Project no. 34721

TAGora

Semiotic Dynamics in Online Social Communities

Plan for using and disseminating the knowledge

Period covered: from 01/06/2006 to 31/05/2007   Date of preparation: 31/05/2007
Start date of project: June 1st, 2006   Duration: 36 months
Due date of deliverable: May 31st, 2007   Actual submission date: May 31th, 2007
Distribution: Public   Status: Final

Project coordinator: Vittorio Loreto
Project coordinator organisation name: "Sapienza" Università di Roma
Lead contractor for this deliverable: "Sapienza" Università di Roma

Contents

Tagora

Chapter 1

Exploitable knowledge and its use

In this section we present the exploitable results, defined as knowledge having a potential for industrial or commercial application in research activities or for developing, creating or marketing a product or process or for creating or providing a service. A list of these results is reported in the following tables:

Current Activities

Exploitable Knowledge (description)	Exploitable product(s) or measure(s)	Sector(s) of application	Timetable for commercial use	Patents or other IPR protection	Owner & Other Partner(s) involved
BibSonomy	Web site (service) and Web server (software)	Public, especially researchers	Preliminary Service was online at the beginning of the project, improvements have been done according to project schedule. No commercial application planned.	–	**UNIK**
Ikoru	Web site (service) and Web server (software)	Internet Service Provider, Software Integrators, Consumer Electronics, Cultural Institutions	Version 1 of Web site online (2007), Demonstrations (2007)	Copyright on software	**SONY-CSL**
Tagster prototype	P2P client application	Academic institutions, small companies	–	–	**UNI KO-LD**

Planned Activities

Exploitable Knowledge (description)	Exploitable product(s) or measure(s)	Sector(s) of application	Timetable for commercial use	Patents or other IPR protection	Owner & Other Partner(s) involved
BibSonomy	Web site (service) and Web server (software)	Public, especially researchers	Currently no commercial application planned	–	**UNIK**
Ikoru	Web site (service) and Web server (software)	Internet Service Provider, Software Integrators, Consumer Electronics, Cultural Institutions	Use for museum (2008)	Copyright on software	**SONY-CSL**
Tagster prototype	P2P client application	Academic institutions, small companies	Currently no commercial application planned	–	**UNI KO-LD**

1.1   BibSonomy

The development of BibSonomy serves two purposes.

Firstly, the academic community will get benefit from it. Research groups world-wide struggle with the problem of publication and citation management and usually implement themselves an individual time consuming solution. With BibSonomy, instead, they may manage their internal and public lists of publications in a simple and efficient way. BibSonomy allows for the interoperation of communities, being its stored entries public by default. Even though there is a large demand for this kind of web tools, a commercial exploitation will be rather difficult to realize, as there are similar competitors on the web (like Connotea and CiteULike), who also have been offering their service free of charge.

Secondly, (and this is the main reason for our engagement in BibSonomy) the system allows us to collect valuable data about its complete usage. These data are an essential brick in our quantitative research, as currently no folksonomy benchmark data are available. The system allows us to implement and experiment our own algorithms, and to study their behavior in a living growing environment. BibSonomy already helped us to achieve a significant position among the *Web 2.0* research community.

1.2   Ikoru

The Ikoru system consists of two parts: the Web site and the Web server, whose brief description follows.

The Web site

The Web site is written in a mixture of HTML and Javascript. It was specifically developed to run experiments and gather data for the TAGora project. The web site currently handles photos and audio but might be extended to include video and other contents as well.

The current running version can be visited at `http://www.ikoru.net`. The Web site is constantly under development as new functionalities become necessary. It is open to the public and free of

charge. Sony CSL shall continue to maintain this Web site for at least the duration of the TAGora project. There has been done little advertisement to promote the site and attract users, so far. In fact, the site is still undergoing a testing phase and we think that we must be able to offer a minimum of quality of service to the users before advertising it massively. Traditionally, partnerships within Sony have been in principle possible, but in this case they might be not trivial to set up because of the current tendency of Sony to outsource Web services.

The Web server

The Web server is written from scratch in C++. That means that it handles the following tasks:

- handle HTTP and SOAP requests,

- manage persistent sessions,

- assure the security (login, validating user input),

- store/retrieve tagging data in a persistent storage,

- receive the uploaded files (testing file type, conversion to Web formats, feature extraction),

- search data based on the text and content (for images),

- generate pages dynamically using key/value substitutions, and

- run the analysis scripts

The reason for the development effort was to allow the Ikoru server to run on a wide range of platforms, from embedded devices and personal computer to full range servers.

This software component can be used by third parties and integrated in larger development projects, for example, as a basis of a Content Management System (CMS) for internal company Web sites. It is also available to consumer electronics and computing equipment businesses that would like to include tagging services on their devices or use it as a value-adding service. Contacts within Sony, in particular with the Sony Cybershot devision in Tokyo and the Digital Imaging Group in the United Kingdom, suggest that a transfer of technology may be feasible. Another possible outlet is in the cultural sector. Contacts with the Zentrum fur Kunst and Medientechnologien in Karlsruhe, Germany, showed that cultural institutions want to open up to a Web audience, in *Web 2.0* fashion. The Ikoru platform could be become a contender in this market.

Sony CSL has nor a precise planning nor a time-table to promote this technology. The current strategy is to increase the usability and reliability of the software through its effective use on the Web site and in small scale projects. In the long run, two options are available, either make it available as Open Source in the hope to attract a large user and developer base, or to seek a partnership for a commercial exploitation.

1.3   Tagster prototype

The development of the peer-to-peer tagging system has two different purposes. The first one is to promote tagging as an easy way for personal data organization and metadata exchange. Therefore, the main goal behind Tagster is to implement a distributed folksonomy system in contrast to the commonly centralized tagging services. We expect that this approach of using tagging for distributed data organization will also find substantial interest in the peer-to-peer community. Thus, the efficient sharing of tagging metadata, which is independent from the actual information objects, is one of the most important aspects.

