideration several semantic-enabled scientific discourse models covered in [15] and scientific information encoding ontologies like the one detailed in [16]. In a similar fashion, the part of the work related with semantic annotations took into consideration the work on [17] and [18]; while the creation of links and enrichment of documents was aimed to extend [19].

As an example of similar approaches, the following are some of the services and platforms that have already been created to aid scientific production and discourse:

- *Search engines*: a search engine specifically tailored for scientific papers, persons and events (e.g. Google Scholar[9]).
- *Conference websites*: these sites contain announcements and concrete information about the venue and event of a conference. They are also mainly used to register (and pay) for participation (e.g. the ESWC2011 site[10])
- *Conference series websites*: these types of sites contain proceeding (i.e. collections of papers), give various announcements and offer special pages for the organizers to share information (e.g. the JCDL conference series site[11]).
- *Submission management*: used to manage the submission process of papers to a conference (e.g. Easychair[12]).
- *Social Networks*: like Facebook[13], which almost unchallenged as a recreational and commercial social network. We would also want to offer scientific-related resources and content as the very center of the social network (e.g Academia.edu[14]).
- *Conference Portals*: like ACM's[15] have the closest approach to what we want to get to, as they includes several of the features we are interested in (albeit sometimes in a reduced or more basic manner).

The wide variety in the approaches and services mentioned shows that many of the functionalities and services that are considered interesting for scientific research are still somewhat dispersed. It would be normal for a researcher, for example, to first search for a paper in scientific search engine, then find it in the proceedings of a conference site and finally resort to email communication to comment or discuss. Besides the minor inconvenience of using multiple sites, the main problem is that no explicit tracking is kept over the whole process or even the interactions or discussions

[9] http://scholar.google.com
[10] http://www.eswc2011.org/
[11] http://www.jcdl.org/
[12] http://www.easychair.org/
[13] http://www.facebook.com/
[14] http://www.academia.edu/
[15] http://portal.acm.org/

that are taking place. Because of this, for example, it is possible that an author is asked the same questions over and over again.

6 Conclusions

As an answer to the previous situation and to help bridge the distance between the conventional and the web-based scientific discourse, with the development and introduction of this scientific discourse platform the aim is to:

1 Offer improved ways for helping the users find the content, persons or events they are looking for (through semantic search and navigation) along with the context and relations that these may have. The specific improvements will be based on the ability to search for title, authors, keywords and even key concepts related to what the user wants to find and to aid the search based on the navigation of the "related items". Search, access to information and interaction with results, will all be integrated in the same platform and aided by it.

2 Offer a way of commenting, tagging and creation relations of all the content in the site (like persons, conferences and papers) in a certified (i.e. approved and validated by the management of the site) way.
This process is not limited in the standard ways predefined by the system and, as such, users are free add their own types of information to the system (e.g. adding a relation "is opposed to" between two papers, when this relation was not predefined in the system).

3 Provide live services during conferences that, through the use of portable devices, would help attendants to find, keep track and take notes about the events happening. All this integrated into the main platform, which would both keep track of this for the user and for computing meaningful statistics about the conference. While similar services to these already exist, they are fairly localized endeavors proving a sort of "your guide during the conference" services. By integrating with the social network we want to *also* achieve "your preparation to the conference" and "this conference continues online" services that have yet to be coherently presented.

4 Reformat the vast volumes of legacy information into a reusable and easily citable format. This would allow the users to browse information from conference that happened over 30 years ago and also have access to information and services similar to the recent conferences (that were specifically tailored for the system).

The current version of the scientific discourse platform is currently undergoing through a trial period where data of the IJCAI conference series (i.e. 4 thousand papers and 7 thousand authors among others) is being introduced and managed with the objective of testing the platform for its upcoming use. Future work includes the analysis of the usage data from the first use cases to plan and implement corrections and feature extensions, along with new applications of the platform to existing conferences and journals.

References

1. Latour, B.: Science in Action: How to Follow Scientists and Engineers through. Harvard University Press, Cambridge Mass., (1987) USA.
2. Kornfeld, W. A., & Hewitt, C.: The Scientific Community Metaphor. IEEE Transactions on Systems, Man, and Cybernetics , 11 (1981).
3. European Commission: Study on the Economic and Technical Evolution of the Scientific Publication (2006).
4. Casati, F., Giunchiglia, F., Marchese, M.: Publish and perish: why the current publication and review model is killing research and wasting your money. ACM Ubiquity (November 2006).
5. Parnas, D. L.: Stop the Numbers Game. Communications of the ACM , 50 (2007).
6. Hearst, M., Design Recommendations for Hierarchical Faceted Search Interfaces, ACM SIGIR Workshop on Faceted Search (2006).
7. Hildebrand, M., van Ossenbruggen, J., Hardman, L.: An analysis of search-based user interaction on the semantic web. Technical Report INS-E0706, CWI (2007)
8. Giunchiglia, F., Kharkevich, U., Zaihrayeu, I., "Concept search". In Proceedings of ESWC, (2009).
9. Giunchiglia, F., Shvaiko, P., Yatskevich, M., S-Match: an Algorithm and an Implementation of Semantic Matching, in proceedings First European Semantic Web Symposium, ESWS (2004).
10. Giunchiglia, F., and Autayeu, A., and Pane, J., (2010) S-Match: an open source framework for matching lightweight ontologies. Technical Report DISI-10-043, Ingegneria e Scienza dell'Informazione, University of Trento.
11. Giunchiglia, F., Biswanath, D., and Vincenzo, M., Faceted Lightweight Ontologies, Conceptual Modeling: Foundations and Applications Lecture Notes in Computer Science, 2009, Volume 5600/2009, 36-51, DOI: 10.1007/978-3-642-02463-4_3
12. Giunchiglia, F., & Chenu-Abente, R. (2009, 06). *Scientific Knowledge Objects V.1.* Technical Report DISI-09-006, Ingegneria e Scienza dell'Informazione, University of Trento. http://eprints.biblio.unitn.it/archive/00001542/
13. Giunchiglia, F., Walsh, T., Abstract theorem proving, Proceedings of the 11th international joint conference on Artificial intelligence - Volume 1, (1989).
14. Giunchiglia, F., Villafiorita, A., Walsh, T., Theories of abstraction, AI Communications Journal, Volume 10, Numbers 3-4, pg. 167-176, (1997).
15. Groza, T., Handschuh, S., Clark, T., Buckingham Shum, S. and de Waard, A.. A short survey of discourse representation models. In: Proceedings 8th International Semantic Web Conference, Workshop on Semantic Web Applications in Scientific Discourse. Lecture Notes in Computer Science, Springer Verlag: Berlin, Washington DC (2009).
16. Sure, Y., Bloehdorn, S., Haase, P., Hartmann, J. and Oberle., D., The SWRC ontology - semantic web for research communities. In EPIA, pages 218-231, (2005).
17. De Waard, A., & Tel, G., The ABCDE Format Enabling Semantic Conference Proceedings. SemWiki, (2006).
18. Groza, T., Handschuh, S., Möller, K., Decker, S., SALT - Semantically Annotated LaTeX for Scientific Publications. Lecture Notes in Computer Science. Volume 4519/2007, pages 518-532, (2007).
19. Rahtz, S., Oberdiek, H., “Hypertext marks in LATEX: a manual for hyperref”, http://www.tug.org/applications/hyperref/manual.html, (2010).