



## First experiences with a Science 2.0 infrastructure

Stefanie Lindstaedt, Peter Kraker, Fridolin Wild, Thomas Ullmann, Erik Duval, Gonzalo Parra

### ► To cite this version:

Stefanie Lindstaedt, Peter Kraker, Fridolin Wild, Thomas Ullmann, Erik Duval, et al.. First experiences with a Science 2.0 infrastructure. 2011. hal-00852079

**HAL Id: hal-00852079**

**<https://telearn.hal.science/hal-00852079>**

Submitted on 19 Aug 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# D6.5

## **First Experiences with a Science 2.0 Infrastructure**

---

Edited by

Stefanie Lindstaedt, Peter Kraker



**STELLAR**

## Amendment History

Version	Date	Contributor(s)	Modification
1	16/09/2010	Peter Kraker	First draft
2	01/11/2010	Stefanie Lindstaedt, Peter Kraker, Fridolin Wild, Thomas Ullmann, Erik Duval Gonzalo Parra	Version for internal review
3	18/11/2010	Stefanie Lindstaedt, Peter Kraker	Final document

*Disclaimer: All information included in this document is subject to change without notice. The Members of the STELLAR Consortium make no warranty of any kind with regard to this document, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The Members of the STELLAR Consortium shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the furnishing, performance, or use of this material.*

WP6 | D6.5

# First Experiences with a Science 2.0 Infrastructure

Stefanie Lindstaedt (KC)

Peter Kraker (KC)

Editor(s)

Fridolin Wild (OUUK)

Thomas Ullmann (OUUK)

Erik Duval (KUL)

Gonzalo Parra (KUL)

Author(s)

public report	final	3	18/11/2010	M20
audience & type	status	version	doc date	due date

science 2.0, mash-up, infrastructure, usage experiences

keywords



# Table of Contents

<b>Amendment History .....</b>	<b>2</b>
<b>Executive Summary .....</b>	<b>6</b>
<b>1 Introduction.....</b>	<b>7</b>
1.1 Goal of Deliverable .....	7
1.2 Scope of Deliverable .....	7
1.3 Structure of Deliverable.....	8
1.4 Related Deliverables .....	8
<b>2 Overview of Science 2.0 Infrastructure .....</b>	<b>10</b>
<b>3 Experiences and Evaluations.....</b>	<b>15</b>
3.1 Researcher Capacity .....	15
3.1.1 Reflection on the collaborative construction of a deliverable with STELLAR wikis .....	15
3.1.2 Inspecting deliverable wikis from logging data over time.....	17
3.1.3 Evaluation of “More!” .....	19
3.2 Doctoral Academy Capacity .....	21
3.2.1 Support for a doctoral community of practice .....	21
3.2.2 Needs Analysis .....	22
3.3 Community Level Capacity .....	23
3.4 Leadership Capacity .....	25
3.4.1 Meeting of the Minds .....	25
3.4.2 Podcasting Infrastructure .....	26
<b>4 Conclusions.....</b>	<b>29</b>
<b>5 References.....</b>	<b>31</b>
5.1 Overview of Papers.....	31
5.2 Additional References.....	35

<b>Annex 1: Papers.....</b>	<b>36</b>
-----------------------------	-----------

# Executive Summary

This deliverable reports on first usage experiences and evaluations of the STELLAR Science 2.0 Infrastructure. Usage experiences were available predominantly for the “mature” part of the infrastructure provided by standard Web 2.0 tools adapted to STELLAR needs. Evaluations are provided for newly developed tools.

We first provide an overview of the whole STELLAR Science 2.0 Infrastructure and the relationships between the building blocks. While the individual building blocks already benefit researchers, the integration between them is the key for a positive usage experience. The publication meta data ecosystem for example provides researchers with an easy to retrieve set of TEL related data. Tools like the ScienceTable, Muse, the STELLAR latest publication widget, and the STELLAR BuRST search show already several scenarios of how to make use of this infrastructure. Especially a strong focus on analytical tools based on publication and social media data seem useful.

In order to highlight the relevance of the infrastructure to the individual capacity building activities within STELLAR, the usage experiences of individual building blocks are then reported with respect to Researcher Capacity (e.g. Deliverable Wikis, More! application), Doctoral Academy Capacity (e.g. DoCoP), Community Level Capacity (e.g. TELeurope), and Leadership Capacity (e.g. Meeting of Minds, Podcast Series). Here we draw from 11 scientific papers published. The reader will find an overview of all these papers in the Appendix. Based on the usage experiences and evaluations we have identified a number of ideas which might be worth considering for future developments. For example, the experiences gained with the Deliverable Wikis show how the modification of the standard Wiki history can provide useful analytical insights into the collaboration of living deliverables and can return the focus on authorship (which is intentionally masked in Wikis, because of their strong notion on the product and not on authors).

We conclude with main findings and an outlook on the development plan and evaluation plan which are currently being developed and which will influence D6.6. Particularly, we close with the notion of a Personal Research Environment (PRE) which draws from the concept of Personal Learning Environments (PLE).

# 1 Introduction

## 1.1 Goal of Deliverable

In this deliverable we report on first usage and evaluation experiences with the STELLAR Science 2.0 infrastructure which was set up within the last 1½ years.

The goal of the deliverable is twofold:

1. We want to provide an overview of how well the STELLAR Science 2.0 infrastructure currently supports the STELLAR instruments, which bottlenecks and which successes could be observed.
2. Based on the identified bottlenecks and success we aim to identify ideas for features which can provide us with pointers for further development and refinement of the STELLAR Science 2.0 infrastructure.

Our examination is based on qualitative as well as quantitative usage data of individual components of the STELLAR Science 2.0 infrastructure as they have been applied to support STELLAR instruments.

Please note: Usage data was predominantly available for existing Web 2.0 applications (e.g. FlashMeeting, Wikis, Open Archive) that we have configured and deployed to support our scientific work within (see lower part of Figure 1). This “mature” part of the STELLAR Science 2.0 infrastructure was available to users from the beginning of the networks and thus has been used extensively. In addition, we also report on a limited number of experiments involving newly developed applications (see upper part of Figure 1) which aim to satisfy a specific STELLAR need or provide the “glue” between standard tools. Since these applications have been developed within the last 1 ½ years (and are still being developed) no real world usage data was available for them.

Based on this deliverable we are currently working on an evaluation plan and a development roadmap. The evaluation plan will examine a variety of possibilities with which we can foster (real world) usage experiences of the newly developed Science 2.0 infrastructure parts. It will also draw from our experience what methods are appropriated to evaluated certain aspects of the infrastructure. This evaluation plan will inform and refine the methods we will choose for Deliverable D6.6. The development roadmap will take the here identified ideas for features into account and aims to streamline the future development activities.

## 1.2 Scope of Deliverable

According to the DoW, this report was to be entitled “Scientific Conference Paper on First Experiences With a Mash-up of Technology Enhanced Services” The deliverable was renamed to “First Experiences with a Science 2.0 Infrastructure” to reflect the extended scope and volume of our work:

- (1) The original scope of “a Mash-up of Technology Enhanced Services” does not reflect the full infrastructure that has been built over the last 1½ years. This infrastructure encompasses not only services (although they play an important part), but also a variety of applications, widgets, and interoperability standards. To do these additional developments justice, the scope was changed to include the whole Science 2.0 infrastructure.
- (2) The deliverable does not consist of a single paper about our experiences with the STELLAR Science 2.0 infrastructure. By now close to 20 papers were published in various conferences and workshops, which address aspects of the STELLAR Science 2.0 infrastructure and their usage. From this pool of papers we selected 11 papers, which have a strong focus on usage experiences and thus best suit this deliverable. This deliverable presents a summary of the most important findings from these papers classified according to different capacity levels of STELLAR instruments. The reader will also find an overview of the original papers in the appendix. In addition, we communicate here new ideas for features which provide pointers for further development of the infrastructure. This deliverable also serves as a draft for a comprehensive infrastructure paper which will be submitted to a leading TEL conference within the next months.

### 1.3 Structure of Deliverable

The deliverable starts out with a short overview of the STELLAR Science 2.0 infrastructure in order to help the reader to put the usage experiences into perspective of the whole. In addition, this overview clearly discusses which parts of the infrastructure are based on standard Web 2.0 tools and which are additional developments.

The next chapters address the different capacity levels of STELLAR Instruments and describe first usage experiences and evaluations for each of them. The usage experiences are drawn from a multi-method approach and constitute the main part of the deliverable:

- Researcher Capacity
- Doctoral Academy Capacity
- Community Level Capacity
- Leadership Capacity

Each of these chapters contains a short description of the tool/application/widget under examination, the usage experience or evaluation results, and the ideas for features which we deduced.

### 1.4 Related Deliverables

In a large integrated project such as STELLAR, individual deliverables cannot stand alone. Instead, this deliverable is closely related to the following deliverables:

- *D6.3 Science 2.0 Mash-Ups (v1)* describes the STELLAR Science 2.0 infrastructure discussed in this deliverable in full (technical) detail.
- *D6.4 Evaluation of the Open Archive* presents a first evaluation of the Open Archive, which is part of the STELLAR Science 2.0 infrastructure.
- *D6.2 Monitoring plan including indicators* reports on a set of quantitative usage measures is identified. Selected usage measures have been employed to evaluate the uptake of the podcasting infrastructure in chapter 3.4.2.
- *D2.1 Podcasting Infrastructure* describes the podcasting infrastructure in detail which we refer to in Chpter 3.4.2.
- *D2.2 Report with the evaluation of the STELLAR Meeting of Minds I and lessons learned for Meeting of Minds II* contains experiences with the usage of Flashmeeting for the Meeting of Minds discussed in Chapter 3.4.1.

## 2 Overview of Science 2.0 Infrastructure

The Science 2.0 Infrastructure consists of two blocks (see Figure 1). The lower block comprises of well-known tools and technologies which are continuously adapted for the needs of the STELLAR NoE. These tools and technologies build the basis for our internal and external work. This deliverable reports predominantly about the usage experiences of these mature Web 2.0 tools which we have gained over the last 1 ½ years.

The upper block comprises of tools and technologies which are being actively developed in the context of STELLAR. They represent missing links between existing tools, and/or operate as analysis and support applications. Since these tools have been developed only recently, no real world usage experiences exist for them. For individual tools we report on available experiment results.

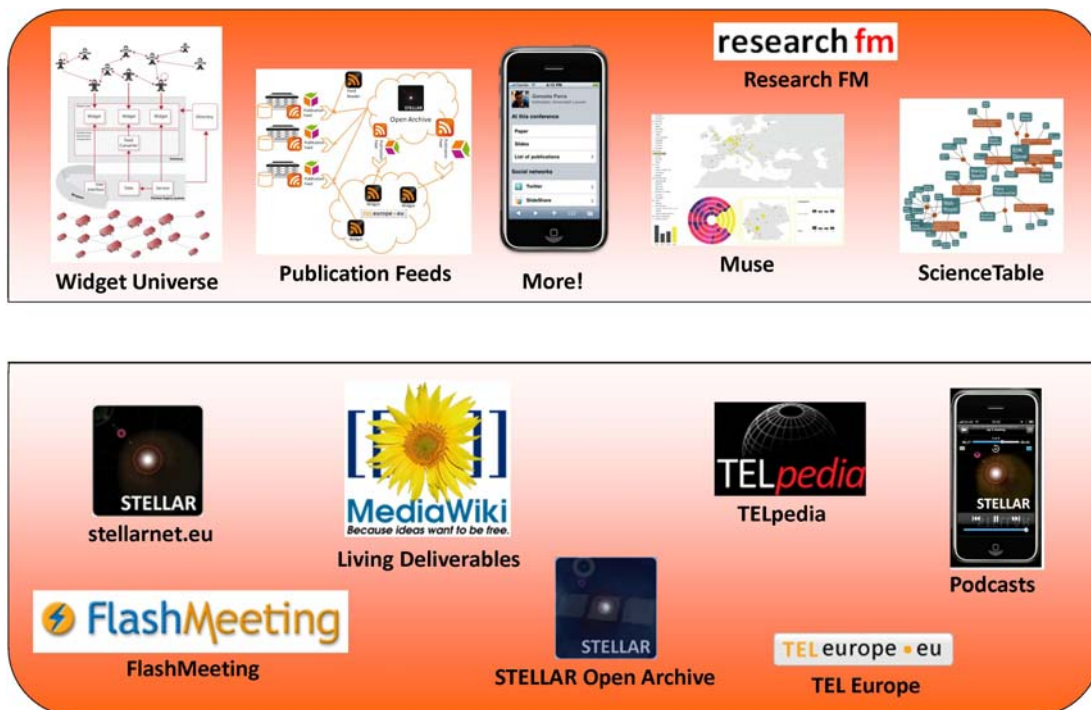


Figure 1: Science 2.0 Infrastructure Blocks

In particular, the tools and technologies in the lower block are:

- The project homepage at <http://stellarnet.eu>, which contains information about the project, as well as a repository of artefacts produced in STELLAR, as well as project management and reporting facilities.
- STELLAR Deliverable Wikis at <http://www.stellarnet.eu/d>, a collection of MediaWikis used to collaboratively produce project deliverables.
- STELLAR Open Archive at <http://oa.stellarnet.eu>, an archive for preprints and references to papers from Technology Enhanced Learning. Its main purpose is to increase the visibility of TEL publications in the international

scientific community. The Open Archive is described in detail in D6.1, first usage experiences are provided in D6.4.