On the other hand - being one of the main goals within Tagora - Tagster will be used to gather comprehensive tagging data from its users. Such information will be very valuable for further

analysis and one question we would like to investigate is whether a different tagging behavior can be observed in a distributed tagging system. Due to the possibility to tag different types of multimedia data we expect to collect richer tagging data in terms of tag co-relations between the media types. This would otherwise only be possible if datasets from different tagging services were integrated which is not always doable in coherent manner. Furthermore, Tagster allows us to implement and test new algorithms in a real system.

An important requirement for the exploitation of Tagster is a stable version of the software to ensure that not only early adopters will be interested but that also common Internet users can be attracted to use the system on a regular basis. Therefore, we are currently focusing completely on improving the prototype and have not planned any action for promoting it yet.

Chapter 2

Dissemination of knowledge

2.1   Overview of dissemination plan

To assure that knowledge generated within the project will be disseminated to a wider scientific and non-scientific audience we conducted and plan to conduct the following actions:

Prepare a White Paper, describing target problems and grand challenges for Semiotic Dynamics Systems. Maximize the dissemination of this paper in the research community as large. Continue improving this document as our understanding of the field and its various challenges get mature;

Foster publications through standard scientific and engineering communication channels. Publish results in the best scientific journals and communicate the results of the project at top conferences. Use all the possible existing communication media to touch the largest possible audience;

Encourage the partners to organize tutorials at major conferences in the different fields that are relevant to the present project. Encourage the partners to contribute to summer schools or other educational activities that touch in particular younger students. Tutorials on subjects related to the project activities will be given by the senior scientists of the network in occasion of international workshops;

Foster the organization of workshop and conferences on themes related to the project;

Disseminate results to the press at large in order to diffuse them as widely as possible;

Establishing, maintaining, updating the project web site and project mailing lists;

Foster the exhibition of demonstrators in industrial exhibitions or in other contexts where the public at large and a broad scientific/engineering audience can get exposure to the ideas of the project. These exhibitions may also represent a platform in which the results of the project are tested;

Foster the preparation of professional dissemination material;

Maximize the availability of tools and experimental platforms developed by the partners during the project to the scientific community as large;

Disseminate the results to the other projects of the Complex System Initiative and to foster synergies and possible joint activities.

The following table summarize the dissemination actions performed and the actions already planned for the next period.

Overview table

Type	Type of audience	Countries addressed	Partner responsible/ involved
1. White Paper	Scientific Community	All	All
2. Publications on International Journals	Scientific Community	All All	All All
3. Publications on Books and Conference Proceedings	Scientific Community	All	All
4. Posters	Scientific Community	All	All
5. Talks, lectures and conference presentations	Scientific Community	All	All
6. Tutorials	Scientific Community	All	All
7. Conferences, Workshop and Summer Schools	Scientific Community	All	All
8. Press, Radio and TV	General public	All	All
9. TAGora Project Website	General public	All	PHYS- SAPIENZA
10. Demonstrations in Industrial and Scientific Exhibitions	Industrial Community	All	All
11. Tools, software, experimental platforms, multimedia presentations presentations	Scientific Community and General public	All	All
12. Dissemination actions toward the other projects of the Complex System Initiative	Scientific Community	All	All
13. Joint Publications	Scientific Community	All	All

2.2   The White Paper

Understanding the relation between the Web and the world involves learning about a complex cycle of interactions, which (to complicate things) occur at radically different scales. Technical developments which are basically to do with computer abilities to pass data between each other, turn out to have strong social effects; computational innovations at the micro scale feed into macro-level effects on the whole of society. An idea, say for an information-sharing protocol, needs a technical engineering design that encapsulates it within a particular social context. The design in context produces a micro-level effect at the level of the individual user's control of his or her computer. But when the number of users of a design within a decentralized structure grows, macro-level effects can be detected, which raise social issues. In large part, these social issues are raised because the social effects were not only not predicted, but they were fundamentally unpredictable.

This is a scenario to which in the last decades scientists have devoted great attention, namely the study of collective phenomena and complex systems. Large systems made up of simple components (for instance atoms or molecules, animal, human or artificial agents) can in fact self-organize themselves, i.e. "acquire a functional, spatial or temporal structure without specific interference from the outside". More precisely, the constituents of such systems are able to develop a complex collective behavior not trivially deducible from the knowledge of the rules that govern their mutual interactions. We then need to analyses these effects in order to understand whether the social effects are good or bad, and to what degree, and also to help determine what new engineering designs might be capable of preserving or enhancing the good effects, or alternatively eliminating the bad. The result, therefore, is a cycle of micro-level engineering and macro-level analysis.

The Web was originally created as a means of two-way communication, writing as well as reading, but as it took off, publishing tools did not develop as strongly as browsers, and reading became more important than writing, a development deplored by many of the more idealistic who looked back to the plain text of the early Internet days with nostalgia. But the recent arrival of tools allowing more information-passing has led to a rebalancing of the Web, and the invention of so-called Web 2.0.

There are micro-level effects of providing tools allowing someone, e.g. to blog, which affect the individual author. However, as the blogosphere, content-hosting sites or recommender systems grow, there are macro-level effects to be noted.

An initial state when writing, editing and publishing were difficult prompted the development of form-based editors, and the individual phenomenon of the wiki. As more users are recruited, increasingly large groups can put together something totally unpredictable such as the extraordinary Wikipedia. Wikipedia is now the first port of call for anyone needing basic information. It is therefore an important rival to traditional undemocratically edited encyclopedias, but also now it is so widely used, its accuracy and veracity are now an issue. The social question is now about how democratic Wikipedia can afford to be.

