Towards Research Collaboration – a Taxonomy of Social Research Network Sites

Angelika C. Bullinger

University of Erlangen-Nuremberg angelika.bullinger@wiso.uni-erlangen.de

Uta Renken

University of Erlangen-Nuremberg uta.renken@wiso.uni-erlangen.de

Stefan H. Hallerstede

University of Erlangen-Nuremberg stefan.hallerstede@wiso.uni-erlangen.de

Jens-Hendrik Soeldner

University of Erlangen-Nuremberg jens.soeldner@wiso.uni-erlangen.de

Kathrin M. Moeslein

University of Erlangen-Nuremberg kathrin.moeslein@wiso.uni-erlangen.de

ABSTRACT

The increase of scientific collaboration coincides with the technological and social advancement of social software applications which can change the way we research. Among social software, social network sites have recently gained immense popularity in a hedonic context. This paper focuses on social network sites as an emerging application designed for the specific needs of researchers. To give an overview about these sites we use a data set of 24 case studies and in-depth interviews with the founders of ten social research network sites. The gathered data leads to a first tentative taxonomy and to a definition of SRNS identifying four basic functionalities *identity and network management, communication, information management,* and *collaboration*. The sites in the sample correspond to one of the following four types: *research directory sites, research awareness sites, research management sites* and *research collaboration sites*. These results conclude with implications for providers of social research network sites.

Keywords

Social network sites, taxonomy, research, collaboration

INTRODUCTION

In the last two decades, a considerable increase in scientific collaboration could be witnessed, with temporary collaboration being formed by scientists across departments, institutions, disciplines, and countries. The percentage of scientific publications that were produced by international collaboration increased from 8% in 1988 to 20% in 2005 (Atkins, Droegemeier, Feldman, Garcia-Molina, Klein, Messerschmitt et al., 2003), publications with authors from more than one institution grew from 40% to 61% in the same timeframe. The US National Science Board (Atkins et al., 2003) sums it up: "[...] research is an increasingly collaborative activity", i.e. effective scientific research requires bringing together experts by overcoming institutional, disciplinary, and national. To make these collaboration projects work, openness and speed are required in communicating, sharing, and validating data, findings, and analysis procedures (Atkins et al., 2003).

With the broad availability of Internet-based communication technologies and new generations of web-based collaboration tools, commonly subsumed by the term "social software" (Avram, 2006; Boulos and Wheeler, 2007; Green and Pearson, 2005; Raeth, Urbach, Smolnik and Zimmer, 2009), the requirements of openness and speed can be effectively fulfilled. As part of the recent web 2.0 developments social software describes various applications that enable or illustrate interaction and relationships between Internet users (Koch and Richter, 2007). Beside weblogs and wikis, social software tools include social network sites and a broad range of further tools (Boulos and Wheeler, 2007). Social software tools offer different support opportunities for research work, ranging from support for the individual researcher to solutions that support communication, coordination, and collaboration of teams (Soeldner, Haller, Bullinger and Moeslein, 2009). Thus, research within an established team or with partners beyond the institution can be supported by these social technologies. This phenomenon has

been described as *open research* (Soeldner et al. 2009) and is related to different trends and initiatives like *e-Research*, *cyberinfrastructure* (Atkins et al., 2003; Lawrence, 2006), *OpenData* (Arzberger et al., 2004; Uhlir and Schroeder, 2007), *OpenAccess* (Harnad and Brody, 2004), *OpenScience* or *eScience* (Goble, 2005; Schroeder, 2007). Subsuming these various trends we face a paradigm shift within the research world which might be comparable to the transition from horses to cars in the transportation domain.

The recent advancements of social software, e.g. enabling mass collaboration and the global search for competencies and common interests, hold the potential to significantly influence a paradigm shift. However, there is only little knowledge on the various social software tools which can support research collaboration. Considering the relative novelty of the topic and its potential impact, it seems mandatory for academia to understand this relation. The research presented here starts to build knowledge in the field, led by the following research question:

Are there, and if so, which are the basic functionalities characterizing and distinguishing social research network sites?

We begin to answer the question by presenting the state of knowledge in the field of social software designed for research collaboration. Next, we present the method and the findings of an empirical study we conducted to better understand functionalities and types of social research network sites. A discussion of future pathways ensues. The paper closes with a conclusion, highlighting future research opportunities.