- TELeurope <http://www.teleurope.eu> is home to the stakeholder network in STELLAR. The social networking platform houses part of the widget universe described below.
- Podcasting Infrastructure which allows for uploading of podcasts and access to these podcasts via the website <http://podcast.stellarnet.eu/> and in iTunes <http://itunes.apple.com/podcast/stellar-compilation-podcast/id391960456>. Its purpose is to inform TEL stakeholders with expertises from the community. The Podcasting Infrastructure is described in detail in D2.1.
- Flashmeeting, a collaborative online video-conferencing system, available at <http://fm.ea-tel.eu>.
- TELpedia at <http://telpedia.stellarnet.eu/>, a MediaWiki for articles on and around TEL. Its purpose is to create an encyclopedia for Technology Enhanced Learning.

The upper part of the infrastructure is described in [1]. It comprises of data services on the server side, applications and widgets on the client side, and interoperability formats to facilitate the exchange between server and client entities. The **Publication Feed Ecosystem** builds upon those pillars [2]. It provides the services infrastructure for exchanging and retrieving, and analyzing publication meta-data. The Publication Feed Ecosystem was devised to disseminate publication data within STELLAR and beyond in a format that is commonly readable by existing Web 2.0 infrastructure. The interoperability format is based on BuRST<sup>1</sup> [Mika2005], which represents publication metadata in RSS feeds with an RDF extension part. On the server side, a suite of services was released to aid publishers in creating, aggregating, and filtering of feeds. On the client side, widgets have been developed for visualization of and interaction with publication data.

The widgets mentioned above are part of the **STELLAR Widget Universe** [1], a collection of widgets which visualize information from all parts of the STELLAR net *“bringing together widgets and services and the legacy systems of the STELLAR partners. [...] All widgets are packaged according to the widget 1.0 specification and can thus not only be run within the reference implementation called Universe, but similarly within STELLAR’s stakeholder platform TEL Europe.”*

Another application that builds on top of this infrastructure is **More!** [3]. More! is a mobile application targeted at conferences. It allows conference participants to obtain further information about the speakers (e.g. bio, slides, papers, social networking data) using a QR code for identification. **ScienceTable** [4] and **Muse** [5] are both multi-touch table top application for the collaborative exploration of co-authorship data from publications. Their main aim is to provide researchers with an

---

<sup>1</sup> Bibliography Management using RSS Technology



overview of the field of Technology Enhanced Learning and its sub-topics. ScienceTable focuses on large-scale visualization of relations between papers and authors. Muse, on the other hand, incorporates location data to visualize the relations between institutions and between countries.

In addition to publications, social media has recently received a lot of attention in the community. There is a huge body of social media artefacts (e.g. blog entries, tweets, presentations), representing scholarly communications taking place in addition to traditional publications. *“The **ResearchFM API** was proposed as a RESTful API to provide publication and social data of authors in a unified way.”* [1]

**SOHARC** is a project initiated at the University of Paderborn (D) that aims to collect identity handles of researchers in different social media systems (twitter, facebook, ...). From the web site (<http://soharc.upb.de/about/>): "It aims at acting as hub for sharing social handles between several initiatives that deal with the analysis of social media activity of researchers. [...] With SOHARC we try to add value to research communities and hope that you'll support us in those efforts." In this context, we are developing extensions to the SWRC specification, so as to provide more details about people (re-using vCard and FOAF) and social media artefacts.

Figure 2 gives an overview on how the various components interoperate. Each institution has its own institutional publication repository as depicted in the upper left corner. This data is converted to publication feeds which are in turn used as an input for the **STELLAR Open Archive**<sup>2</sup>. The data aggregated in the Open Archive can be accessed from the semantic web, with publication feeds, and via OAI-PMH<sup>3</sup>. Additionally, raw data from conferences and journals is processed and exposed via BuRST (depicted in the middle from left to center).

Social media data is fetched from various sources, including the wikis from the living deliverables, TELpedia, and FlashMeeting. All these data sources are being aggregated in a central repository (denoted “All Pub” in Figure 2). This central repository takes care of disambiguation and matching. The ResearchFM API provides unified access to the data contained in the central repository. Applications, like the ScienceTable and More! (depicted to the right) make use of this data for the purpose of analysis and support. They are also being used to power widgets in the Widget Universe, which in turn can be included in the TEL Europe.

---

<sup>2</sup> Bear in mind that this is not exclusive – indeed anyone can subscribe to the feeds, and develop services and widgets for the ecosystem.

<sup>3</sup> Open Archives Initiative – Protocol for Metadata Harvesting (see <http://www.openarchives.org/pmh/>)

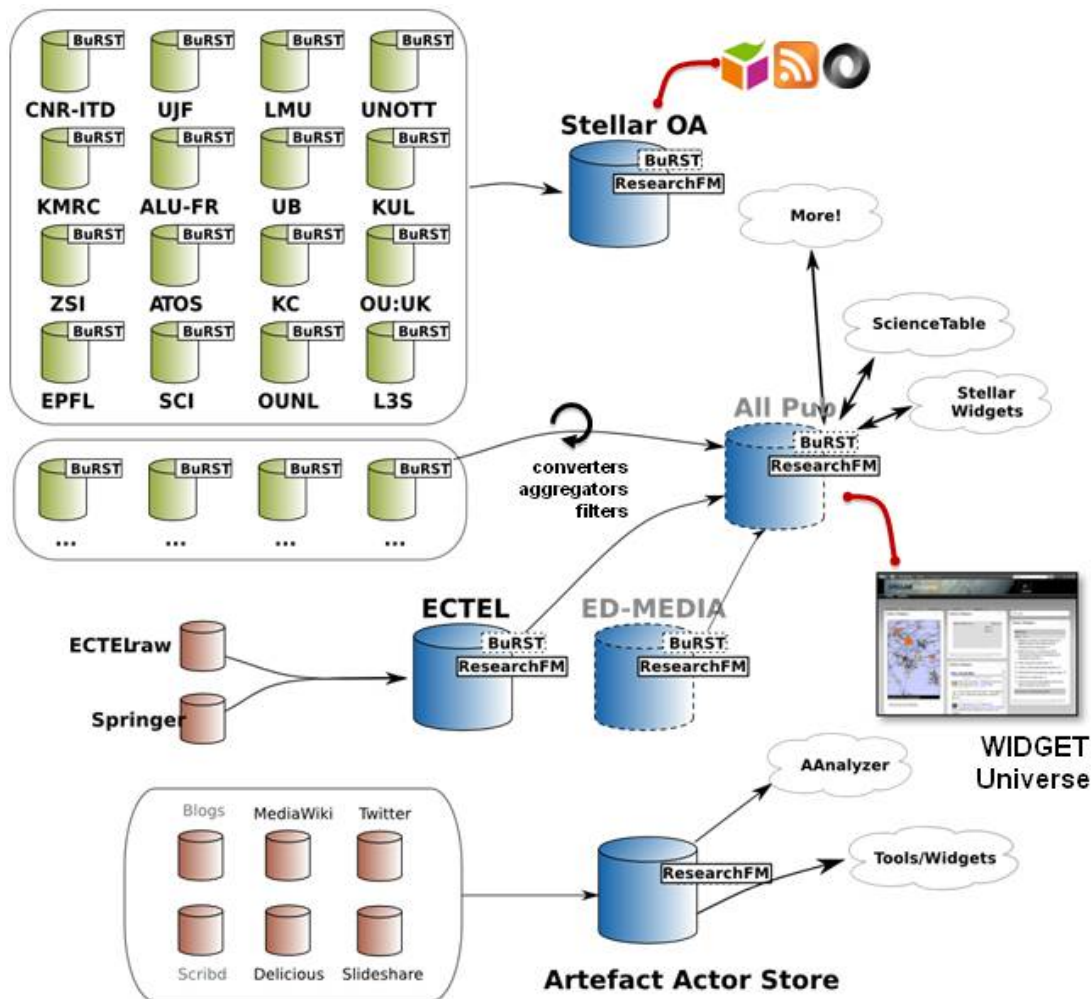


Figure 2: The Components of the TEL Research 2.0 Information Infrastructure.

By now, there are eight publication feeds available; six of them represent institutional feeds and two of them are published by journals and conferences.<sup>4</sup> As a consequence, the Open Archive has recently seen a rise to 3,685 references.<sup>5</sup> There are certain partner feeds that only fulfill the minimum requirement for publication metadata which might impact the usefulness of the archive. One of the challenges will be to improve the quality of the data in the feeds. Nonetheless, the first users and special interest groups are starting to use the widgets provided on TELeurope. To facilitate this trend, new versions of the publication feed widgets are being developed looking to increase the visibility of feed content on the platform.

One of the widgets mentioned above is the Publication Feed Visualization Widget - see Figure 3 for a screenshot. The widget is based on Simile. It displays publication feed items (A) in a faceted browser view, allowing users to filter for source, author,

<sup>4</sup> See [http://www.stellarnet.eu/d/6/3/Directory\\_of\\_BuRST\\_feeds](http://www.stellarnet.eu/d/6/3/Directory_of_BuRST_feeds) for details.

<sup>5</sup> Date: 15/11/2010. See <http://oa.stellarnet.eu/> for details.

and year (B). This is complemented by two further navigational features, namely a timeline (C) and a tag cloud (D – only the onset is shown in the figure). A full text search (E) completes the range of available instruments. [2]

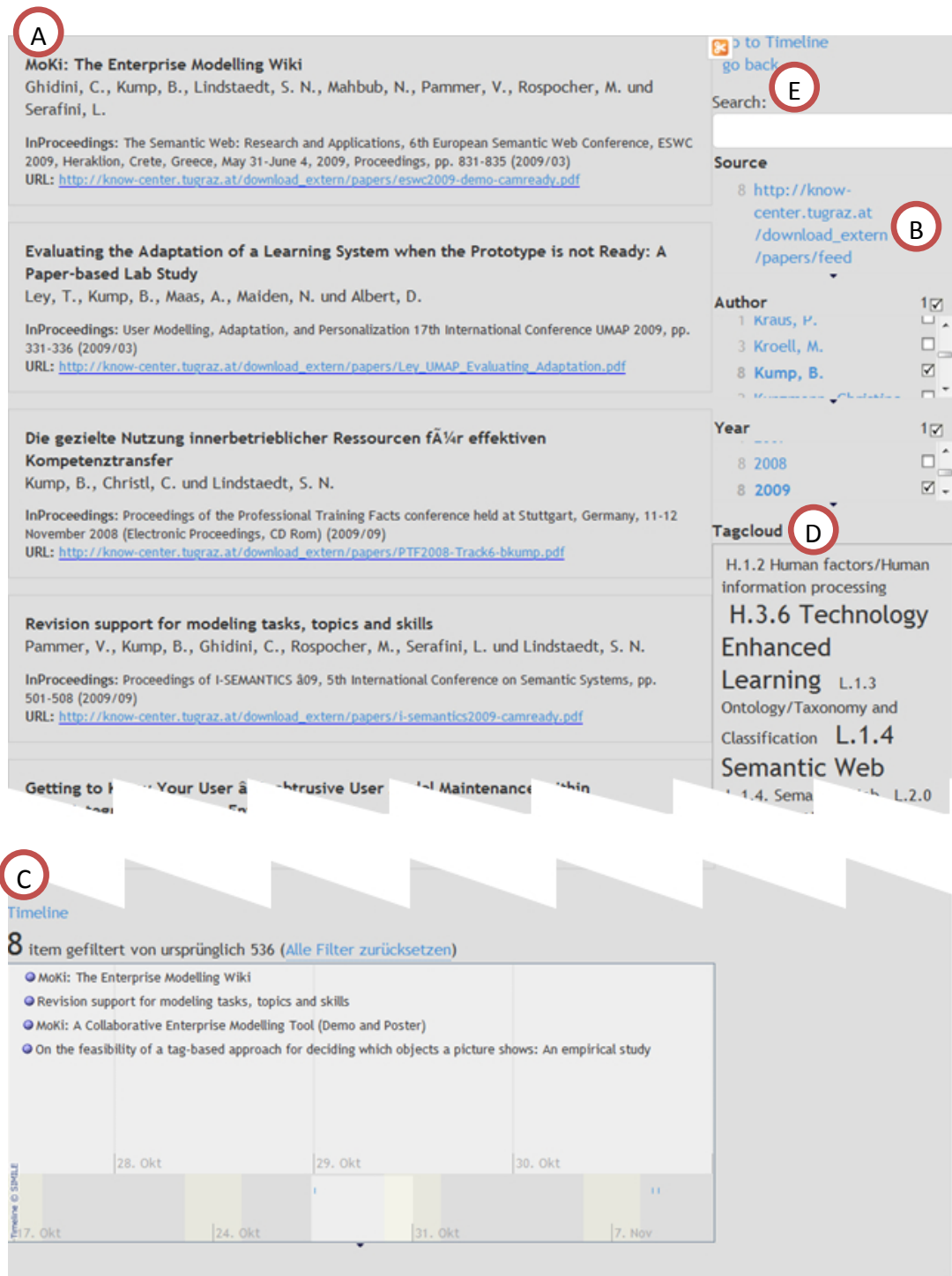


Figure 3: Publication Feed Visualization Widget

## 3 Experiences and Evaluations

This chapter addresses the different capacity levels of STELLAR Instruments and describes first usage experiences and evaluations for each of them: Researcher Capacity, Doctoral Academy Capacity, Community Level Capacity, and Leadership Capacity. Each of the sub-chapters is structured as follows:

- short description of the tool/application/widget under examination
- the usage experience or evaluation results
- ideas for features which we deduced.

### 3.1 Researcher Capacity

Here we start out with our experiences on “Deliverable Wikis” as an internal tool to support the STELLAR NoE for collaborative writing of deliverables and knowledge management and the possibility of turning them into living deliverables. We contrast the two approaches to report about the practices emerging from the Deliverable Wikis in [7] and [8]. We then turn to the evaluation of More! [3], as an example for an external tool that supports researchers.

#### 3.1.1 Reflection on the collaborative construction of a deliverable with STELLAR wikis

STELLAR Deliverable Wikis are a collection of wikis intended for the collaborative production of project deliverables. The Deliverable Wikis are based on Mediawiki and offer usual wiki functionality, such as article history and discussion pages.