Blogs, Wikis, and Social Bookmark Tools have rapidly emerged on the Web creating a new scenario that radically change the knowledge production process. We have virtually unbounded storage capabilities and essentially no limits in our ability to interact with other peers. This new knowledge production process is impacting on all aspects of knowledge creation on all types of knowledge and the Web is becoming the most extensive knowledge repository that ever existed. The reason for this immediate success is the fact that no specific skills are needed for participating. A new paradigm is actually gaining impact very quickly in large-scale information systems: Collaborative Tagging. In new web applications (e.g. Flickr `www.flickr.com`, Connotea`www.connotea.org`, Citeulike `www.citeulike.org`, Delicious `de.icio.us`, etc.) people no longer make passive use of online resources. They take instead an active role and enrich resources with semantically meaningful information. Such information consists of terminology (or tags) freely associated by each

user to resources and is shared with users of the online community. Despite its intrinsic anarchist nature, the dynamics of this terminology system spontaneously leads to patterns of terminology common to the whole community or to subgroups of it. Surprisingly, this emergent and evolving semiotic system provides a very efficient navigation system through a large, complex and heterogeneous body of information.

At the moment, there exists no foundational research for these systems, and they provide only very simple structures for organizing knowledge. Individual users create their own structures, but these can currently not be exploited for knowledge sharing. The aim of TAGora is twofold. On the one hand it is important performing analysis of Web 2.0 phenomena to determine what social effects we are now seeing from tagging structures, recommendation systems, and other content that is emergent from large-scale use of social software. On the other hand it is crucial providing theoretical foundations for upcoming Web 2.0 applications and to investigate further applications that go beyond bookmark- and file-sharing.

We strongly believe we are facing a unique opportunity to exploit and give theoretical foundations to the recent, though extremely rapid, developments of emergent semantics in Web-based applications. The common effort of researchers in many different fields could provide the right trigger to face and tame the challenges of Web Science.

And of course one among the most important challenging tasks of this project is the development and fostering of a new culture of interaction between IT and complex systems science. This will require that gaps will be bridged from both sides.

2.3   Publications on international journals:

2.3.1   Published or in Press

- Cattuto, C., Loreto, V. & Pietronero, L., *Semiotic dynamics and collaborative tagging*, Proceedings of the National Academy of Sciences United States of America, 2007, **104**, 1461.

- Cattuto, C., Loreto, V. & Servedio, V. D. P., *A Yule-Simon process with memory*, Europhysics Letters, 2006, **76**, 208-214.

- Cattuto, C., Schmitz, C., Baldassarri, A., Servedio, V. D. P., Loreto, V., Hotho, A., Grahl, M., Stumme, G., *Network Properties of Folksonomies*, AICOM Special Issue on *Network Analysis in Natural Sciences and Engineering*, 2007.

2.4   Publications: Books & Conference Proceedings

2.4.1   Published or in Press

- Abbasi R., Staab S., Cimiano P., *Organizing Resources on Tagging Systems using T-ORG*. Proc. Workshop on "Bridging the Gap between Semantic Web and Web 2.0" at the 4th European Semantic Web Conference ESWC 2007, Innsbruck, Austria, June 7, 2007

- Aurnhammer, M., Hanappe, P. & Steels, L., *Augmenting Navigation for Collaborative Tagging with Emergent Semantics*, Proc. of ISWC2006 International Semantic Web Conference, Springer, LNCS, 2006.

- Dellschaft, K. & Staab, S., *On How to Perform a Gold Standard based Evaluation of Ontology Learning*, Proc. of ISWC-2006 International Semantic Web Conference, Springer, LNCS **4273**, 2006.

Tagora

- Hotho, A., Jäschke, R., Schmitz, C. & Stumme, G. *Das Entstehen von Semantik in BibSonomy*, Social Software in der Wertschöpfung, Nomos, 2007 (in press).

- Hotho, A., Jäschke, R., Schmitz, C. & Stumme, G., Hochberger, C. & Liskowsky, R. editors, *Emergent Semantics in BibSonomy*, Proc. Workshop on Applications of Semantic Technologies, Informatik 2006, Gesellschaft für Informatik, 2006, P-94.

- Hotho, A., Jäschke, R., Schmitz, C. & Stumme, G., Avrithis, Y. S., Kompatsiaris, Y., Staab, S. & O'Connor, N. E. Editors, *Trend Detection in Folksonomies*, Proc. First International Conference on Semantics And Digital Media Technology (SAMT), Springer, 2006, **4306**, 56-70.

- Cattuto, C., Loreto, V. & Pietronero, L., *Semiotic Dynamics in Online Social Communities*, Proceedings of the European Conference on Complex Systems ECCS06, Oxford (UK), 2006.

- Schmitz, C., Grahl, M., Hotho, A., Stumme, G., Cattuto, C., Baldassarri, A., Loreto, Servedio, V. D. P., *Network Properties of Folksonomies*, Proceedings of the WWW2007 International World Wide Web Conference, 2007.

- Szomszor, M., Cattuto, C., Alani, A., O'Hara, K., Baldassarri, A., Loreto,V., Servedio, V. D. P., *Folksonomies, the Semantic Web, and Movie Recommendation*, "Bridging the Gap between Semantic Web and Web 2.0" at the 4th European Semantic Web Conference ESWC 2007, Innsbruck, Austria, June 7, 2007

2.4.2 Submitted for Publication

- Cattuto, C., Baldassarri, A., Servedio, V. D. P., Loreto, V., *Vocabulary growth in collaborative tagging systems*, submitted to the European Conference on Complex Systems ECCS07, 2007.

- Cattuto, C., Baldassarri, A., Servedio, V. D. P., Loreto, V., *Investigating Community Structure In Social Tagging Systems*, submitted to the European Conference on Complex Systems ECCS07, 2007.