SOCIAL SOFTWARE AND RESEARCH

Social software denotes a class of web-based tools that allow for information exchange, mass interaction, and collaboration (Green and Pearson, 2005; Plotnick, White and Plummer, 2009; Raeth, Urbach, Smolnik and Zimmer, 2009). In the last years, these applications have become pervasive and part of everyday life, both in private and business contexts. Existing frameworks have attempted a categorization along three basic functionalities (Koch and Richter, 2007): identity and network management, information management, and communication. Existing social software applications represent these three basic functionalities to a varying degree. Wikis for example lean more towards the functionality of information management, whereas the focus of instant messaging tools lies on communication. Figure 1 illustrates the three basic functionalities and illustrates a sample of social software applications, e.g. wikis, weblogs, social bookmarking, social tagging services, yellow pages and social network sites.

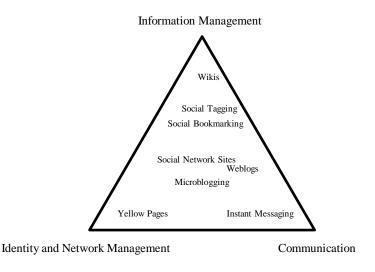


Figure 1. Basic Functionalities of Social Software (cf. Koch and Richter, 2007)

Among the different applications of social software social network sites have continuously grown, well-known examples are facebook.com or myspace.com. Facebook.com attracted 50 million new users between July and September 2009, reaching a total number of 350 million users (Zuckerberg, 2009).

Social network sites (SNS) have been defined with respect to their functionalities: "a web-based service that allows individuals to 1) construct a public or semi-private profile within a bounded system, 2) articulate a list of other users with whom they share a connection, and 3) view and traverse their list of connections and those made by others within the system" (Boyd and Ellison, 2008). Consequently, SNS integrate the basic functionalities listed by Koch and Richter (2007). They allow for *identity and network management*, *information management*, and *communication* with peers. These three

functionalities are needed as effective tools for collaborative research, thus turning social network sites into particularly useful social software tools to support collaboration (Moeslein et al., 2009). Recently, many social network sites have emerged especially addressing researcher and attracting a wide population of users.

To delineate social network sites for researchers from hedonic social network sites (e.g. facebook.com), we adhere to Boyd and Ellison's definition of SNS (2008). However, the utilitarian aspect of collaboration – relevant for researchers – is not yet covered by this definition. We hence add the basic functionality *collaboration* and propose the following definition:

Social research network sites (SRNS) are a web-based service that allows individual researchers to 1) construct a public or semi-public profile within a bounded system (*identity*), 2) articulate a list of other researchers with whom they share a connection and communicate (*communication*), 3) share information with other researchers within the system (*information*) and 4) collaborate with other researchers within the system (*collaboration*).

We will use this literature based definition throughout the paper and underpin its elements by presenting empirical data.

DATA COLLECTION AND ANALYSIS

To define our sample of SRNS in a yet unexplored field, we undertook a pyramiding approach (Hippel, Franke and Pruegl, 2009): Starting off with 24 case studies of social research network sites which had been identified via a keyword-based search on the web (Moeslein et al., 2009), we asked our interviewees to name additional SRNS they would designate as relevant to the market.

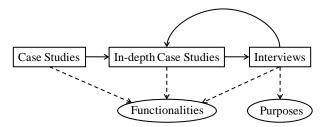


Figure 2. Data Collection and Analysis Process

The data collection and analysis were carried out in four steps (cf. Figure 2): As a first step three experts in the field conducted linear analytical in-depth case studies on the functionalities (Yin, 2009) of the previously identified 24 sites (for a detailed description of this study cf. Moeslein et al., 2009). Continuously, further SRNS had been added until saturation could be reached with ten cases: At this point additional SRNS provided only marginally new information concerning functionalities (Glaser and Strauss, 1967). In step two, the authors used a semi-structured interview guideline to conduct indepth phone or face-to-face interviews with founders or board members of the ten SRNS which had been analyzed in step on. These interviews covered the topics *purpose*, *users* and *usage* of the site as well as *functionalities* and their relative importance in the system. One exemplary question among others was: "What are your site's functionalities to support researchers?" The interviews lasted between 45 and 90 minutes, they were recorded on tape and subsequently transcribed. These transcripts were analyzed by three parties with the help of the qualitative research software ATLAS.ti, following content analysis procedures to code data (Mayring, 2002; Miles and Huberman, 1994; Ryan and Bernhard, 2000). In step three, results from the interviews were merged with the case studies from step one to compare the results using a process of analyst triangulation (Yin, 2009). Finally, if data collected from the various sources were inconsistent or contradictory, we went back to the interviewee to clarify issues (step four). Table 1 summarizes our sample of ten social research network sites¹.