Joubert and Sutherland [7] reflect on the collaborative production of deliverable D1.1 with help of the STELLAR Wikis. The goal of this D1.1 deliverable was to develop a research vision and strategy for STELLAR NoE. This turned out to be a particular delicate matter, because of the large amount of partners (16) from nine different countries and the interdisciplinary nature of the endeavor.

The notes from a first face-to-face meeting were used as a starting point for online collaboration. Team members were assigned for each grand challenge to provoke other members of STELLAR to contribute to the wiki. This was supported by virtual meetings conducted with the video-conferencing software Flashmeeting, and a visiting researcher of UJF at the lead university (University of Bristol). The final document was produced by copying the text to a Word document, and editing it in there. A draft document was devised and circulated once more among the STELLAR community, also to account for the author debate *“to check that any contributions they had made had been represented in the way they wanted.”* After 2 internal peer reviews it was submitted.

About 20 people contributed to the wiki, although the authors note that *“sometimes a contribution under one name represented a collation of several contributions from an institution so it could be argued that there were more contributors. The majority of the contributions were made by a small number of people, usually within a short time frame.”*

Although individuals were very willing to contribute statements and clarify them later, *“that contributions were seldom expanded, by either the original authors or other colleagues, with arguments, examples or references.”* Overall, *“[...]the levels of engagement with other people’s contributions was disappointing. There was little evidence of individuals challenging other people’s contributions or questioning what they had said, but typically were more concerned with phrasing and style.”* People were concerned to interfere with other people’s contributions, and some even checked with the original authors. One other concern was that the last edit has the point and everything else is not immediately visible to the viewer. Some suggested to use a word processor instead, because one could instantly see the changes and check back with the original author via the accept/reject feature, even if that meant that the full changes history was lost.

Notwithstanding the problems discussed above, the authors conclude that the wiki usage was a success because they were able to gather the views of the community. However, they found it hard to engage people; problems and suggested solutions are summed up in Table 1.

<b>Problem</b>	<b>Suggested solution</b>
Wiki pages were overloaded with text, which seemed to hinder discussion.	Reduce text that is on one page and include prompts.
The goals of the wiki were not clearly communicated and therefore it was not clear what should be contributed.	Communicate the goals better.
Issues with authorship as outlined above.	Encourage people to use the “Discussion” feature.
Missing trust within the community due to the early stage of the project.	This is a matter of time.

Table 1: Community engagement: problems and suggested solutions

In the end the authors discuss the wisdom of the crowd approach that they took and conclude: *“we could not aggregate all the voices while remaining faithful to the Research 2.0 philosophy underpinning our project. It may be that listening to the multiple voices of the crowds is at odds with forming an aggregation and it may be that we have to re-think how we conceptualise an ‘aggregation’ (particularly an aggregation of visions).”*

Through this experience the following additional ideas for features could be identified:

- Authors need an overview of which parts of the wiki are already of a very good quality and which parts still require additional work. In order to do so we will incorporate a number of information quality and maturity indicators (developed within the EU-funded IP MATURE) into the STELLAR wiki and apply them to the scientific writing process. Examples for such indicators include that a page has not been changed for a long period after extensive editing, and that it is often referred to by other pages.

- Authors need an overview of which parts of the wiki represent the outcomes of discussions and interactions between a multitude of authors and which parts represent the view of only one or a few authors. In order to do so we will incorporate a number of interaction indicators (developed within the EU-funded IP MATURE) into the STELLAR wiki and apply them to the scientific writing process. Examples for such indicators include that a page has been the subject of many discussions, and that it was created/refined in a meeting.
- The scientific discussions need to be closely linked to the text which is discussed. We will examine how this link can be best established. For this we might utilize the interaction indicators (above) as a trigger point.

### 3.1.2 Inspecting deliverable wikis from logging data over time

Wild and Ullmann [8] evaluate another aspect of the STELLAR Wikis. They inspect logging data from the wikis to determine, if the deliverables collaboratively produced in the wikis qualify as “living deliverables”. According to the authors, *“[t]he assumption behind living documents is that knowledge construction processes are continuous and deliverables are artefacts of an underlying, continuous collaboration process.”*

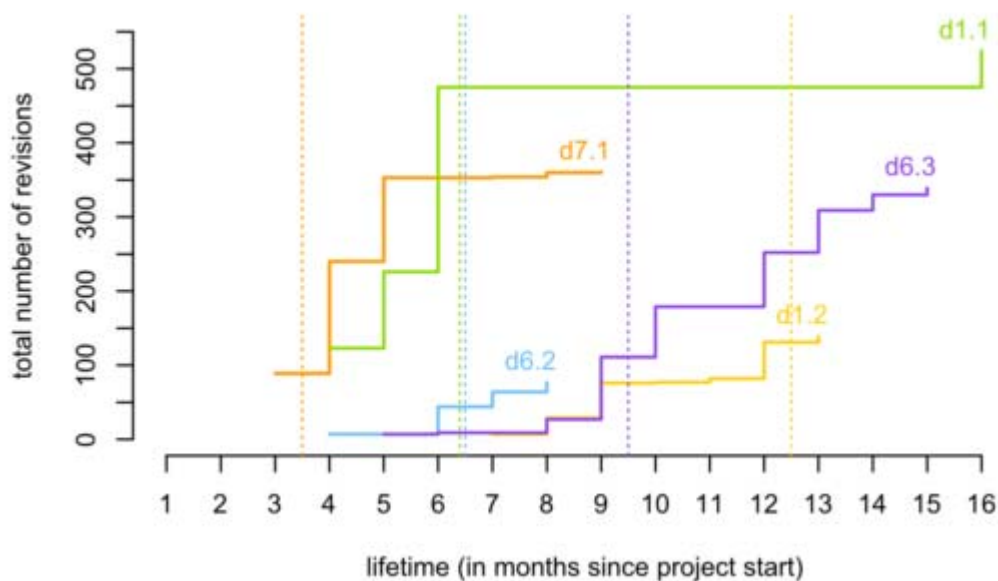


Figure 4: Total number of edits (cumulated) for each deliverable over time

Therefore, the authors inspect the history of each of the wikis. Figure 4 shows the cumulated revisions for five deliverables over the first 18 months of the project. The dotted lines represent the deadlines. *“The three deliverables d6.2 (blue), d6.3 (purple), and d1.2 (yellow) show a very steady increase over time, whereas*

particularly the early deliverables d7.1 (orange) and d1.1 (green) experience their most busy editing processes around the time of their deadline.”<sup>6</sup>

For deliverable D6.3, the relationship between authors and their contributions to pages is visualized before (see Figure 5) and after the deadline (see Figure 6). These figures show how the focus shifted from structure of the deliverable (home page), and scenarios and use cases to technical topics, like tools, services, and widgets.



Figure 5: Authors (green/dark grey) and their contributions to pages (orange/light grey) in D6.3 before the deadline

<sup>6</sup> For some of the deliverables, editing activity related to the deadline is still after the deadline (but – timewise – close). This is due to late revisions and synchronization activities that update the online version with changes applied to their print version (all deliverables were moved to a word version at a certain point in time, when the final printed document was prepared and polished).



Figure 6: Authors (green/dark grey) and their contributions to pages (orange/light grey) in D6.3 after the deadline

As for the other deliverables: *“D7.1 again exposes a larger network of pages (but with a smaller number of contributors), whereas D1.1 is significantly reduced in the number of contributors (but still showing a larger number of edits). The deliverable D6.2 shows a star pattern of authors editing the main page and D1.2 ceased its activity with its delivery deadline.”*

The authors conclude that there is an afterlife for most of the deliverables. *“At least for the one deliverables we have analysed this in more depth and collaboration beyond the deadline exposes a large co-authorship network, accompanied by shift in focus.”*

Through this experience the following additional ideas for features could be identified:

- Re-introduce a focus on authorship and responsibilities, which are not foreseen for normal Wikis, but are essential for scientific writing.
- Enhancing the capabilities of the standard history view of Wikis to enable collaboration views of chosen timeslots.
- Improve feedback via an enhanced Wiki history view.
- Increase support for export (and import) for generating the print versions.
- Better support for managerial editing processes (e.g. flagging of pages that need more work).

### 3.1.3 Evaluation of “More!”

More! [3] is a mobile application targeted at conferences. It enables participants to retrieve further information about a speaker by scanning an individual QR (Quick Response) code with their mobile phone. This code is displayed e.g. on the



speaker's slides. After successfully scanning the QR code, participants are taken to the speaker's page in the More! web application. There they are presented with information about the speaker, such as institutional details, presentations, and social media tools. This information is provided by a first implementation of the ResearchFM API described in Chapter 2.

More! was evaluated in two different settings: one evaluation took place in a controlled setting with 20 university students of Computer Science at K.U. Leuven; the other evaluation was conducted in the field at the STELLAR Alpine Rendez-Vous 2009 with 15 conference participants. The data in both settings was collected using questionnaires and unstructured interviews and discussions. In addition, More! was also evaluated within the EC-TEL 2010 conference.

Parra and Duval evaluated two aspects of user satisfaction: usability and functionality of the application. Table 2 below shows results from the controlled scenario (Setting 1) and the field (Setting 2). Values range from 1 to 5, with 5 being the highest agreement value.

Table 2: Results from the "More!" evaluation

Aspect	Setting 1 (N=20)		Setting 2 (N=15)	
	Avg.	Std. dev.	Avg.	Std. dev.
Know more of the speaker's work	4,44	0,51	4,25	0,50
Easily follow speaker's presentation	4,06	0,64	3,00	1,15
Easy to use	4,44	0,51	4,75	0,50
Satisfaction with tool design	4,28	0,46	4,00	0,82

The authors conclude that More! is simple and easy to use, and that it actually allows the user to know more about the speaker. With regards to the practices the authors report that they found some unexpected use cases in the evaluation data:

- *"deciding if a presentation is worth attending by quickly checking the presentation slides and paper,*
- *catching up and following the presentation more easily, even if a part of the talk was missed by an attendee,*
- *elaborating more and better questions by having the presentation slides at hand."*

We are currently exploring which additional features would enable us to support those unexpected use cases better. In order to evaluate the impact of More! on the research community even further, another evaluation was conducted at EC-TEL 2010. Three aspects were considered: usefulness, satisfaction and ease of use. The data was gathered using questionnaires, which were requested to be filled in after attending a conference presentation and using the tool. The questionnaire is based on the work presented by [Lund2001].

10 researchers participated in this survey, where all of them were familiar with the Web 2.0 term and have used at least 4 different tools. Also, they are familiar with smartphones and own one themselves. On the survey, the evaluated aspects were represented by a range of agreement from 1 to 7, with 7 the highest agreement

value. On a general level the tool is considered to be useful with an average 5.5. The participants stated their satisfaction with the tool with an average of 5.2. Finally, they consider More! easy to use with an average of 5.7.

As a final part of the evaluation, the researchers were requested to list the positive and negative aspects of the application. Among the positive feedback, they liked the idea of combining the data in one place. The major negative aspect encountered was the problems encountered with the QR codes and the decoding process.

## 3.2 Doctoral Academy Capacity

Here, we begin with the first ideas for supporting doctoral students from [9]. We contrast this with outcomes from workshop with PhD students conducted in [10].

### 3.2.1 Support for a doctoral community of practice

Gillet et al. [9] describe the two instruments STELLAR maintains with regard to the doctoral community. The first instrument is the STELLAR Doctoral Academy which organizes face-to-face events. This is complemented by the Doctoral Community of Practice (DoCoP). The DoCoP aims at complementing the face-to-face events relying on the Science 2.0 infrastructure and *“offering them [the PhD students] opportunities to share, discuss, and receive feedback on their research by peers and experts.”*

The authors of said paper reflect on the practice of PhD candidates in the community, and conclude that two features influencing TEL research are especially relevant in this context: (1) its relative youth which makes it hard to track down resources and experts; (2) the interdisciplinary nature of the field which *“requires multidisciplinary competences infrequently owned by a single individual, the role of the community as a coach to PhD advisors can be seen as stronger in TEL than in other research fields.”* This shows again the importance of supporting the DoCoP properly and highlights the importance of interaction in the community, as well as peer or expert recommendation, possibly supported through the Science 2.0 portal.”

Then Graaasp is introduced, a social software, which would help to overcome those difficulties. It has since been included as a widget in the STELLAR Widget Universe (see Chapter 2). Graaasp allows users (actors) to produce, edit, share and annotate assets in a collaborative environment. Assets can be text files, RSS feeds, wikis, video files, and audio files. The system then produces recommendations for other users consisting of relevant actors, assets, and group activities.

The authors conclude that Graaasp, or a similar social software would provide members of the doctoral community of practice with an environment where they could interact. They see the greatest challenge in bootstrapping such an environment, which requires dedicated PhD students and their supervisors. *“We suggest that once the ‘ball is rolling’, the value of the tool will grow and be recognized by all individuals in the DoCoP, in STELLAR and in the network of stakeholders.”*

### 3.2.2 Needs Analysis

In a workshop at the JTEL Summer School 2010, a needs analysis of PhD students towards a community of practice was conducted. *“[D]octoral candidates worked in groups of 5-6 people and were asked to discuss how they would wish to receive support for their doctoral work in terms of personal support, awareness support, tools for collaboration and the characteristics of a doctoral community of practice that would be of value to them. Each group then presented their findings, explained their results and engaged in discussions about their thoughts with the other participants of the workshop.”*

Results were split into needs on an individual and on a community level.