- M. Grahl, A. Hotho, G. Stumme: Conceptual Clustering of Social Bookmarking Sites. Submitted to the 7th International Conference on Knowledge Management, Graz, 2007

- Aucouturier J.-J., Pachet F., Roy P., Beurivé A. (2007) *Signal + Context = Better Classification*, submitted to the International Conference on Music Information Retrieval (ISMIR).

2.5 Posters

- M. Aurnhammer, *Evolving Texture Features by Genetic Programming*, 9th European Workshop on Evolutionary Computation in Image Analysis and Signal Processing, Valencia (Spain), 2007.

- C. Cattuto, A. Baldassarri, V. D. P. Servedio, and V. Loreto, *Investigating Community Structure In Social Tagging Systems*, Spring Meeting of the German Physical Society (DPG), Regensburg (Germany), March 26-30th, 2007.

- Franz T., Saathoff C., Görlitz O., Ringelstein C., Staab S. (2006), SEA: Introducing the Semantic Exchange Architecture, presented at the European Semantic Web Conference 2006, Budva Montenegro.

- Aurnhammer M., *Evolving Texture Features by Genetic Programming*, 9th European Workshop on Evolutionary Computation in Image Analysis and Signal Processing, 10th European Conference on Genetic Programming, Valencia (Spain), 11-13 April, 2007.

- Hanappe P., *The Ikoru system: a testbed for collaborative tagging and content-based analysis*, poster at the Sony CSL OpenHouse, Tokyo, June 2007.

2.6 Talks, lectures and conference presentations

- Staab S., Topic Communities in Peer to Peer Networks. Talk at the workshop on *Semantic Network Analysis* at the European Semantic Web Conference 2006, June 11-15th, Budva, Montenegro.

- Nigel Shadbolt, panel on *The Next Wave of the Web*, World Wide Web Conference, Edinburgh, 2006.

- G. Stumme: Semantic Network Analysis of Ontologies. Fachgruppentreffen FGML, 7.10.2006, Hildesheim

- Görlitz O., Tagora: Analysing Folksonomies in Online Communities. Talk at the workshop on *Formation of Social Networks in Social Software Applications*, GI Annual Meeting 2006, October 2-10th, Dresden, Germany.

- Schmitz, C..: *Trend Detection in Folksonomies*, Proc. First International Conference on Semantics And Digital Media Technology (SAMT), Athens, Greece, December 6-8, 2006.

- Staab S., Ontology Learning. Invited Talk at the Inaugural Workshop of the Language, Interaction and Computation Lab at the Center For Mind / Brain Sciences (CIMeC), May 19, 2007, Rovereto, Italy.

- C. Cattuto, *Stochastic Models of Tagging Behavior*, International workshop on "Cooperation in Selfish Systems incorporating TagWorld I", University of Bologna Residential Center, Bertinoro (Italy), May 26-29th, 2006.

- C. Cattuto, *Semiotic Dynamics in Online Social Communities*, European Conference on Complex Systems ECCS06, Saïd Business Schoo, Oxford (UK), September 25-29th, 2006.

- C. Cattuto, *Complex Dynamics and the new Web Language*, 2005 Erice Prize for Science and Peace, Pontifical Academy of Sciences, The Vatican, December 21st, 2006.

- C. Cattuto, *Stochastic Models for Collaborative Tagging*, Noise in Physical Systems (N.i.P.S.) Workshop, Castello di Monterone, Perugia (Italy), March 10th, 2007.

- C. Cattuto, DPG Conference, Working Group Physics of Socio-Economic Systems, *Stochastic Models for Social Tagging Behavior*, Spring Meeting of the German Physical Society (DPG), Regensburg (Germany), March 26-30th, 2007.

- C. Cattuto, *Structure and Evolution of Collaborative Tagging Systems*, Collaborative Knowledge Management workshop (CoKM2007) at the Professional Wissensmanagement conference (WM2007), Potsdam (Germany), March 28-29th, 2007.

- C. Cattuto, *Semiotic dynamics in online communities*, Seminar on Language, Evolution, and the Brain (SLEB), International Institute of Advanced Studies (IIAS), Kyoto (Japan), April 23-27th, 2007.

Tagora

- C. Cattuto, *Emergent semantics in online communities*, Language Engineering Group, Chinese University of Hong Kong, Hong Kong, May 4th, 2007.

- V.D.P. Servedio, DPG Conference, Section Dynamics and Statistical Physics, *Ranking and Community detection in unweighted networks*, Spring Meeting of the German Physical Society (DPG), Regensburg (Germany), March 26-30th, 2007.

- V. Loreto, DPG Conference, Working Group Physics of Socio-Economic Systems, *Non-equilibrium phase transitions in negotiation dynamics*, Spring Meeting of the German Physical Society (DPG), Regensburg (Germany), March 26-30th, 2007.

- V. Loreto, *Simon's like models with memory for tagging behaviour*, Conference on New Directions in Complex Systems, Buyukada, Istanbul (Turkey), September 3-9, 2006.

- V. Loreto, *Stochastic models for tagging behaviour*, Language simulations Conference, Warsaw (Poland), September 11-14, 2006.

- V. Loreto, *Cultural route tothe emrgence of linguistic agreement*, EPFL, Lausanne, November 13, 2006.

- V. Loreto, *Complex systems approach to Language Games*, International Conference on Complex systems: from Physics to Biology and the social sciences, Lisbona (Portugal), November 22-25, 2006.

- V. Loreto, *Characterization of Complexity*, V Seminaire Transalpin de Physique "Dynamics and statistics in complex systems", Villa Gualino, Torino - Italy, February 26 - March 2, 2007.