-

¹ Due to confidentiality reasons the statements of the interview partners will be codified in the following. The numbering is random.

Homepage	Description	Interview Partner	# of Members ²
scholarz.net	Management and execution of own or group projects	Daniel Koch, CEO	4.000
collabrx.com	Collaboration on the analysis of shared data	Jeff Shrager, Ph.D., CTO	unknown
ec.europa.eu/ euraxess/	Brokerage of exchange opportunities and information for visiting researchers	Sohail Luka, Ph.D., Policy Officer in the DG Research	1.000
academia.edu	Directory of researchers and of overview of organizations	Richard Price, Ph.D., CEO	120.000
laboratree.org	Online collaboration and management of research teams	Sean Mooney, CEO	7.000
mendeley.com	Online and offline management and recommendation of publications	Victor Henning, CEO	100.000
mynetresearch.com	Document and project management for research groups	Bay Arinze, CEO	12.500
scispace.com	Document management and group communication	Ian Frame, Owner	700
researchgate.net	Networking to maintain awareness of a topic or network	Ijad Madisch, M.D., Ph.D., CEO 250.000	
citeulike.org	Reference collection and recommendation	Kevin Emamy, Owner	275.000

Table 1. Overview of the Sample

FINDINGS

The analysis of the case studies and the interviews leads to three sets of findings on how the social software SRNS can support collaborative research. First, basic functionalities of SRNS are highlighted; second, four typical configurations are presented; and third, we derive strategic implications for the providers of SRNS.

Basic Functionalities of SRNS

Data from the interviews reveal that the developers of social research network sites are convinced that their platforms *differ significantly from existing social network sites* which some interviewees consider to be insufficient for researchers' needs: "Users didn't actually want Facebook, because Facebook was already there and they could have used it, if they wanted to." (interviewee 7). Subsequently, these differences will be presented along the framework of Koch and Richter (2007).

According to the classification scheme for social software, *information management* tools allow for data structuring making wikis a prominent example. This need can also be found within the SRNS and can be illustrated by one SRNS which has explicitly been developed to support *information management*: "It all started due to a personal need when I was doing my PhD. I am doing my PhD in economics and was on the lookout for a tool to manage my data efficiently." (interviewee 1). Beyond the individual scope, SRNS can also support *information management* within a group as explains one interviewee: "Furthermore, we wanted to provide a way for the folks at the university and people at my group to disseminate documents and datasets among each other." (Interviewee 3)

Identity and network management is another main functionality of social software. Facebook.com and other SNS are the most prominent examples for this aspect of social software. They enable self-presentation and facilitate personal contact management, aspects which can also be found on social research network sites according to the explanation of one SRNS developer: "Some sort of facebook.com for researchers, that's what's needed, helping one to quickly find people with specific competencies and qualifications." (interviewee 8). However, SRNS profile details are designed to display the scientist's experience with certain research methods or show the publications of a user. Another issue which can be addressed by identity and network management functionality is gaining an overview about actors in the same research field: "The

² Approximation.

second big thing that happened was when I finished my PhD, I discovered two other doctoral students who worked on the same problem for three years and we had never discovered or heard about each other." (interviewee 2).

Messaging tools prominently represent the functionality *communication* which is another typical social software element. This feature is also present in almost all SRNS under scrutiny: "For example, built into the site, there is some sort of an internal messaging tool, which is basically internal e-mail." (interviewee 7).

In addition to the three basic functionalities of social software (Koch and Richter, 2008), the SRNS in our sample comprehend new tools to support researchers, thus exceeding the offerings made by hedonic social network sites. Multiple research-specific features, like knowledge management tools, citation management, and paper recommendation engines have been implemented and are combined in various ways on the existing platforms. These tools enable researchers to collaboratively structure, coordinate and conduct their work online. One of the interviewees stresses: "We don't like to call it a social network, [...] because the objective is not only socialization. That's not why [researchers] are coming online. The reason they are coming online is to do work." (interviewee 6). Thus, we argue that the classification scheme for social software should be extended by an additional basic functionality, namely *collaboration* as in our proposed definition of SRNS.