*“In addition, we asked the 4 groups to consider what kind of awareness support may be helpful in research communities with respect to contributing to increased productivity. With awareness we mean the state or quality of being aware of the current themes, projects, events and researchers including their background within the field of TEL and one’s own position within it.”*

The results from this need analysis are summed up in Table 3 and Table 3: Needs of doctoral candidates on individual and community level

Individual Level	Community Level
Peer-review of artefacts	Information modeling
Methodology	Researcher information
Problem solving	Futuregazing
General feedback	Networking
Jobs/internships/exchange programs	Guidelines for community management
F2F meetings	Sharing testbeds/datasets
Information management	Peer groups
	Collaborative filtering

Table 3: Needs of doctoral candidates on individual and community level

Personal Level	Research Level
Research background	Artefacts/publications
Expertise/competencies	State-of-the-art topic
Projects	Opinions from peers
Social media handles	

Table 4: Awareness needs of doctoral candidates on personal on a research level

From the discussion, three suggestions in correlation with practices emerged:

1. Participants noted that the actual communities of practice that emerge from events such as the summer school are smaller groups of 6 to 10 people. *“The Ph.D. candidates suggested that these smaller communities of practice should be supported not by a particular tool or service, since the community members would decide on those depending on their needs and habits, but rather by the provision of guidelines on collaboration, including the use of existing Web 2.0 tools for research and community management.”*
2. *“The second conclusion the participants drew was that the sustainability of a community of practice, based on the philosophy underpinning Research 2.0, would be highly dependent on individuals dedicated to it. [...] Participants recommended a community facilitator to keep the flow of information going and the community members active in participating.”*
3. The students did not see an immediate need for a new tool or service as they already use a number of tools. Any new tool or service, however, needs to be easy-to-use and it needs to have collaboration as well as communication functions.

The authors conclude that: *“Ph.D. candidates working within the field of TEL feel they have sufficient Web 2.0 tools at their disposal but would appreciate more support in terms of their use as well as finding and filtering information relevant to their research.”*

For the Science 2.0 infrastructure, this means that there is no need to design a new tool or application specifically targeted at doctoral candidates. Nevertheless, we need to adapt existing tools and technologies to cater for the special needs of PhD students. Further research needs to be conducted to deduct concrete user requirements, but possible changes include:

- Adapted facilities for the Doctoral Community of Practice on TEL Europe (e.g. special set of aggregation widgets, groups targeted at PhD students).
- Additional filtering capabilities for publication feeds to reflect the needs of PhD students.

### 3.3 Community Level Capacity

TEL Europe is a social networking platform aimed at all stakeholders in Technology Enhanced Learning. It is based on elgg and offers numerous functionalities. Key features include:

- Users are able to maintain a personal profile and friend other individuals on the platform.
- Users can open up groups, enabling them to start discussions with like-minded members, upload files, and collaboratively edit pages.
- Users are able to keep a personalized dashboard which can be populated with widgets from the STELLAR Widget Universe (see Chapter 2).

In [11], Barak et al. sum up their experiences with TEL Europe in the first four months after the launch of this stakeholder platform. The authors note that with

the three key principles in mind – openness, collaboration and sustainability – TELeurope provides a number of benefits to potential members:

- networking opportunities and being part of a larger community of shared interests,
- increasing the visibility of personal profiles,
- receiving news of projects, events, promotional opportunities, leading developments in TEL and other communications,
- access to resources such as reports and ‘grey’ literature ,
- access to expert discussions and opinions, and
- opportunities for funding, collaboration and employment.

At the more collective level, the community could potentially benefit from:

- collective lobbying power,
- access to test beds on a regional scale,
- expert reviews,
- EU- wide TEL research presence, and
- a “neutral zone” to discuss field related matters; a way of reducing barriers between research, innovation, policy and practice; and contributing to the development of the research agenda related to TEL.

These benefits, however, have not lead to the uptake the authors had hoped for. The authors attribute this to the fact that existing processes and best practices of establishing a stakeholder network have not been taken into account during the creation of TEL Europe. Therefore, *“it has not been successful in clearly communicating some of the key features of the network to its members.”* The authors also note that there are issues concerning usability, complexity, and flexibility. They would like to see more use cases to be supported.

The authors conclude that *“[a]fter one year of activity, and four months since the launch of TELeurope, the STELLAR stakeholder network is experiencing a tension between an organisation-centric and a system-whole approach. On the one hand, the network is defined as a multi-stakeholder network, aimed at providing benefits to the TEL field as a whole, rather than to STELLAR or a specific group of stakeholders inside the network. On the other hand, activities so far have had an organisation-centric feel – trying to demonstrate STELLAR’s achievements more than the actual progress on facilitating inter-disciplinary dialogue. In order to achieve its goals, TELeurope should take a step towards a more network-focused view, possibly by applying participatory design.”*

Through this experience the following additional ideas for features could be indentified:

- Improve the overall structure of the platform in terms of content, activitiy, and network visibility to make the purpose of the platform clear to the user.

- Increase the visibility of attractive sections like CfPs, funding opportunities, and open positions.
- A personalized user experience – perhaps an interface adapted to the background or priority group of the stakeholder.

### 3.4 Leadership Capacity

We begin with the evaluation of the usage of Flashmeeting in the Meeting of Minds as described in D2.2. In terms of Podcasting, we present data from Google Analytics about usage of the podcasting infrastructure.

#### 3.4.1 Meeting of the Minds

Flashmeeting is an online video-conferencing system developed by the Open University's Knowledge Media Institute (KMi). It incorporates audio and video communication, as well as a text chat, a whiteboard, URL and file sharing facilities, and a voting system. [Scott2007]

Flashmeeting supports the work of the network and is extensively used throughout the project. As of the time of writing of this deliverable, more than 240 virtual meetings have been held. An average meeting lasts 62 minutes, has five participants, and has been replayed three times. Within the subsequent section, one particular series of meetings will be investigated in more detail: the Meeting of Minds flashmeeting series.

The first Meeting of the Minds was held from April 21, 2010 to April 23, 2010. It was scheduled to take place at Schloss Dagstuhl, Germany; due to the outbreak of the Icelandic volcano Eyjafjallajökull and subsequent travel hindrances it was held in five virtual meetings using the online video-conferencing system Flashmeeting. In D2.2 the organizers reflect positively on the experience:

*“Video-Conferencing and especially FlashMeeting is surprisingly suitable for such open discussions in a group of 15 participants. Although participants regretted not being able to meet face-to-face, most felt that the last-minute re-organization as a series of FlashMeetings allowed us to still achieve the intended result. One main advantage of the approach is the possibility to record and meta-analyze the meetings as also the documentation of chat logs and document exchange between the participants.”* (STELLAR Deliverable D2.2)

The experience lead to several changes and especially expansions of the concept “Meetings of the Minds”:

- Virtual MoM meeting will be conducted every two or three months, we are currently in the process of scheduling the upcoming virtual MoM;
- Virtual MoM meetings can bring together more STELLAR participants and advisory board members than previously expected;
- Face-to-face MoMs can be better focused and prepared based on the results obtained in the virtual MoMs;

- Recordings of virtual MoMs will be made open to all STELLAR members and the TEL community beyond. This will support the distribution of discussions and results obtained to the STELLAR community (and beyond).

Through this experience the following additional ideas for features could be identified:

- One important challenge we encountered with the first virtual MoM is to raise awareness within the STELLAR community (and beyond) of the discussions and results obtained. We will examine a number of different approaches of feeding the insights gained back into the network.

### 3.4.2 Podcasting Infrastructure

The Podcasting infrastructure allows for uploading of podcasts and access to these podcasts via the project website and via iTunes. To date, there are 35 podcasts available, which were arranged in 10 lists.

The Podcasting infrastructure is tracked in Google Analytics. By 2010/09/08 the feed located at <http://www.stellarnet.eu/feeds/podcasts.xml> was downloaded 3,229 times<sup>7</sup>.

Table 5 shows a overview of podcast lists and the number of page views until 2010/10/17:

Podcast	Page Views	Unique Page Views
Series: TELstars interviews	168	140
Series: Science 2.0 workshop	117	99
Series: Vania Dimitrova on EC-TEL 2010	68	59
Series: Peter Brusilovsky Adaptive LS	64	62
Series: Ways of conducting SNA	47	36
Series: Voices JTEL summer school 2009	41	27
Series: Ways of conducting SNA	37	29
Series: Neuroscience and TEL	21	16
Series: e-teaching.org Podcasts 2009	14	9
Series: Ulrike Lucke SIG e-learning	10	8

Table 5: Podcast Statistics from stellarnet.eu

<sup>7</sup> Downloads of the Open Archive were excluded, because the podcasting feed is fetched everytime someone opens the homepage.

Table 6 and Figure 7 show an overview of actual downloads of single podcasts from 2010/09/04 to 2010/11/05:

File	Requests
/kmi/podcasts/24_20100706_peterb.mp3	159
/kmi/podcasts/19_law-usability2.m4v	68
/kmi/podcasts/19_klamma-wisdom-of-crowds.m4v	66
kmi/podcasts/27_20100921_14-09-2010_10-34-33_vania_ectel_v4.mp3	35
/kmi/podcasts/28_20100923_horizon-report-teil1.mp3	
/kmi/podcasts/25_20100712_sne_julia_eberle_and_kris_lund.mp3	34
/kmi/podcasts/17_20091228_gillet.m4v	33
/kmi/podcasts/22_anna.m4v	30
/kmi/podcasts/17_20091228_ochoa.m4v	29
/kmi/podcasts/22_hazan2.m4v	29
Other (including index pages)	399

Table 6: Actual podcast downloads (Source: <http://residentialschoolblogs.open.ac.uk/stellar/report/>)

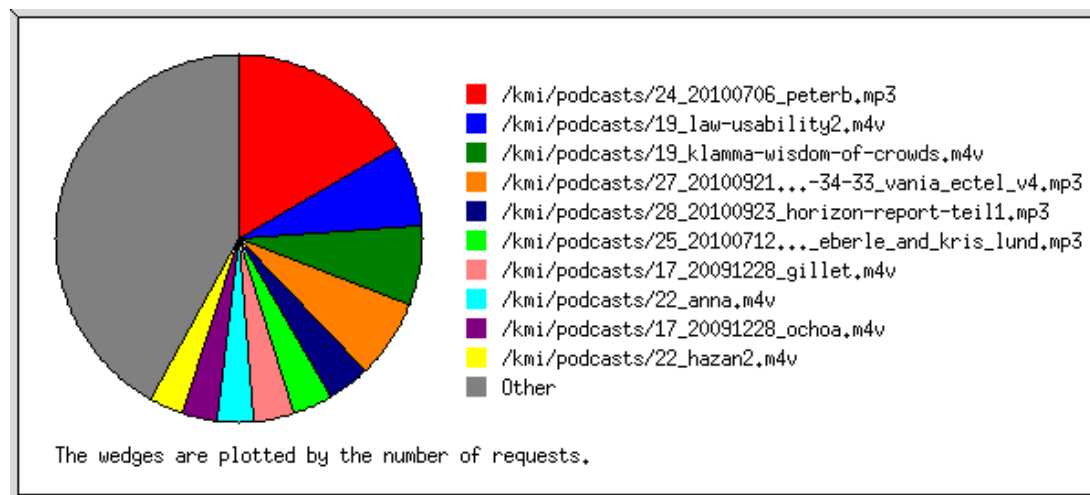


Figure 7: Distribution of actual podcast downloads (Source: <http://residentialschoolblogs.open.ac.uk/stellar/report/>)

The podcasts are provided in two ways. First of all they are distributed over the popular iTunes podcasting infrastructure. This makes them accessible on a wide range of systems (operating systems, mobiles) and programs. The podcasts are also distributed via a video player based on Flash. This has the additional benefit that they can be used more easily for mash-ups, such as the widget universe describe above.

Monitoring usage of the podcasts required several modifications of the webserver logging mechanisms, as Google Analytics cannot track access information about downloads and streaming (progressive downloads). It turned out that obtaining



reliable usage statistics is much harder than expected. We are currently examining the usage data in detail.

Through this experience the following additional ideas for features could be identified:

- Multi-channel advertisement strategy of the podcasts (TEurope, stellarnet, dedicated blogs, etc.) to create more visibility.
- Improve logging mechanisms to obtain reliable usage data of downloads and streaming.

## 4 Conclusions

There is already a considerable number of papers on the STELLAR Science 2.0 infrastructure, published at EC-TEL and ED-MEDIA. They fall into two main categories. In the first category, there are papers of a conceptual nature, which describe the technical side of the infrastructure. These papers report on the set up of the infrastructure based on the immediate needs of the STELLAR NoE.

In the second category there are evaluation papers. These papers incorporate data from logging, quasi-experimental set-ups, and workshops. Currently, the evaluation papers are largely of an exploratory, descriptive and qualitative nature. They usually shed light on a single application or a certain use case. The results presented in these papers provide valuable insights on the emerging Science 2.0 practices in STELLAR and the applicability of the infrastructure components to the STELLAR instruments. We will build on these reports in two ways: (1) to refine and expand the existing infrastructure (development roadmap) and (2) to design future evaluations of the infrastructure (evaluation plan).

The **development roadmap** will take the here identified ideas for features into account and aims to streamline the future development activities. The main lessons learned from the usage experiences are:

- in many cases no new tools/applications are needed to support scientific work, however target group specific settings and filters (for personalization) as well as connectors between tools are needed
- functionalities need to be provided in small increments such as in widgets, these allow individual users and user groups to select the features they need and to combine them in innovative ways
- in order to support the scientific work of whole communities and groups intelligent “control panels” are needed which provide a condensed overview of the different ongoing activities of individual researchers, in addition “gardening environments” will help to support the emergence (maturing) of collaboratively defined structures and visions

These findings lead us to the notion of **Personal Research Environments** as an innovative way to improve individual scientific work. The idea of Personal Research Environments (PRE) is very similar to the idea of Personal Learning Environments (PLE) however adapted to the application area of scientific work. PRE rely on the same technical approaches (e.g. widgets) and face many of the same issues (not every researcher is willing/able to invest time in the construction of his/her own research environment). We will organize our development plan around this notion of PRE and take into account the main lessons learned (above). In order to speed up the development of a PRE we will specifically examine research into PLEs (e.g. MATURE) and its counterpart Community Learning Environments (CLE). While PREs focus on supporting individual research work, Community Research Environments (CREs) will focus on the community and “gardening” aspects.