- V. Loreto, *Language Games: a physicist point of view*, Seminar on Language, Evolution, and the Brain (SLEB), International Institute of Advanced Studies (IIAS), Kyoto (Japan), April 23-27th, 2007.

- V. Loreto, *Cultural route to the emergence of linguistic categories*, Language Engineering Group, Chinese University of Hong Kong, Hong Kong, May 4th, 2007.

- Szomszor, M., Cattuto, C., Alani,A., O'Hara, K., Baldassarri, A., Loreto,V., Servedio, V. D. P., *Folksonomies, the Semantic Web, and Movie Recommendation*, "Bridging the Gap between Semantic Web and Web 2.0" at the 4th European Semantic Web Conference ESWC 2007, Innsbruck, Austria, June 7, 2007.

- Schmitz, C.: *Network Properties of Folksonomies*, Challenge on Collaborative Knowledge Creation at the WWW2007 International World Wide Web Conference, Banff, Canada, May, 8–12, 2007.

- Hotho, A.: *BibSonomy - A social publication reference management system*. Challenge on Collaborative Knowledge Creation at the WWW2007 International World Wide Web Conference, Banff, Canada, May, 8–12, 2007.

- A. Hotho: BibSonomy - Research around a Folksonomy System, Informatik–Kolloquium Magdeburg, 31.5.2007

- Hanappe P., *Introduction to Collaborative Tagging*, Università IUAV di Venezia, Italy, June 2006.

- Hanappe P., *Collaborative Tagging and Semiotic Dynamics*, presentation at the Hochschule fur Gestaltung, Karlsruhe, Germany, November 2006.

2.7   Tutorials

UniK provides four online tutorials for BibSonomy. They explain what the links around a BibSonomy entry mean, how to integrate BibSonomy into ones browser, how to add keywords to ones bookmarks, and how to provide metadata for your publication. The tutorials are available online at `http://www.bibsonomy.org/help/tutorials`.

2.8   Conferences, Workshops and Summer Schools

Completed

- Gerd Stumme and Christoph Schmitz were co-organizers of a workshop on Semantic Network Analysis that was co-colocated with the European Semantic Web Conference ESWC-2006, June 12, 2006 at Budva, Montenegro.
  (`http://www.kde.cs.uni-kassel.de/ws/sna2006/`)

- Andreas Hotho co-chaired the "Workshop on Web Mining 2006 (WebMine)" that was co-located with the European Conference on Machine Learning/International Conference on Principles of Knowledge Discovery in Databases 2006, September 18th, 2006, Berlin, Germany.
  (`http://www.kde.cs.uni-kassel.de/ws/webmine2006/`)

- Steffen Staab co-organized the 15th International Conference on Knowledge Engineering and Knowledge Management, EKAW 2006, Podebrady, Czech Republic, October 2-6 2006.
  (`http://ekaw.vse.cz/`)

- Vittorio Loreto (PHYS-SAPIENZA) organized the II Bagnovignoni meeting on *Semiotic Dynamics: Grammar*, Bagnovignoni (Siena), Italy, 23-26 October 2006.

- Steffen Staab co-organized the 1st International Conference on Semantic and Digital Media Technologies, Athens, Greece, 6-8 December 2006.
  (`http://www.samt2006.org/`)

- Harith Alani and Gerd Stumme organized a workshop on Social and Collaborative Construction of Structured Knowledge at the World Wide Web Conference 2007.
  (`http://km.aifb.uni-karlsruhe.de/ws/ckc2007`).

- Andreas Hotho co-chaired the Workshop "Bridging the Gap between Semantic Web and Web 2.0" that was co-located with the 4th European Semantic Web Conference ESWC 2007, June 7, 2007, Innsbruck, Austria.
  (`http://www.kde.cs.uni-kassel.de/ws/eswc2007/`)

Planned

- Vittorio Loreto (PHYS-SAPIENZA) is the vice-chairman of the XXIII IUPAP International Conference on Statistucal Physics, STATPHYS 23, Geona, Italy, 9-13 July 2007.
  (`http://www.statphys23.org`).

Tag☉ra

- Vittorio Loreto (PHYS-SAPIENZA) and Luc Steels (SONY-CSL) are organizing the *International School on Complexity: Course on Statistical Physics of Social Dynamics: Opinions, Semiotic Dynamics, and Language*, jointly with Ettore Majorana Foundation and Center For Scientific Culture, Erice, Italy 13-20 July, 2007.
  (`http://pil.phys.uniroma1.it/erice2007`).

- Gerd Stumme is co-organizing the "Web Mining 2.0 Workshop" that will be co-located with the 18th European Conference on Machine Learning / 11th European Conference on Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD) 2007, September 21, 2007 - Warsaw, Poland.
  (`http://www.kde.cs.uni-kassel.de/ws/ecmlpkdd2007/`)

- Ciro Cattuto (PHYS-SAPIENZA) is organizing a satellite workshop of the European Conference on Complex Systems 2007 (ECCS 2007) on *Social websites: complex dynamics and structure*, Dresden, Germany, 5th October 2007.

- Andreas Hotho is co-organizing the "International Workshop on Data Mining in Web 2.0 Environments" held in conjunction with the IEEE International Conference on Data Mining (ICDM 2007) on October 28, 2007 in Omaha, United States.
  (`http://www.kde.cs.uni-kassel.de/ws/Web2DM`)

- Steffen Staab is co-organizing the 6th International Conference on Ontologies, DataBases, and Applications of Semantics (ODBase), Vilamoura, Portugal, November 25-30, 2007.
  (`http://www.cs.rmit.edu.au/fedconf/index.html?page=odbase2007cfp`)

- Steffen Staab is the general chair of the International Conference on Intelligent User Interfaces, Canary Islands, Spain, January 13-16, 2008.
  (`http://www.iuiconf.org/`)

- The TAGora project is organizing a Dagstuhl Seminar on *Social web communities* (seminar n.08171), Dagstuhl, Germany, 20-25 April 2008.