A Taxonomy of SRNS

In addition to the basic functionalities, the conducted interviews reveal that the purposes of the sites differ and that this difference seems to have an influence on the provided functionalities. Hence, in order to derive a taxonomy of SRNS, we clustered the sites according to their purpose and functionalities in step three of our data analysis. Four types of SRNS were revealed which are introduced subsequently.

The first type, *research directory sites*, focuses on the identification of researchers according to certain criteria, e.g. his research agenda or special competencies in a field, theory or method. These sites seem to facilitate first contact. As one interviewee puts it:

"The most important thing for these sites is how much information is collected and not just how much, but the quality of the information that is collected for each new member, because that is what enables you to have very sophisticated searches to actually identify the specific skills for the persons you want to work with." (interviewee 6)

To populate the directory, sites of this type allow a researcher to present his or her comprehensive profile to the scientific world. Functionalities supporting *identity management* and *communication* are hence very well established. Academia.edu is a typical representative of this type.

The second type supports researchers to keep informed about news in their network or their field of research. We call these sites *research awareness sites*, as these sites allow researchers to maintain their profile, supply detailed information on their current work and interests, as well as follow other users they are interested in to keep track of their activities. One interview partner says:

"You have a profile, you have updated your profile with new papers and updates and the conferences you are going to, whatever it is and then people who are following you can see your updates and similarly you have your own newsfeed and you can follow more people. You know, you can see what Stephen Hawking is thinking about, for instance. He posted an update on the site yesterday." (interviewee 2)

Tools accounting for the functionalities *identity and network management* as well as *information management* are well developed in this type. On these sites (e.g. ResearchGATE), researchers can create a single point of awareness on topics and people they are interested in.

Sites focusing on the support of a researcher's daily work are coined *research management sites*. They for instance provide tools to collect and manage references or propose references to a researcher. Another possible application is the improved management of research e.g. by tools to structure ideas. Hence, some SRNS "[...] feature range and main benefit is more on the productivity tool side" (interviewee 4).

To sum it up, these SRNS support individuals and teams in carrying out their research by providing supportive tools focusing on *information management* functionalities. A typical representative is Mendeley.

As the fourth type we identify *research collaboration sites*. These focus on the support of (virtual) collaboration by facilitating a joint research process. One interviewee explains:

"We believe that we're focusing on trying to enable the collaboration to happen, not the development of the collaboration itself. [...]. Whether [collaboration] would be at the university, whether that would be at the same department, whether that would be at the other side of the world." (interviewee 3)

To achieve this goal, sites of this type (e.g. Collabrx) focus on tools to support online collaboration functionalities. This requires opportunities to work synchronously on a shared scientific dataset and to develop data analysis scripts collaboratively, which can be executed on the SRNS.

Figure 3 characterizes the four types of SRNS along the four basic functionalities.

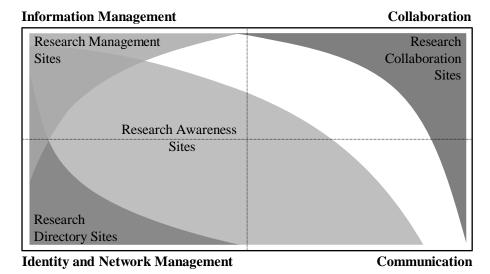


Figure 3. Categorization of SRNS Types According to the Extended Framework

As the figure shows, the four types do overlap in the basic functionalities *identity and network management*, *communication*, *information management*, and *collaboration*. Additionally, sites cannot exclusively be assigned to one type, but rather have a dominant type.

DISCUSSION

Given the four types of social research network sites, data analysis revealed possible future pathways for each type.

Research directory sites are outstanding in the simplicity they offer their users. Entry barriers are hence extremely low. Additionally, research directory sites hold a viable business model by offering details of registered researchers to institutions seeking for academic personnel. These advantages might protect research directory sites from a hostile takeover by research awareness sites. Such takeover can be expected as the two types share a major set of functionalities, while research awareness sites typically hold a larger user base.