On the technical side, widgets have the potential to provide researchers with little mini-applications, which they can embed in a variety of platforms. Due to the open standard of this widgets, re-usability and re-mashing of functionalities allows researchers to tailor and create widgets for new research challenges, and allows them to share them with other researchers over the STELLAR directory of widgets. Additional to this, the widget approach allows to tailor platforms like TELeurope in ways, which are not supported by default. This is especially important for the area of research with its high dynamic and changing requirements.

The **evaluation plan** will examine a variety of possibilities with which we can foster (real world) usage experiences of the newly developed Science 2.0 infrastructure parts (PRE). It will also draw from our experience what methods are appropriated to evaluated certain aspects of the infrastructure. This evaluation plan will inform and refine the methods we will choose for Deliverable D6.6.

## 5 References

### 5.1 Overview of Papers

#	Authors (Event and Year)	Title	Abstract
<b>Infrastructure</b>			
1	Ullmann T., Wild F., Scott P., Duval E., Parra G., Reinhardt, W., Heinze N., Kraker P., Fessl A., Lindstaedt S., Nagel T., Gillet D. (Proceedings of the Poster Track at EC-TEL 2010)	Components of a Research 2.0 Infrastructure	In this paper, we investigate the components of a Research 2.0 infrastructure. We propose building blocks and their concrete implementation to leverage Research 2.0 practice and technologies in our field, including a publication feed format for exchanging publication data, a RESTful API to retrieve publication and Web 2.0 data, and a publisher suit for refining and aggregating data. We illustrate the use of this infrastructure with Research 2.0 application examples ranging from a Mash-Up environment, a mobile and multitouch application, thereby demonstrating the strength of this infrastructure.
2	Kraker P., Fessl A., Hoefler P., Lindstaedt S. (Research 2.0 Workshop at EC-TEL 2010)	Feeding TEL: Building an Ecosystem Around BuRST to Convey Publication Metadata	In this paper we present an ecosystem for the lightweight exchange of publication metadata based on the principles of Web 2.0. At the heart of this ecosystem, semantically enriched RSS feeds are used for dissemination. These feeds are complemented by services for creation and aggregation, as well as widgets for retrieval and 3visualization of publication metadata. In two scenarios, we show how these publication feeds can benefit institutions, researchers, and the TEL community. We then present the formats, services, and widgets developed for the bootstrapping of the ecosystem. We conclude with an outline of the integration of publication feeds with the STELLAR Network of Excellence <sup>1</sup> and an outlook on future developments.
<b>Applications and Visualizations</b>			
3	Parra G., Duval E. (Proceedings of the ED-MEDIA 2010)	More! A Social Discovery Tool for Researchers	Science 2.0 is the result of Web 2.0 tools and trends influencing the research area. In this paper, we focus on a scenario where a researcher is interested in the topic and speaker at a conference: finding more information about them is far from instantaneous. Thus, we identified a need

			to not only find speaker and paper information during the presentation, but also to subscribe to feeds that keep the attendee informed about ongoing work from the speaker. This work presents the development of a mobile application that groups all the relevant information of a speaker in a way that can be easily exposed and integrated in the normal workflow of the audience. The result is a frictionless blending of the face-to-face event that a conference or workshop presentation represents with the rich interaction and alerting services that a web2.0 environment provides.
4	Vandeputte B., Duval E. (Research 2.0 Workshop at EC-TEL 2010)	Research at the table	In this paper we describe how we want to take advantage of the rapid developments in technology to assist researchers in doing research. More specifically in exploring the publication space. For this purpose we have designed and developed a prototype application to take advantage of large displays with multi touch enabled input. We describe the current state, the next steps and how we will to evaluate it. To conclude we give an outlook on further possibilities and challenges that lay ahead.
5	Nagel T., Duval E. (Research 2.0 Workshop at EC-TEL 2010)	Visualizing the origins and connections of institutions based on co-authorship of publications	This paper introduces Muse, an interactive visualization of publications to explore the collaborations between institutions. For this, the data on co-authorship is utilized, as these signify an existing level of collaboration. The affiliations of authors are geo-located, resulting in relations not only among institutions, but also between regions and countries. We explain our ideas behind the visualization and the interactions, and briefly describe the data processing and the implementation of the working prototype. The prototype focuses on a visualization for large tabletop displays, enabling multiple users to explore their personal networks, as well as emerging patterns in shared networks within a collaborative public setting. For the prototype we used the publication data of the EC-TEL conference.
6	Fisichella M., Herder E., Marenzi I., Nejd W. (Proceedings of the ED-MEDIA 2010)	Who are you working with? - Visualizing TEL Research	Author Co-Citation Analysis (ACA) provides a principled way of analyzing research communities based on how often authors are cited together in scientific publications. In this paper, we

		Communities	present preliminary results based on ACA to analyze and visualize research communities in the area of technology enhanced learning, focusing on publicly available citation and conference information provided through CiteseerX and DBLP. We describe our approach to collecting, organizing and analyzing appropriate data, as well as the problems which have to be solved in this process. We also provide a thorough interpretation of the TEL research clusters obtained, which provide insights into these research communities. The results are promising, and show the method's potential as regards mapping and visualizing TEL research communities, making researchers aware of the different research communities relevant for technology enhanced learning, and thus better able to bridge communities wherever needed.
	<b>Researcher Capacity</b>		
7	Joubert M., Sutherland R. (Research 2.0 Workshop at EC-TEL 2010)	Research 2.0: Drawing on the wisdom of the crowds to develop a research vision	This paper describes and reflects upon taking a 'Research 2.0' approach to developing a 'vision and strategy statement' for a network of researchers involved in researching Technology Enhanced Learning (TEL). It relates how the statement was developed first by collecting content from colleagues within the network through face to face meetings and contributions to a wiki and then by creating a coherent linear text document which further developed the content on the wiki. It discusses the risks inherent in the approach and outlines the strategies taken to address the risks. It suggests that, although the approach taken was successful, the success was limited owing to factors including a) limited engagement by the community with other people's contributions, b) a reluctance to amend other people's contributions and c) the difficulty of aggregating the multiple voices within the community while retaining faithfulness to the philosophies underpinning a 'Research 2.0' approach.
8	Wild F., Ullmann T. (Research 2.0 Workshop at EC-TEL 2010)	The afterlife of "living deliverables": angels or zombies?	Within the STELLAR project, we provide the possibility to use living documents for the collaborative writing work on deliverables. Compared to 'normal'

			<p>deliverables, 'living' deliverables come into existence much earlier than their delivery deadline and are expected to 'live on' after their official delivery to the European Commission. They are expected to foster collaboration. Within this contribution we investigate, how these deliverables have been used over the first 16 months of the project. We therefore propose a set of new analysis methods facilitating social network analysis on publicly available revision history data. With this instrumentarium, we critically look at whether the living deliverables have been successfully used for collaboration and whether their 'afterlife' beyond the contractual deadline had turned them into 'zombies' (still visible, but no or little live editing activities). The results show that the observed deliverables show signs of life, but often in connection with a topical change and in conjunction with changes in the pattern of collaboration.</p>
	<b>Doctoral Academy Capacity</b>		
9	Gillet D., Helou S., Joubert M., Sutherland R. (Science 2.0 Workshop at EC-TEL 2009)	Science 2.0: Supporting a Doctoral Community of Practice in Technology Enhanced Learning using Social Software	<p>This paper focuses on the ways in which STELLAR supports doctoral candidates through the establishment of a Doctoral Community of Practice (CoP) in Technology Enhanced Learning as a STELLAR doctoral integration instrument for the doctoral stakeholder community. This TEL Doctoral CoP (DoCoP), officially established in Autumn 2009, is also instrumental in bringing together actors of engineering education research in academic institutions, as proposed in the USA by Streveler. The paper discusses possible ways in which the DoCoP could be developed through the innovative use of Web 2.0 technologies, by outlining the characteristics of one such technology and describing the ways in which an imaginary PhD candidate might use the technology in their PhD journey.</p>
10	Heinze N., Joubert M., Gillet D. (Research 2.0 Workshop at EC-TEL 2010)	Connecting Early Career Researchers: Investigating the Needs of Ph.D. Candidates in TEL Working with Web 2.0	<p>This article describes the results of a case study conducted amongst 21 doctoral candidates and three senior researchers at the Joint European Summer School on Technology Enhanced Learning 2010. The study aims to analyse the needs of early career researchers working within the field of TEL in geographically distant communities, particularly with respect to</p>

			online collaboration, communication and information exchange. This study can be seen as a needs analysis on support structures to enable research 2.0 in TEL among young researchers.
	<b>Community Level Capacity</b>		
11	Barak N., Burgos, D., Camilleri, A., Vries F., Specht M., Windrum C. (Proceedings of the EC-TEL 2010)	Modelling a Stakeholder Community via a Social Platform: The Case of teleurope.eu	Past attempts at creating stakeholder networks for specific fields of research or industrial sectors have shown to be a resource-consuming and timeconsuming process, which requires continuous monitoring and political efforts, as well as the trial-and-error deployment of technological tools. Still, these networks are thought to be an efficient and essential communication instrument for addressing challenges and building capacities. The EU FP7 STELLAR Network of Excellence has the mission of establishing a network for Technology Enhanced Learning (TEL) stakeholders, and has decided to do so via an online social community called TELeurope. In this paper we provide an overview of some relevant experience in establishing collaborative networks in the fields of business sciences, learning networks and communities of practice and reflect on our experience thus far with TELeurope.

## 5.2 Additional References

[Lund2001] Lund, A. Measuring usability with the USE questionnaire. *Usability Interface*. 2001;2(8). Available at: [http://www.stcsig.org/usability/newsletter/0110\\_measuring\\_with\\_use.html](http://www.stcsig.org/usability/newsletter/0110_measuring_with_use.html)

[Mika2005] 1. Mika P, Klein M, and Radu Serban. Semantics-based Publication Management using RSS and FOAF. In: Decker S, Park J, Quan D, Sauermann L, eds. *Proc. of Semantic Desktop Workshop at the ISWC, Galway, Ireland, November 6*. Vol 175.; 2005. Available at: [http://ceur-ws.org/Vol-175/6\\_mika\\_swwu\\_final.pdf](http://ceur-ws.org/Vol-175/6_mika_swwu_final.pdf).

[Scott2007] Scott P, Tomadaki E, Quick K. The shape of live online meetings. *International Journal of Technology, Knowledge & Society*. 2007;3(4):1-16. Available at: <http://oro.open.ac.uk/8641/>.



# Annex 1: Papers



This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

# Components of a Research 2.0 Infrastructure

Thomas Daniel Ullmann<sup>1</sup>, Fridolin Wild<sup>1</sup>, Peter Scott<sup>1</sup>, Erik Duval<sup>2</sup>, Bram Vandeputte<sup>2</sup>, Gonzalo Parra<sup>2</sup>, Wolfgang Reinhardt<sup>3</sup>, Nina Heinze<sup>4</sup>, Peter Kraker<sup>5</sup>, Angela Fessl<sup>5</sup>, Stefanie Lindstaedt<sup>5</sup>, Till Nagel<sup>6</sup>, Denis Gillet<sup>7</sup>

<sup>1</sup>KMi, The Open University, Walton Hall, Milton Keynes, United Kingdom  
{f.wild, t.ullmann, peter.scott}@open.ac.uk

<sup>2</sup>Departement Computerwetenschappen, Katholieke Universiteit Leuven, Leuven, Belgium  
{erik.duval, bram.vandeputte, gonzalo.parra}@cs.kuleuven.be

<sup>3</sup>University of Paderborn, Fürstenallee 11, Computer Science Education Group, 33102 Paderborn, Germany  
wolle@upb.de

<sup>4</sup>Knowledge Media Research Center, Konrad-Adenauer-Straße 40, Tuebingen, Germany  
n.heinze@iwm-kmrc.de

<sup>5</sup>Know-Center and Graz University of Technology, Knowledge Management Institute, Austria  
{pkraker, slind, afessl}@know-center.at

<sup>6</sup>Fachhochschule Potsdam, Potsdam, Germany  
nagel@fh-potsdam.de

<sup>7</sup>Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland  
denis.gillet@epfl.ch

**Abstract.** In this paper, we investigate the components of a Research 2.0 infrastructure. We propose building blocks and their concrete implementation to leverage Research 2.0 practice and technologies in our field, including a publication feed format for exchanging publication data, a RESTful API to retrieve publication and Web 2.0 data, and a publisher suit for refining and aggregating data. We illustrate the use of this infrastructure with Research 2.0 application examples ranging from a Mash-Up environment, a mobile and multitouch application, thereby demonstrating the strength of this infrastructure.

**Keywords:** research 2.0; infrastructure; mash-ups; #Res2TEL

## 1 Research 2.0

In technology-enhanced learning (TEL), the use of Web 2.0 technologies is now actively researched under banners such as "Learning 2.0" [1], "Personal Learning Environments" [2] or "Open Learning Environments" [3] and the like. In our Research 2.0 work, we aim to leverage the same opportunities for research on TEL. Research 2.0 can be defined as the application of new practices that focus on opening up the research process to broaden participation and collaboration with the help of new technologies that are able to foster continuous engagement and further development.

The basic idea is that, as researchers in technology-enhanced learning, we already know how to make use of for example blogs, wikis and forums to enhance collaborative work, but a full Research 2.0 framework might provide us with a much more powerful structure to make our research more effective.