2.9 Press Radio & TV.

2.9.1 Press

- Swaminathan, N., *Tag, You're It: Scientists Describe Collaborative Tagging Sites like Del.icio.us*, ScientificAmerican.com, 23 January 2007
  (`http://www.sciam.com/article.cfm?articleid=5021D304-E7F2-99DF-33DDD86F3B3ECA20`).

- Bohannon, J., *Folk Wisdom for Web Sites*, ScienceNOW, 23 January 2007
  (`http://sciencenow.sciencemag.org/cgi/content/full/2007/123/5`).

- Progetto italiano studia nuovo linguaggio rete: Un gruppo INFM-CNR studia lo sviluppo del fenomeno "tagging", che genera nuove modalità di categorizzazioni, *Corriere della Sera*, 19 July 2006,
  (`http://www.corriere.it/Primo_Piano/Scienze_e_Tecnologie/2006/07_Luglio/19/web2.shtml`).

- Al via un progetto di ricerca per studiare la dinamica della informazione sul web, *BUR.it - Bollettino Universit & Ricerca*, 18 July 2006,
  (`http://www.bur.it/2006/N_G_061926.php`).

- Al via un progetto di ricerca per studiare la dinamica dell'informazione sul web, *Sussidiario*,
  (`http://www.sussidiario.it/notizie/ultimora/messages/12581.shtml`).

- Comportamento online, sotto c'è la matematica, *Corriere dell'Università e del Lavoro*, 5-18 February 2007, p. 6.

- Dopo i bloggers arrivano i "taggers": il nuovo fenomeno del web tra i nuovi punti del progetto "Tagora" del CNR, *ITnews*, 31 July 2006, (http://www.itnews.it/news/2006,182,dx,0731,135847.html).

- Dopo i bloggers arrivano i "taggers": il nuovo fenomeno del web tra i nuovi punti del progetto "Tagora" del CNR, *pc-facile.com*, 2 August 2006, (http://www.pc-facile.com/news/dopo_bloggers_arrivano_taggers/).

- Dotta, G., Tagora, studio italiano sulla folksonomy, *HTML.it* ,19 July 2006, (http://webnews.html.it/news/leggi/4468/tagora-studio-italiano-sulla-folksonomy/).

- Frittella L., Verso il web del futuro: Lo studio "Tagora" è rivolto alle dinamiche dell'informazione in Internet, *Rai.it*, 31 July 2006, (http://www.rai.it/news/articolonews/0,9217,1067143,00.html).

- Gallavotti B., *Come si spia il futuro di Internet: un progetto europeo indaga la folksonomia, il fenomeno di organizzazione spontanea dei contenuti. Fisici e neuroscienziati impegnati in un'impresa interdisciplinare, ispirata dalla logica dei sistemi complessi*, La Stampa, 28 March 2007 (http://www.tagora-project.eu/wp-content/2007/05/bohannon_folksonomy_sciencenow.pdf).

- Gruppo studia linguaggio e dinamiche della rete, *Broadcast&Video*, n. 373, 21 July 2006, p. 12.

- Guidolanda F., Tagora, lo studio del Web 2.0, *InternetNews*, 24 July 2006, (http://internetnews.blog.dada.net/permalink/270736).

- Manacorda E., Internet: Dimmi come cataloghi... *Galileo - Giornale di scienza e problemi globali*, 4 August 2006, (http://www.galileonet.it/Magazine/mag0630/0630_2.html).

- Molinari C., Il web del futuro, *Il web del futuro - Cached Blog*, 1 August 2006, (http://www.cached.it/blog/archives/248-Il-web-del-futuro.html).

- Progetto "TAGora" sulle nuove dinamiche del web, *Città della Scienza*, 3 August 2006, (http://www.cittadellascienza.it/news_dal_mondo/index.cfm?sw=1&id_cat=6&ID_LANCI=5078).

- Progetto"Tagora": Semiotic Dynamics in on-line Social Communities?, *Ricerca Italiana*, 25 July 2006, (http://www.ricercaitaliana.it/news/news_ente-redazione_ri.htm(linkaPDFINFM)).

- Sante J. Achille, Europe studies the Web 2.0 Phenomenon with TAGora, *Multilingual Search*, 20 July 2006, (http://www.multilingual-search.com/527/20/07/2006/en/).

- TAGora: ecco dove ci porterà il web, *La Stampa Web*, 31 July 2006, (http://www.stampa.cnr.it/documenti/agenzie/2006/Lug/31_lug_06_02.txt(adkronos)).

- TAGora: esplorare le dinamiche del web, *ADV News 24h*, 3 August 2006, (http://www.advmagazine.net/advnews-24h/index.php?p=11015048441707).

- TAGora: esplorare le dinamiche del web, *Amici Amici Magazine*, n. 1518, 3 August 2006.

Tagora

- TAGora: esplorare le dinamiche del web, , *Jugo*, 31 July 2006,
  (`http://www.jugo.it/modules.phpname=News&file=article&sid=4118`).

- TAGora: un progetto italiano verso il web semantico, *Googlisti.com*, 26 July 2006,
  (`http://googlisti.com/tag/tagora/`).

- Tagora, studio italiano sulla folksonomy, *HTML.it* ,
  (`http://webnews.html.it/news/stampa/4468/tagora-studio-italiano-sulla-folksonomy/`).

2.9.2 Talks at non-technical events, Radio

- Vittorio Loreto, Il volo delle Oche, Radio24, 1st July 2006.
  (`http://ilsole24ore.com/radio24_audio/oche010706.rm`).

- C. Cattuto, panel on *Collaborative collection making* at the Sony Intensive Science event, La Maison Rouge, Paris, October 6-10th, 2006.