Concerning *research awareness sites*, we see a need to improve identity management by a) improve profile information and b) provide powerful search tools. This might enable the companies running *research awareness sites* to take over *research directory sites*. Furthermore, our data indicates potential to improve the integration of research management tools, e.g. reference management and advanced recommender systems. As one interviewee partner puts it:

"Up to now, we have never played the same game. Twelve or fourteen weeks ago we put the reference repository online. Now we move towards their [research directory site] strategic direction." (interviewee 8)

Research management sites capitalize networks effects e.g. by recommending potentially relevant literature on the basis of other user's preferences. To fulfill their potential, they should increase their number of members. The interviewees show an interest to develop the research management sites towards research awareness sites: "We intend to integrate more functionalities for the community." (interviewee 4). However, we see a second possible pathway by a merger with a research awareness site to combine a large user base and the capacity to capitalize network effects.

The most focused platforms are *research collaboration sites* which provide particular collaboration tools for highly specific groups of researchers. This leads to a strong usage pattern by registered members, but limits possibilities to capitalize network effects. Accordingly, our analysis indicates potential advantages if *research collaboration sites* are integrated as subcommunities within larger *research awareness sites* or *research management sites*.

On a more general level, we see business opportunities for each type of SRNS in addressing the privacy needs of specific organizations (e.g. industrial research departments or universities) that do not want to make their research visible, e.g. due to intellectual property concerns. This is in line with the current development that several social research network sites offer sub-communities as isolated silos or in a protected environment with a transparent passage from the protected sub-community to the global, open network.

"We noticed that a lot of institutions implemented Web 2.0 platforms, which were not really being used due to the abundance of different platforms. Thus, we had the idea of creating these sub-communities and to integrate them with the global platform. Data are stored on a different platform, behind the firewalls of the universities, but it's essentially the same system with a separate area. It also brings increased awareness to the university from the global system ... we get a lot of requests for such sub-communities." (interviewee 8)

Strategic consideration of future pathways seems particularly relevant considering the strong increase in new registrations on the sites. This indicates a change from prototypical realizations used by early adopters to a more established application for research collaboration. As one interviewee emphasizes:

"The research productivity gains are too great not to use this technology. So right now, I would say that we are just passing the early adopter stage ... we're now beginning to see a faster rate of adoption, more users in different countries." (interviewee 6)

Table 2 below provide	se a evnoncie an	d a categorization	of the analyz	and SPNS according	or to the ab	ove developed taxonomy.
1 abic 2 below provide	o a symopsis an	id a categorization	i or the anaryz	accordin	ig to the at	ove developed taxonomy.

Homepage	Preliminary SRNS type	Important functionalities		
scholarz.net	Research Management Site Identity and Network Management/ Information Management/ Collabora			
collabrx.com	Research Collaboration Site Collaboration			
ec.europa.eu/euraxess/	Research Directory Site	Identity and Network Management		
academia.edu	Research Directory Site	Identity and Network Management/ Communication		
laboratree.org	Research Collaboration Site	Collaboration		
mendeley.com	Research Management Site/ Research Awareness Site	Information Management/ Identity and Network Management		
mynetresearch.com	Research Management Site	Identity and Network Management/ Information Management/ Collaboration		
scispace.com	Research Collaboration Site	Collaboration		
researchgate.net	Research Awareness Site	Identity and Network Management/ Information Management		
citeulike.org	Research Management Site	Information Management		

Table 2. Categorization of the Sample

CONCLUSION

This paper presents a definition of the evolving social software application *social research network sites* along with the four basic functionalities – *identity and network management, communication, information management,* and *collaboration.* In addition, it suggests a first tentative taxonomy according to the purpose their founders had in mind and the basic functionalities they provide. Implications for providers of SRNS complete the paper. Thus, it contributes to research in the field of collaborative work, focusing on research collaboration.

Results of the paper need to be tempered with its limitations. Our research focuses on data gathering from providers of social research network sites. A comprehensive study analyzing current or potential users and their requirements is necessary. This seems particularly important as no user survey regarding acceptance of functionalities has been carried out by the SRNS in our sample. Furthermore, the presented taxonomy needs to be further examined (e.g. by a further analysis of the distinct types).

However, we see an exciting new emerging field that promises to change the way scientists 1) construct and maintain their public or semi-private profile within a bounded system (*identity and network management*), 2) identify other researchers with whom they share a connection and communicate (*communication*), 3) share information with other researchers (*information management*) and 4) collaborate (*collaboration*).