The proposed components of a research infrastructure build upon the ideas of Research 2.0. By now, the focus is on individual practice and especially on the information management of publication and social media data. Based on this foundation, future extension will strengthen collaborative and community practice for a full “Research 2.0” framework.

The paper is organized as follows. We first outline the three main components of the research information infrastructure. It follows an outline of a publication format, of services for publication and Web 2.0 data, and a publisher suit. The interplay between these components is shown with three applications, which are built on top of the infrastructure. Finally, we conclude and give a forecast about the next development steps.

## 2 Components of a TEL Researcher Information Infrastructure

The architecture of the infrastructure foresees three cornerstones [4]. (1) On the server side, services provide the backing data for the tools and widgets. The data are retrievable through a RESTful API. (2) On the client-side, widgets are combined into a coherent user experience with the help of a mash-up environment. Mobile and multitouch applications use their own environment. (3) Widgets are administered in a directory, thereby subjecting the management of the portfolio to conscious maintenance and development. The fundament of the infrastructure tying these three pillars together is a set of interoperability formats.

Based on these cornerstones of Research 2.0 architectures we implemented data services, tools and widgets, using interoperability formats. We begin with the description of a publication exchange format. This defines a minimum set of guidelines easing the usage across different systems and partner infrastructures. It follows two data services approaches, one for research data including publication data and Web 2.0 data, and a publisher suit. These services are accessible for the use in tools and widgets. We outline three of them, which especially show the strength of the Research 2.0 mash-up architecture for the use in different application fields, including desktop, multitouch and mobile applications. We begin with the interoperability format.

**Publication feeds:** In order to facilitate the exchange of bibliographic data across the TEL community we use the concept of publication feeds. They are used for a lightweight exchange of publication metadata in a format commonly readable by existing Web 2.0 infrastructure. Hence, they can easily be combined, aggregated, visualized and re-released. This allows for inclusion of external parties who can expose their publication data through publication feeds as well. An institution only needs to export its publication metadata once to automatically update all the

subscribers to this feed (e.g. the STELLAR<sup>1</sup> Open Archive<sup>2</sup>). Publication feeds are RSS 1.0 feeds enhanced with elements from the SWRC and DC ontologies. The feeds are based on the BuRST format [5]. The basis for the publication feed are RSS 1.0, RDF, DC 1.1, SWRC 0.3, and BuRST 0.1. Modifications were applied where the format was outdated or underspecified.

**ResearchFM service:** The ResearchFM API was proposed as a RESTful API to provide publication and social data of authors in a unified way. Publication data shed light on of communication and collaboration of a research community, e.g. through analysis of co-authorship, co-citations and conference themes. With social media content, there is an unfathomable amount of data being generated almost constantly on the Web from research communities aside from the “official” publications. Heinze et al. [6] point out a number of Web 2.0 tools that are actively used during the daily work of researchers. However, in many community and group work situations the awareness of others is essential for effective and efficient work. This can be especially true in conference settings, since they provide the time and space for exploring new themes, finding like-minded researchers, or finding out what is being discussed online about one’s own work. Reinhardt et al. [7] propose the model of Artefact-Actor-Networks (AANs) to store, analyse and visualise the actions, connections and structure of individuals within research communities on both social and artefact level. Therefore, they monitor the community’s activities on social media sites based on given tags or given online handles and analyse the content of the gained artefacts. Every artefact is stored together with its metadata, semantic annotations and connections to other artefacts in a semantic database. Furthermore, the relations to actors referring to an artefact (e.g. creating, linking, retweeting, forwarding, discussing about, favouring, tagging) are stored and allow analysing the nexus of a community starting from any artefact or actor in the Artefact-Actor-Network. Furthermore, it allows the identification of semantically similar artefacts or actors from their respective content, extending the possibilities of co-citation measures or co-authorship relations.

As all the collected data is very similar on the one hand, and the tools and widgets use this data in a similar way on the other hand, it became apparent that a lot of benefit could come from a common API in terms of interconnectivity and reusability.

**Services for publication data:** A suite of publisher services was released to aid institutions and individuals in producing, aggregating and refining publication feeds in producing, aggregating and refining publication feed. The services include a BibTeX converter as well as a feed merger and a feed filtering service: these services can be mashed together, e.g. by using DERI pipes<sup>3</sup>. Additional to the data from the STELLAR Open Archive further TEL specific publication data has been gathered, namely the publication data of two conferences EC-TEL and ED-MEDIA, with others to follow. This will help to feed more data into the Archive, and form an interesting foundation for tools and widgets to build upon. To have easy access to this data, all tools and widgets will be able to use the unified ResearchFM service.

---

<sup>1</sup> <http://stellarnet.eu>

<sup>2</sup> <http://oa.stellarnet.eu/>

<sup>3</sup> <http://pipes.deri.org>

Build upon the data services and interoperability format three applications are used to demonstrate the wide usage of the Research 2.0 infrastructure.

**STELLAR Widget Universe:** Builds upon the mash-up idea. It uses Elgg<sup>4</sup>, an open source networking and publishing software, as showcasing platform for bringing together widgets and services and the legacy systems of the STELLAR partners. The widgets are delivered through the Wookie widget engine<sup>5</sup>. A plugin for Elgg enables to embed the widgets into Elgg (plugins for Wordpress, Moodle, LAMS exist as well). Researchers can arrange a widget per drag-and-drop on their dashboard. A list shows the gallery of all available widgets from the STELLAR directory. After the selection, the widget is automatically instantiated and can be used by the researcher. All widgets are packaged according to the widget 1.0 specification<sup>6</sup> and can thus not only be run within the reference implementation called Universe, but similarly within STELLAR's stakeholder platform TELEurope<sup>7</sup>.

**ScienceTable:** While the widget universe is browser based, the ScienceTable is a multitouch tabletop application for the collaborative exploration of publication data. This tool allows for an interactive exploration of co-authorship relations. Its layout is completely dynamic, based on a spring graph algorithm. The ScienceTable can be interesting for a researcher exploring his own collaborations or exploring the clusters of co-operating authors in the field. In order to start navigation, search for a specific author is supported. Exploration happens through zoom, pan, drag and tap gestures on a large multi-touch tabletop. Extensions towards citation data are planned for the near future.

The **More!** application [8]: This application is build for mobile devices. Its purpose is to let researchers find information about for example a speaker at a conference and to subscribe to feeds from social tools that keep the attendee informed about ongoing work from the speaker. The application exposes the following information:

- Speaker: full name, photo, e-mail, affiliation and publication list
- Current presentation: slides and paper
- Social tools: Twitter, SlideShare, blog, Delicious, LinkedIn, and Facebook

The following figure gives an overview of the above outlined components of the Research 2.0 information infrastructure. The publication data are collected through the publication feed format. These data and social media data are retrievable through the ResearchFM API, which serves as the backing data for the applications, like the STELLAR universe, the ScienceTable, the More! application and many more.

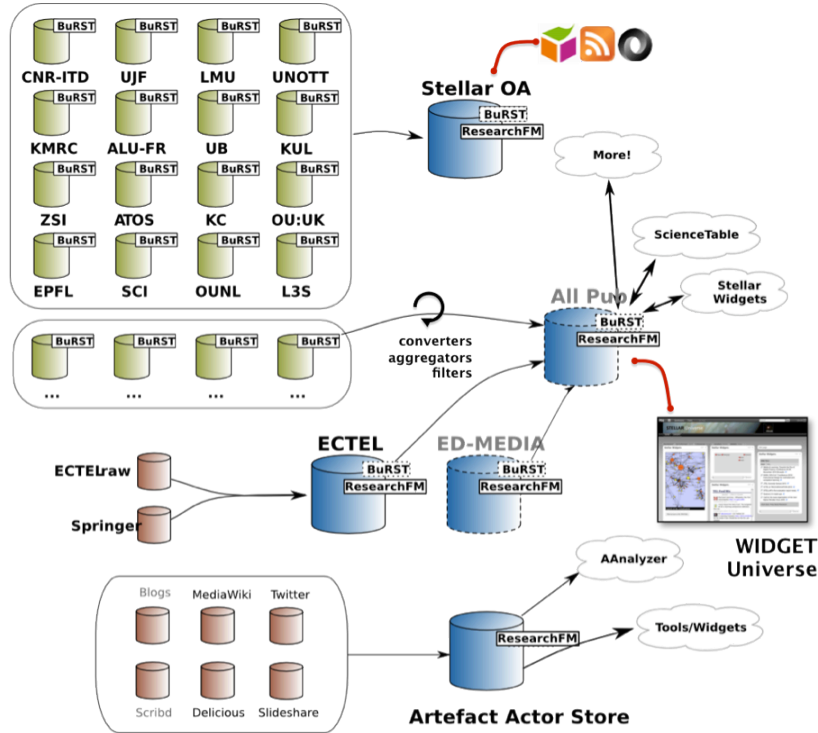
---

<sup>4</sup> <http://elgg.org>

<sup>5</sup> <http://incubator.apache.org/wookie/>

<sup>6</sup> <http://www.w3.org/TR/widgets/>

<sup>7</sup> <http://www.teleurope.eu/>



**Fig. 1.** The Components of the TEL Research 2.0 Information Infrastructure

### 3 Conclusions and Future Work

We proposed a mash-up infrastructure allowing for continuous innovation, by recombining and repurposing existing technology, and showed concrete implementations. With this, the first steps towards a Research 2.0 framework have been made. The outlined Research 2.0 architecture can help to support the practices of researchers providing them with tools to discover and develop their research field.

The Research 2.0 infrastructure lays the foundation for researchers to experience new practices and provides a rich set of data (publication and social media data) to explore further possibilities. Overall, broadening participation means broadening communication and therefore Research 2.0 must aim at supporting research communities in information processing creating more awareness amongst the members of a research community.

While the components of the infrastructure by now focus on the practice of information provision and distribution, for a full Research 2.0 framework further practices, like collaborative and community practice need to be taken into account. They will serve as a further testbed helping to determine extension and modification

6 Ullmann, Wild, Scott, Duval, Vandeputte, Parra, Reinhardt, Heinze, Kraker, Fessler, Lindstaedt, Nagel, Gillet

needs. However, with the use of Mash-Up environments we see suitable support for the later two, allowing users to engage in collaboratively in a personal research Mash-Up environment.

Although the concepts outlined here focus on the domain of technology-enhanced learning, they might very well apply to several other domains.

## ACKNOWLEDGMENT

Special thanks go to Nicolas Balacheff and Jerome Zeiliger from the Laboratory of Informatics Grenoble for their work on the Open Archive. This work was carried out as part of the STELLAR network of excellence, which is funded by the European Commission (grant agreement no. 231913). The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

## REFERENCES

1. Ala-Mutka, K., Bacigalupo, M., Kluzer, S., Pascu, C., Punie, Y., Redecker, C.: Learning 2.0: The Impact of Web 2.0 Innovation on Education and Training in Europe, <http://ipts.jrc.ec.europa.eu/publications/pub.cfm?id=2139>.
2. Wilson, S., Liber, O., Johnson, M., Beauvoir, P., Sharples, P., Milligan, C.: Personal Learning Environments: Challenging the dominant design of educational systems. *Journal of e-Learning and Knowledge Society*. 2, (2007).
3. Hannafin, M., Land, S., Oliver, K.: Open learning environments: Foundations, methods, and models. *Instructional-design theories and models: A new paradigm of instructional theory*. 2, 115–140 (1999).
4. Wild, F., Ullmann, T.D., Scott, P.: The STELLAR Science 2.0 Mash-Up Infrastructure. Presented at the 10th IEEE International Conference on Advanced Learning Technologies, Sousse, Tunisia (2010).
5. Mika, P.: Bibliography Management using RSS Technology (BuRST), <http://www.cs.vu.nl/~pmika/research/burst/BuRST.html>.
6. Heinze, N., Bauer, P., Hofmann, U., Ehle, J.: Kollaboration und Kooperation in verteilten Forschungsnetzwerken durch Web-basierte Medien – Web 2.0 Tool in der Wissenschaft. *GMW2010*. (2010).
7. Reinhardt, W., Moi, M., Varlemann, T.: Artefact-Actor-Networks as tie between social networks and artefact networks. In *Proceedings of CollaborateCom 2009*, November 2009.
8. Parra, G. Duval, E.: More! A Social Discovery Tool for Researchers. *ED-MEDIA 2010-World Conference on Educational Multimedia, Hypermedia & Telecommunications*. Accepted. (2010).



---

# Feeding TEL: Building an Ecosystem Around BuRST to Convey Publication Metadata

Peter Kraker<sup>1</sup>, Angela Fessler<sup>1</sup>, Patrick Hoefler<sup>1</sup> and Stefanie Lindstaedt<sup>1</sup>

<sup>1</sup> Know-Center Graz, Inffeldgasse 21a,  
8010 Graz, Austria  
{pkraker, afessler, phoefler, slind}@know-center.at

**Abstract.** In this paper we present an ecosystem for the lightweight exchange of publication metadata based on the principles of Web 2.0. At the heart of this ecosystem, semantically enriched RSS feeds are used for dissemination. These feeds are complemented by services for creation and aggregation, as well as widgets for retrieval and visualization of publication metadata. In two scenarios, we show how these publication feeds can benefit institutions, researchers, and the TEL community. We then present the formats, services, and widgets developed for the bootstrapping of the ecosystem. We conclude with an outline of the integration of publication feeds with the STELLAR Network of Excellence<sup>1</sup> and an outlook on future developments.

**Keywords:** science 2.0, web 2.0, mashups, services, widgets, feeds

## 1 Introduction

Recently, developments under the paradigm of Science 2.0 have received a lot of attention [1]. Researchers are embracing the capabilities of Web 2.0 tools and technologies, such as blogs, wikis, and social networking sites, to support their research. Using Web 2.0 for scientific work has numerous potential advantages: it possibly leads to shorter feedback cycles, enhances the communication between researchers, and yields a higher penetration of ideas. One of the prerequisites for the introduction of a modern Science 2.0 in the field of Technology Enhanced Learning is the wide-spread access to resources, data, and publications for the whole community [2].