- V. Loreto , *La nascita del consenso nei sistemi complessi: l'esempio del linguaggio*, round table on "La Nascita del Consenso" at the Roma Opera Theatre, Roma, May 31st, 2007.

- C. Cattuto, *Il web come sistema complesso: la gestione collaborativa della conoscenza* , round table on "La Nascita del Consenso" at the Roma Opera Theatre, Roma, May 31st, 2007.

- C. Cattuto, *La complessità nel cyberspazio*, "Apriamo la mente. Lazio, terra di scienza" science festival, Poggio Mirteto (Italy), May 22nd, 2007.
  (`www.culturalazio.it/binary/lazioCultura/agenda/programmi_Mela.1178712093.pdf`)

- C. Cattuto, *La complessità nel cyberspazio*, "Apriamo la mente. Lazio, terra di scienza" science festival, Viterbo (Italy), May 26th, 2007.
  (`www.culturalazio.it/binary/lazioCultura/agenda/programmi_Mela.1178712093.pdf`)

- G. Stumme, C. Schmitz: Jagd auf die Hommingberger Gepardenforelle Suchmaschinen und Soziale Software. Presentation for B Braun AG, Kassel, 19.9. 2006

BibSonomy has been announced on several mailing lists to attract more users. The list of mailing lists includes:

— dbworld

— kdnet-members@iais.fraunhofer.de

— wi@aifb.uni-karlsruhe.de

— ak-kd-list@aifb.uni-karlsruhe.de

— fgml@cs.uni-kassel.de

— fg-db@informatik.uni-rostock.de

— fca-list@aifb.uni-karlsruhe.de

— orgmem@aifb.uni-karlsruhe.de

— dl@dl.kr.org

— kaw@science.uva.nl

— community@mlnet.org

— web_graph_algs@yahoogroups.com

— webir@yahoogroups.com

— ontoweb-list@lists.deri.org

— semanticweb@yahoogroups.com

— seweb-list@lists.deri.org

— cg@conceptualgraphs.org

— kweb-all@lists.deri.org

— all-prolearn@agws.dit.upm.es

— ml@isle.org

— AI-SGES@JISCMAIL.AC.UK

— machine-learning@yahoogroups.com

— mlearn@googlegroups.com

— Web-Mining@googlegroups.com

— Machine-Learning@googlegroups.com

— Data-Mining@googlegroups.com

— INDUCTIVE@LISTSERV.UNB.CA

2.10 TAGora Project Website

As part of the project presentation task, a web site has been set up as a single reference point for all the public activities of the TAGora project. The second-level domain `tagora-project.eu` was registered and reserved for the TAGora project. A dedicated Linux server was purchased and is currently hosted in the computer room of the PHYS-SAPIENZA team. The server is running the open-source Apache (`http://www.apache.org`) web server and related services. The TAGora web site, `http://www.tagora-project.eu` (Fig. 2.1) is managed by using Wordpress (`http://wordpress.org`), the popular open-source content management system (CMS). The Wordpress installation was highly customized to fit the project needs, both in terms of graphical layout and in terms of content structure. The front page of the web site provides an outline of the TAGora mission and a list of the participating institutes. The main navigation bar (see Fig. 2.1, top) provides access to the following sections:

- **consortium**, `http://www.tagora-project.eu/consortium`: this section gives an overview of the TAGora consortium, highligthing the complementarity of the partners in relation to the project mission. A list of the TAGora teams follows, providing details about the areas of expertise of each team and its position within the international research community.

- **people**, `http://www.tagora-project.eu/people`: this section provides short bios of all the persons involved in the TAGora project.

Tagora

Figure 2.1: `http://www.tagora-project.eu`

- **research**, `http://www.tagora-project.eu/research`: gives a more technical overview of the scientific and technological objectives of the TAGora research agenda, focusing on emergent metadata, data analysis, modeling and simulation.

- **products**, `http://www.tagora-project.eu/products`: this section showcases the applications developed in the framework of the project (currently the Ikoru image-tagging application by Sony CSL and the BibSonomy social bookmarking system developed at the University of Kassel).

- **contact**, `http://www.tagora-project.eu/contact`: provides contact information for the project. This includes the email address for project inquiries, `info@tagora-project.eu`, the address of the project coordinator, and information access to the TAGora mailing list for announcements and project news, available at `http://lists.tagora-project.eu/listinfo/news`.

In addition to the above section, three more dynamic website sections are available, whose content is constantly updated to reflect the progress of the project:

- **blog**, `http://www.tagora-project.eu/blog`: this section is a full-blown blog for the TAGora project, co-authored by several project members. The blog is used to announce all kinds of events and news involving the TAGora project, covering scientific results, participation to conferences and workshops, project-sponsored events, press releases, articles in the press and so on. RSS and Atom feeds are provided to allow better exchanges with the blogging community.

- **outreach**, `http://www.tagora-project.eu/outreach`: this section is a subset of the TAGora blog, focusing on blog entries that deal with press releases and articles in the press.

- **publications**, `http://www.tagora-project.eu/publications`: provides a live list of papers published by members of the TAGora project. For each article, the full bibliographic information is given, together with an PDF version of the paper. The publication list is constantly updated by polling BibSonomy: project members post their new articles to BibSonomy, under the grop "TAGora", and tag them as *tagorapub*. The TAGora website polls the BibSonomy RSS feed for the TAGora group and when it detects new articles tagged with *tagorapub*, it loads them into Wordpress and published them. This kind of integration is an original design deployed by the TAGora project.

All the content included in the above three sections (blog entries, press entries, publication entries) is tagged, and a live tag-cloud is provided in the sidebar of the website (Fig. 2.1, right). Blog posts have editorially given tags, while the tags of publication entries are automatically pulled from BibSonomy. Tags can be effectively used to browse the whole dynamic content of the TAGora website. Standard text search of the website content is also available.