ACKNOWLEDGMENTS

We would like to thank all researchers who are with us on our journey towards the world of open collaboration in research. This applies particularly to our interviewees for the openness with which they answered our questions. This study would have been impossible without support by the German Federal Ministry of Education and Research (project: BALANCE of Flexibility and Stability in the World of Research, FKZ 01FH09153).

REFERENCES

- 1. Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., et al. (2004). Promoting Access to Public Research Data for Scientific, Economic, and Social Development. *Data Science Journal*, 3, 135-152.
- Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-Molina, H., Klein, M. L., Messerschmitt, D. G., et al. (2003). Revolutionizing Science and Engineering Through Cyberinfrastructure. Retrieved from http://www.nsf.gov/cise/sci/reports/atkins.pdf on January 25th, 2010.
- 3. Avram, G. (2006). At the Crossroads of Knowledge Management and Social Software. *Journal of Knowledge Management*, 4, 1, 1-10.
- 4. Boulos, M. N. and Wheeler, S. (2007). The Emerging Web 2.0 Social Software: an Enabling Suite of Sociable Technologies in Health and Health Care Education. *Health Information and Libraries Journal*, 24, 2-23.
- 5. Boyd, D. M. and Ellison, N. B. (2008). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13, 1, 210-230.
- 6. Glaser, B. G. and Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. Aldine, Chicago.
- 7. Goble, C. (2005). Putting Semantics into e-Science and Grids. In *Proceedings of the First International Conference on e-Science and Grid Computing (e-Science'05)*.
- 8. Green, D. and Pearson, J. (2005). Social Software and Cyber Networks: Ties That Bind or Weak Associations within the Political Organization? In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*. IEEE.
- 9. Harnad, S. and Brody, T. (2004). Comparing the Impact of Open Access (OA) vs. Non-OA articles in the Same Journals. *D-Lib Magazine*, 10, 6, 2-6.
- 10. Hippel, E. V., Franke, N. and Pruegl, R. (2009). Pyramiding: Efficient Search for Rare Subjects. *Research Policy*, 38, 9, 1397-1406.
- 11. Koch, M. and Richter, A. (2007). Enterprise 2.0 Planung, Einfuehrung und erfolgreicher Einsatz von Social Software im Unternehmen. Oldenbourg, Muenchen.
- 12. Lawrence, K. A. (2006). Walking the Tightrope: The Balancing Acts of a Large e-Research Project. *Computer Supported Cooperative Work (CSCW)*, 15, 385-411.
- 13. Mayring, P. (2002). Qualitative Content Analysis Research Instrument or Mode of Interpretation?, in M. Kiegelmann, *The Role of the Researcher in Qualitative Psychology*. Tuebingen, 139-148.
- 14. Miles, M. B. and Huberman, A. M. (1994). Qualitative Data Analysis: an Expanded Sourcebook. Sage, Thousand Oaks.
- 15. Plotnick, L., White, C. and Plummer, M. M. (2009). The Design of an Online Social Network Site for Emergency Management: A One Stop Shop. In *Proceedings of the 15th AMCIS, San Francisco*.
- 16. Raeth, P., Urbach, N., Smolnik, S. and Zimmer, C. (2009). Towards Assessing the Success of Social Software in Corporate Environments. In *Proceedings of the 15th AMCIS, San Francisco*.
- 17. Ryan, G. W. and Bernhard, R. H. (2000). Data management and analysis methods, in N. K. Denzin and Y. S. Lincoln, *Handbook of qualitative research*. Thousand Oaks, Sage, 796-802.

- 18. Schroeder, R. (2007). e-Research Infrastructures and Open Science: Towards a New System of Knowledge Production? *Prometheus*, 25, 1-17.
- 19. Soeldner, J., Haller, J., Bullinger, A. C. and Moeslein, K. M. (2009). Supporting Research Collaboration on the Needs of Virtual Research Teams, in *Proceedings of the 9th WI 2009*, Vienna, 275-284.
- 20. Uhlir, P. F. and Schroeder, P. (2007). Open data for global science. Data Science Journal (Open data issue), 6, 36-53.
- 21. Yin, R. K. (2009). Case Study Research: Design and Methods. Sage, Thousand Oaks.
- 22. Zuckerberg, M. (2009). An Open Letter from Facebook Founder Mark Zuckerberg. Retrieved from http://blog.facebook.com/blog.php?post=190423927130 on January 18th, 2010.