In this paper we present an ecosystem for the exchange of publication data based on existing Web 2.0 infrastructure. At the heart of this ecosystem, semantically enriched feeds based on the popular RSS format [3] are used as a means for lightweight exchange of information on the web. They can easily be combined, aggregated, visualized, and republished. Hence, publication feeds have the advantage

---

<sup>1</sup> STELLAR [4] is an EU-funded Network of Excellence, which aims at unifying the diverse community in the field of Technology Enhanced Learning in Europe.

to provide important scientific data in a format widely used by existing Web 2.0 infrastructure.

To facilitate the opening of institutional archives, easy-to-use tools are needed. Web services are especially apt for this, since they are the cornerstone of Web 2.0, allowing for loosely coupled systems and simple syndication [5]. Whereas the services aid the producer in generating a publication feed, widgets let the recipient consume and manipulate these feeds. Users can collectively contribute to the data base by adding their own feeds; they can help identify good publications by rating them, and interact with each other by leaving comments. A visualization widget provides them with filtering and searching facilities for the aggregated data.

This paper consists of three sections. At first, we introduce two scenarios for the usage of publication feeds in research from a personal and an organizational perspective. Then, we present the pillars of the ecosystem, namely the adapted BuRST format, a suite of web services for feed producers, and several widgets for feed consumers. Finally, we conclude with an overview of the integration of the ecosystem into the STELLAR Network of Excellence and an outlook on future developments.

## **2 Scenario**

In the following section we present two scenarios which illustrate the benefits of the presented ecosystem. These scenarios emphasize on lightweight dissemination, visualization, and navigation of semantically-enriched scientific publication feeds in the style of Web 2.0.

### **2.1 Scenario 1: Semi-automated dissemination of publication feeds**

Sandra is a supervisor at a TEL research institution dedicated to professional learning. She is responsible for collecting the publications of her group. Therefore, her assistants keep a BibTeX file of their publication metadata, which is periodically uploaded to a common server. Sandra is interested in a wider dissemination of this data, but unfortunately she cannot get her assistants to enter the publication data over and over again into other repositories. Hence, she is looking for a way to automate dissemination. Since publication data is already available in several BibTeX files, she uses a dedicated BibTeX converter to convert these files into publication feeds. The resulting individual feeds are then merged into a single feed with the help of the Publication Feed Merger. Due to the fact that there are also publications not related to TEL in the feed, a Publication Feed Filter is applied. Sandra now publishes this feed so that all interested parties that support the BuRST format can subscribe to it.

## **2.2 Scenario 2: Explorative research on publication feeds**

Kurt is an early-career researcher interested in professional learning. He wants to find out about the most influential publications, recently trending topics, and interesting conferences in the field. Therefore, he joins a special interest group dedicated to professional learning on a social networking platform. Sandra and other users have already added their institutions' publication feeds to this group. The individual publications are presented as blog posts, which can be rated and commented on. Kurt now has an overview of the top rated publications and the discussions revolving around them.

Kurt then opens the "Publication Visualization" widget from within the special interest group. He is presented with a faceted browsing view containing all publication metadata from the feeds. A tag cloud aggregated from the keywords is additionally shown to Kurt. He then restricts the data to certain years to see the changes in the tag cloud. This allows him to reflect on the trending topics.

Next, Kurt restricts the publication type to conference proceedings. Now, all proceedings titles are presented to him, alongside the corresponding articles. From the keyword tag cloud, he chooses a topic that he finds interesting. This supplies Kurt with a list of conferences that are important for that specific topic.

## **3 Publication Feed Ecosystem**

In this section, we present the three initial pillars of the publication feed ecosystem: the adapted BuRST format, a suite of web services for feed producers, and several widgets for feed consumers.

### **3.1 Publication Feeds**

Publication feeds are RSS 1.0 feeds, enhanced with elements from the SWRC<sup>2</sup> and DC<sup>3</sup> ontologies. These feeds are an adaption of the BuRST<sup>4</sup> format, proposed by Peter Mika [6]. The bases for BuRST [7] are RSS 1.0 [2], RDF [8], DC 1.1 [9], and SWRC 0.3 [10]. Modifications were applied where the format was outdated or underspecified. It is, for example, not possible to express affiliation in FOAF<sup>5</sup> other than by providing the URL of the institution. As this is not always feasible, the affiliation attribute of SWRC is suggested to represent this data in free text. A complete reference of the publication feed format can be found at [11].

---

<sup>2</sup> Semantic Web for Research Communities

<sup>3</sup> Dublin Core

<sup>4</sup> Bibliography Management using RSS Technology

<sup>5</sup> Friend of a Friend

See below for an exemplified item representation. The item is divided into two parts:

1. A native RSS part
2. A RDF extension part (highlighted in grey)

Both parts are linked through the `burst:publication` property. Information given in the RSS part of the item is mainly intended for display purposes (e.g. in RSS feed readers or widgets), and for processing in other tools which can deal with RSS (e.g. Yahoo! Pipes). The RDF extension part describes the publication in a semantically much more sophisticated way. This part is intended for tools and services that are able to process and display BuRST feeds (see sections 3.2 and 3.3), as well as semantic web applications that understand RDF.

Example of a publication represented in a BuRST feed.

```
<item rdf:about="http://know-center.tugraz.at/papers/16"
xml:lang="en">
  <title>A Storyboard of the APOSDLE Vision</title>
  <link>http://www.aposdle.tugraz.at/content/download/288/1411/file/1
indstaedt_mayer_APOSDLE_poster_p.pdf</link>
  <description>Lindstaedt, S. N., Mayer, H. (2006): A Storyboard of
the APOSDLE Vision.</description>
  <dc:date>2009-10-27T14:40:18+01:00</dc:date>
  <burst:publication>
    <swrc:InProceedings>
      <swrc:title>A Storyboard of the APOSDLE Vision</swrc:title>
      <swrc:author>
        <swrc:Person>
          <swrc:name>Lindstaedt, Stefanie N.</swrc:name>
        </swrc:Person>
      </swrc:author>
      <swrc:author>
        <swrc:Person>
          <swrc:name>Mayer, Harald</swrc:name>
        </swrc:Person>
      </swrc:author>
      <swrc:booktitle>Proceedings of the First European Conference
on Technology Enhanced Learning</swrc:booktitle>
      <swrc:year>2006</swrc:year>
      <swrc:month>10</swrc:month>
    </swrc:InProceedings>
  </burst:publication>
</item>
```

The publication feed format serves two purposes: firstly, it can be understood by existing Web 2.0 infrastructure, which is capable of processing and visualizing RSS feeds. Secondly, it has the expressive power of RDF to describe publication metadata and to link entities through URIs. The example given contains a minimum set of attributes, especially addressing the "what?", "who?", "where?", and "when?". The available vocabulary is much larger, because the whole SWRC ontology can be used to markup publication metadata.

### **3.2 Publisher Services**

The Publication Feed Publisher Services are a suite of helper services aiding individuals as well as institutions in producing, aggregating, and refining publication feeds. Services are one of the cornerstones of Web 2.0, allowing for loosely coupled systems and simple syndication [3]. The publisher services were designed according to the needs of institutions as described in scenario 1. At the moment there are three services available (via [12]):

1. The *BibTex Converter* translates BibTex to the publication feed format. It takes any BibTex file as input and converts it into a publication feed. Optionally, certain other metadata can be set, e.g. the publisher of the feed.
2. The *Publication Feed Merger* combines two or more publication feeds and ensures that item URIs are unique. If two items have the same URI, but different content, the more recent version prevails. It takes two or more publication feeds as input and provides a single publication feed as output.
3. The *Publication Feed Filter* selects relevant publications from a feed, according to a given taxonomy. It follows the "filter in" approach, which means that all publications containing one or more keywords in the taxonomy are included in the filtered feed. The Publication Feed Filter takes a publication feed and a taxonomy file as input and returns a filtered publication feed.

All publisher services were written in PHP. They are free for everyone to use, and there is no registration or API key required. To help with the orchestration of these services, a DERI Pipes [13] Installation is available at [14], along with a frontend to the BibTex converter [15].

### **3.3 Subscriber Widgets**

The Publication Feed Subscriber Widgets are a suite of widgets for the visualization of and the interaction with publication feeds. They were designed according to the needs of researchers described in scenario 2. Specifically there are two widgets already implemented:

1. The *Publication Feed Integration Widget* was designed as a plugin to the social networking platform system Elgg [16]. It is based on Blogextend [17] and the Simplepie RSS Feed Integrator [18]. The widget allows members of an Elgg platform adding publication feeds to groups. The publications contained in these

feeds can be accessed via a common group blog. As pictured in Figure 1, individual publications are being visualized as blog post entries. Users are able to rate each publication and engage in discussions with each other by posting comments.

2. The *Publication Feed Visualization Widget* is available as a native Elgg widget and in a Wookie [19] version. It visualizes publication feed items in a faceted browser view based on Simile [20]. The faceted browser currently allows for filtering the publication feeds along the dimensions authors, publication years, and keywords, but this could easily be expanded to include other fields contained in the feeds. The filtering mechanisms are complemented with a full text search. Furthermore, a timeline visualization orders publications chronologically and allows users to intuitively browse through them. A tag cloud helps with detecting the most important keywords for a given collection of publications.

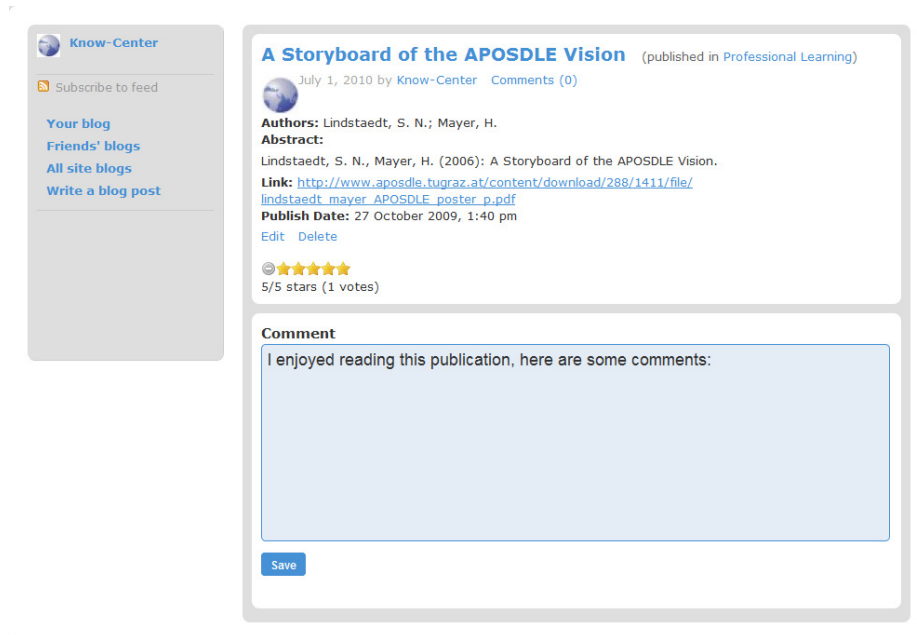


Fig. 1. Rating and commenting features of the Publication Feed Integration Widget

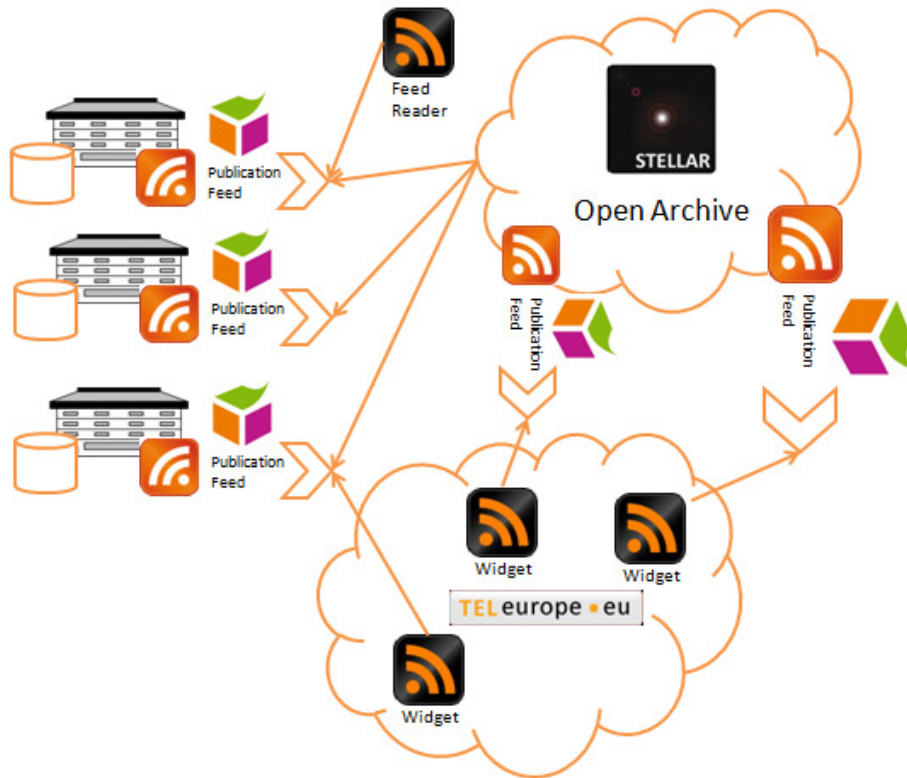
## 4 Integration into the STELLAR Network of Excellence

The publication feed ecosystem is being integrated with the STELLAR Network of Excellence. See Figure 2 for an overview of the proposed concept.