Updated access statistics to the TAGora website are available at

`http://www.tagora-project.eu/stats` (password-protected).

As of April 2007, the TAGora websites is visited by about 1,100 unique users per month.

2.11   Demonstrations in Industrial and Scientific Exhibition

SONY-CSL met with the following people and corporate divisions to discuss collaborative tagging and demonstrate the Ikoru system (see Task 2.2):

- Steve Amagai (Sony Corporation, Tokyo)

- Gerald Reitmayr (Senior Manager E-Business at Digital Imaging Sony Europe)

- Mr Ozawa (Product Planing Dept. VBD, ITCNC, Sony Corporation)

- Akikazu Takeuchi of So-Net (Sony Communication Network Corporation, Tokyo)

- Masanao Tsutsui (Sony CyberShot, Digital Imaging Business Group, Tokyo)

- Dr Sato (Mobile Product R&D Group, Sony Corporation, Tokyo)

- Dr Toshino (PAO Group, Sony Corporation, Tokyo)

Planned actions

- We plan to present a position paper about TAGora at the Complexity Research Projects (CRP) Forum in Dresden (Germany), October 6th, 2007.

2.12   Other dissemination actions

- Aurnhammer M., Hanappe P., Linke A., Brunello M., Graziani S., and the students of the Faculty of Arts and Design, Università IUAV di Venezia, ClaVES undergraduate degree in Visual Arts and Theatre, *Linking Linke*, La Maison Rouge, Paris, Ocrober 2006.

Tagora

- Aurnhammer M., Hanappe P., Linke A., Brunello M., Graziani S., and the students of the Faculty of Arts and Design, Università IUAV di Venezia, ClaVES undergraduate degree in Visual Arts and Theatre, *Linking Linke*, Sony ExploraScience Museum, Tokyo, December 2006 - January 2007.

- Sony CSL, *Intensive Science*, brochure for the 10th anniversary of the Sony Computer Science Laboratory, Paris, October 2006.

- Linke A., Hanappe P., Tisselli E., Rich A., Kuhn W., Mende D., Pocock P., Korn S., students of the curator class at the HfG Karlsruhe, *Interactive Photo Archive* (working title), Century of the Consumer Exhibition, Zentrum fur Kunst und Medienkunst, planned for October 2007.

- Linke A., Hanappe P., Tisselli E., Rich A., Kuhn W., Mende D., Pocock P., Korn S., students of the curator class at the HfG Karlsruhe, *Interactive Photo Archive* (working title), Institute for Contemporairy Art and Tought, Athens, Greece, planned for October 2007.

- Debatty R., *Net_Dérive, the city as instrument*, We make money not art, 19 October 2006, `http://www.we-make-money-not-art.com/archives/009042.php`

- Rheingold H. *Street Ethnographers*, Smart Mobs, 22 May 2007, `http://www.smartmobs.com/archive/2007/05/22/street_ethnogra....html`

- Green J.-A., *canal*MOTOBOY: Empowering people to tell their stories*, Networked_performance, 11 May 2007, `http://www.turbulence.org/blog/archives/004266.html`

- Wellington C., *Motoboys transmitem de celulares*, Overmundo, 10 May 2007, `http://www.overmundo.com.br/agenda/motoboys-transmitem-de-celulares`

- Tisselli E., (SONY-CSL) *Tagging dynamics at Zexe.net*, report and analysis about the online collaborative tagging project, zexe.net (`http://zexe.net`). Document available by fall, 2007.

- Aurnhammer M., Hanappe P., *Linking Linke and Collaborative Tagging*, panel discussion, Intensive Science Exhibition, La Maison Rouge, Paris, October 2006.

- Steels L., Pachet F., Hanappe P. *Collective Intelligence and Sustainable Development*, Solutions for a Sustainable Worls Workshop, Camargue, France, May 2007.

2.13 Dissemination actions toward the other projects of the Complex System Initiative

- Vittorio Loreto (PHYS-SAPIENZA) is the vice-chairman of the XXIII IUPAP International Conference on Statistical Physics, STATPHYS 23, Geona, Italy, 9-13 July 2007. (`http://www.statphys23.org`).

- Vittorio Loreto (PHYS-SAPIENZA) and Luc Steels (SONY-CSL) are organizing the *International School on Complexity: Course on Statistical Physics of Social Dynamics: Opinions, Semiotic Dynamics, and Language*, jointly with Ettore Majorana Foundation and Center For Scientific Culture, Erice, Italy 13-20 July, 2007. (`http://pil.phys.uniroma1.it/erice2007`).

- Ciro Cattuto (PHYS-SAPIENZA) is organizing a satellite workshop of the European Conference on Complex Systems 2007 (ECCS 2007) on *Social websites: complex dynamics and structure*, Dresden, Germany, 5th October 2007.

2.14   Joint Publications

- Cattuto, C., Schmitz, C., Baldassarri, A., Servedio, V. D. P., Loreto, V., Hotho, A., Grahl, M., Stumme, G., *Network Properties of Folksonomies*, AICOM Special Issue on *Network Analysis in Natural Sciences and Engineering*, 2007.

- Schmitz, C., Grahl, M., Hotho, A., Stumme, G., Cattuto, C., Baldassarri, A., Loreto, Servedio, V. D. P., *Network Properties of Folksonomies*, Proceedings of the WWW2007 International World Wide Web Conference, 2007.

- Szomszor, M., Cattuto, C., Alani, H., O'Hara, K., Baldassarri, A., Loreto, V. and Servedio, V. D. P., *Folksonomies, the Semantic Web, and Movie Recommendation*. In Proceedings of 4th European Semantic Web Conference, Bridging the Gap between Semantic Web and Web 2.0 (2007) (in press), Innsbruck, Austria.