As a first step, all partners within STELLAR are asked to produce a publication feed. In the process, they are able to use the publisher services described in section 3.2 to generate their feeds. The published feeds are in turn being used to update the STELLAR Open Archive (SOA) [21], an open access platform dedicated to collecting

and distributing TEL-related publications as well as the accompanying metadata. Therefore, the SOA subscribes to all of the feeds generated by the partners. The SOA is not only an archive, but it also acts as an aggregator of feeds, allowing to export all or parts of the collected publications as publication feeds. As shown in Figure 2, other tools, which are able to process RSS (such as feed readers) are able to subscribe to the publication feeds as well.

At the same time, the subscriber widgets described in section 3.3 are being deployed to TEL Europe. TEL Europe [22] is a social networking platform based on Elgg for all stakeholders in Technology Enhanced Learning in Europe, operated by STELLAR. With these widgets, users on TEL Europe are able to add relevant publications to a group in subscribing to any publication feed. The feeds might be coming from the SOA, from individual partner institutions, or indeed from any publisher of such a feed (e.g. a special interest group). The members of the group are then able to start a discussion around particular publications, and they may also add a rating. Additionally, they can visualize all feeds available on the platform for search, exploration, and trend scouting.



**Fig. 2.** Overview of the integration of the ecosystem in STELLAR

## **5 Conclusion and Outlook**

In this paper, we presented an ecosystem for the lightweight exchange of publication metadata contributing to the prerequisites for a modern Science 2.0. In two scenarios, we showed how publication feeds can benefit researchers, institutions, and the TEL community. We described the main building blocks of the ecosystem, being (1) the feed format, (2) publisher services, and (3) subscriber widgets. Lastly, we outlined the adoption of the ecosystem by the STELLAR Network of Excellence.

The adoption process has not been finished yet, but the first results are promising. Four partners in STELLAR are actively developing BuRST feeds. Some of them have already been submitted to the STELLAR Open Archive which recently experienced a boost in the number of publications to 1038<sup>6</sup>. The two subscriber widgets have been deployed to TEL Europe and the first special interest groups are starting to use them.

There are certain challenges regarding the publication feed format, which have not been explicitly addressed in the first version. First, the vocabulary of SWRC could be enhanced to include more metadata, e.g. the Digital Object Identifier (DOI) of a publication. Secondly, URIs for authors and institutions would help to manage the entities in the network, and to detect duplicates. URI assignment can either be carried out by the individual institutions or a central repository. With a central repository there is no need to match corresponding entities from various sources, but it also imposes the burden of creating and maintaining said repository.

There are some possible enhancements concerning the existing services and widgets as well. For the Publication Feed Merger, it would make sense to implement a more sophisticated conflict management. This could be done by taking into account the richness of the metadata, as well as the source of information. In the Publication Feed Visualization Widget, additional fields will be added to the existing facets. Furthermore, there is no possibility for end users to correct errors in feed entries. This functionality, however, would rather have to be implemented with a large aggregator of feeds, such as the SOA.

Generally, harvesting and processing of RSS is an open issue. RSS feeds need to be fully retrieved under most circumstances; one is not able to restrict the data to just the new/updated items like in dedicated harvesting protocols, such as OAI-PMH<sup>7</sup>. To overcome this deficiency, we are investigating the integration of the PubSubHubbub protocol [23] into the ecosystem. In the PubSubHubbub protocol, each publisher declares a hub. Subscribers register with that hub, which in turn notifies the subscribers of new and updated items. This avoids repeated polling of the publisher's feed and relieves the subscriber from retrieving the whole feed on update.

Due to its decentralized architecture, the publication feed ecosystem can be extended by anyone. In the future, we expect to see other interested parties contributing their own components. This openness helps making the ecosystem adaptable by other research communities and is a precondition for its sustainable future.

---

<sup>6</sup> On 24/06/2010

<sup>7</sup> Open Archives Initiative - Protocol for Metadata Harvesting



## 6 Acknowledgement

This work was carried out as part of the STELLAR Network of Excellence, which is funded by the European Commission (grant agreement no. 231913). This contribution is partly funded by the Know-Center, which is funded within the Austrian COMET program – Competence Centers for Excellent Technologies – under the auspices of the Austrian Federal Ministry of Transport, Innovation and Technology, the Austrian Federal Ministry of Economy, Family and Youth, and the State of Styria. COMET is managed by the Austrian Research Promotion Agency FFG.

## 7 References

1. Waldrop, M: Science 2.0: Is Open Access Science the Future? *Scientific American*, 5 (298), 46--51 (2008).
2. Kieslinger, B., Lindstaedt, S.N.: Science 2.0 Practices in the Field of Technology Enhanced Learning. Science 2.0 for TEL Workshop, EC-TEL (2009)
3. RDF Site Summary (RSS) 1.0, <http://web.resource.org/rss/1.0/spec>
4. STELLAR: The Network for Technology Enhanced Learning, <http://www.stellarnet.eu/>
5. O'Reilly, T.: What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Online at <http://oreilly.com/web2/archive/what-is-web-20.html> (2005)
6. Mika, P., Klein, M. and Serban, R.: Semantics-based Publication Management using RSS and FOAF. In: *Proceedings of the Poster Track, 4th International Semantic Web Conference*, Galway (2005)
7. Mika, P.: Bibliography Management using RSS Technology (BuRST). Online at <http://www.cs.vu.nl/~pmika/research/burst/BuRST.html> (2005)
8. RDF - Semantic Web Standards, <http://www.w3.org/RDF/>
9. DCMI Metadata Terms, <http://dublincore.org/documents/dcmi-terms>
10. SWRC Ontology v0.3, [http://ontoware.org/swrc/swrc\\_v0.3.owl](http://ontoware.org/swrc/swrc_v0.3.owl)
11. Publication Feeds Format 1.0, [http://www.stellarnet.eu/d/6/3/Publication\\_feeds\\_format\\_v1.0](http://www.stellarnet.eu/d/6/3/Publication_feeds_format_v1.0)
12. Publication Feeds Publisher Services, <http://stellar.know-center.tugraz.at/services>
13. DERI Pipes: Open Source, Extendable, Embeddable Web Data Mashups, <http://pipes.deri.org/>
14. DERI Pipes@Know-Center, <http://stellar.know-center.tugraz.at:8080/pipes/>
15. BibTeX to Publication Feed Converter, <http://stellar.know-center.tugraz.at/html/convert.php>
16. Elgg – Open Source Social Networking Engine, <http://elgg.org/>
17. Blogextended, <http://community.elgg.org/pg/plugins/antifm/read/230708/blogextended-132>
18. SimplePie RSS Feed Integrator, <http://community.elgg.org/pg/plugins/costelloc/read/37480/simplepie-feed-integrator/>
19. Apache Wookie, <http://getwookie.org/>
20. SIMILE Widgets, <http://www.simile-widgets.org/>
21. Stellar Open Archive, <http://oa.stellarnet.eu/>
22. TEL Europe, <http://www.teleurope.eu/>
23. pubsubhubbub, <http://code.google.com/p/pubsubhubbub/>

# More! A Social Discovery Tool for Researchers

Gonzalo Parra, Erik Duval  
Dept. Computer Science  
Katholieke Universiteit Leuven  
Belgium  
{gonzalo.parra, erik.duval}@cs.kuleuven.be

**Abstract:** Science 2.0 is the result of Web 2.0 tools and trends influencing the research area. In this paper, we focus on a scenario where a researcher is interested in the topic and speaker at a conference: finding more information about them is far from instantaneous. Thus, we identified a need to not only find speaker and paper information during the presentation, but also to subscribe to feeds that keep the attendee informed about ongoing work from the speaker. This work presents the development of a mobile application that groups all the relevant information of a speaker in a way that can be easily exposed and integrated in the normal workflow of the audience. The result is a frictionless blending of the face-to-face event that a conference or workshop presentation represents with the rich interaction and alerting services that a web2.0 environment provides.

## Introduction

One of the initial goals of early Internet development was to enable information sharing between researchers. This original aim is now again the focus with much activity around it: “Science 2.0” is the result of “Web 2.0” tools and trends influencing how we carry out research (Shneiderman 2008; Codina 2009; O’Reilly, T. 2005). The effects are visible on how researchers experiment, get feedback on their work, and interact with their community (Waldrop 2008; Reinhardt et al. 2009). Researchers are starting to embrace different types of social tools, in order to pose questions, provide answers, share knowledge, initiate debates, etc. The hope is that the communication and collaboration possibilities offered by these tools can accelerate and improve the way science is being done.

There are several types of social tools actively used by researchers. These can be categorized as follows (Cabezas-Clavijo et al. 2008): scientific blogs, magazines 2.0, reference managers, social taggers, mashups, social networks and sharing. In this paper, we focus on the scenario where a researcher is attending a conference presentation and is interested in the topic and the speaker: finding more information about them is currently done in an ad hoc way; where the researcher either talks to the speaker or uses a search engine to find information about him and his paper, his home page, blog, publications list, etc. A big drawback is that this process is far from instantaneous: the attendee may search during the conference session, or write down that he wants to follow up later on, or make a mental note that he should talk to the presenter and inform his team afterwards. Oftentimes, this follow-up doesn’t happen. And even if it does, it may no longer be as useful as the attendee may no longer be at the conference and it may be difficult to find the relevant information to begin with.

Thus, we identified a need to easily find speaker information during the presentation, and to subscribe to feeds from social tools that keep the attendee informed about ongoing work from the speaker. We have addressed this need through a mobile application, called “More!”. The structure of this paper is as follows: we first present related work that has been done in this area. In the following section, the design and implementation of our tool is covered. An evaluation of the usability and functionality of the tool is presented and its results are analyzed. Finally, we include conclusions and opportunities for further work.

## Related Work

Communication and collaboration between researchers is a key activity in the way we do science. Before researchers can start collaborating, an initial process of discovery must take place. There are several ways on which this discovery process can take place: reading a paper or assisting to a conference and later use a search engine to find more information about his author, or research colleagues can introduce other researchers. Tools like (Noovo 2009;

Mendeley 2009; ResearchGATE 2009; Academia.edu 2009) help a researcher to find people with similar research interests.

On the other hand, as researchers also become more active on web2.0 tools, there are several approaches to consolidate or aggregate the outcome from different social tools. Applications like (FriendFeed 2009) and (Socialthing! 2009) create one checkpoint for updates or trends by aggregating all the activities of colleagues or friends on a variety of web2.0 platforms. In addition to aggregation, it is possible to power social recommendation from web2.0 services (Drachsler et al. 2009).

Some science2.0 tools focus more on the sharing of data sets (Infochimps 2009 ; Myexperiment 2009), bibliographical data (Mendeley 2009 ; AcaWiki 2009 ; BibSonomy 2009 ; Citeulike 2009) or data mashups and visualizations (AuthorMapper 2009 ; Glasser et al. 2009).

Other tools focus on enhancing the experience of participating in scientific gatherings, assisting people with planning the sessions they will attend (PAWS 2009), recording the sessions and making them accessible for later or remote viewing (VideoLectures.net 2009). And some mobile applications tools that can be used at a conference workshop or seminar (Chen 2009 ; Arbogast 2009).

An outstanding example of a discovery process from a different domain is the Shazam music identification service. This mobile application allows identifying a song by just capturing a small sample of it (music fingerprint). If the identification process is successful, the application provides information about the song (artist, title, album) and relevant links to other applications like YouTube and iTunes. Inspired by this application, we have developed “More!” in order to provide a new approach to facilitate discovery and connection among researchers participating in an event.

## **The More! Application**

We have developed an application that group relevant information on a speaker in a way that can be easily exposed and integrated in the normal workflow of the audience. In this context, it is important that the application is mobile, as the intent is that researchers would use it while attending a conference. The design and implementation of the application takes into account software quality attributes (Bass et al. 2003), like extensibility, configurability and portability.

The application is called “More!”, as it provides more information about the speaker. In order to provide the desired portability, which means the ability to run on multiple hardware and software configurations; we decided to create “More!” as a web application. Although it is optimized for viewing on a mobile device, the application can also be used from a regular computer with a web browser.

It is important to consider that the application relies on referring Web2.0 tools, and this kind of tools will keep appearing. That is the reason why the application design considers the extensibility feature for future growth. The social tools will be presented in a list that can increase or be adapted over time, without changing the user interface.

The configurability relates to both extensibility and reusability, as it refers to modifications or configuration changes of components after the deployment. The “More!” application was designed to be a dynamic web application. The data of the speaker and its referred social tools are not a part of the main code of the application and will be linked to an external storage technology (e.g. a database).

Shazam records a fragment of a song in order to create a music fingerprint that identifies a song. The equivalent approach in the “More!” application is Quick Response (QR) codes to represent the speaker fingerprint (ISO/IEC 2006). These codes provide a high capacity for encoding data and readability from any direction in 360° (omni-directional). The QR code encodes a URL that resolves to the speaker page on the More! web application. A full interaction diagram is presented in (Fig. 1).

Finally, we expose the following information for each speaker:

- speaker: full name, photo, e-mail, affiliation and publication list;
- current presentation: slides and paper;
- social tools: Twitter, SlideShare, blog, Delicious and Facebook.

This selection was done considering the provision of the regular information that a researcher expects from a speaker, but also including data about his current work and social tools referring and to his previous, current and future work. The selected social tools include technologies that cover different ways to collaborate and share among researchers. While Twitter and blog exposes the speaker’s thoughts on different levels, SlideShare presents a collection of his previous presentations. Delicious allows the sharing web site interests and Facebook provides a different channel of communication between researchers.

In this way, the attendee may access some personal details about the speaker, as well as the paper and slides of the current presentation. Moreover, he can ‘follow’ the speaker on some of the mainstream Web2.0 social tools, in order to stay informed of new work by the speaker.

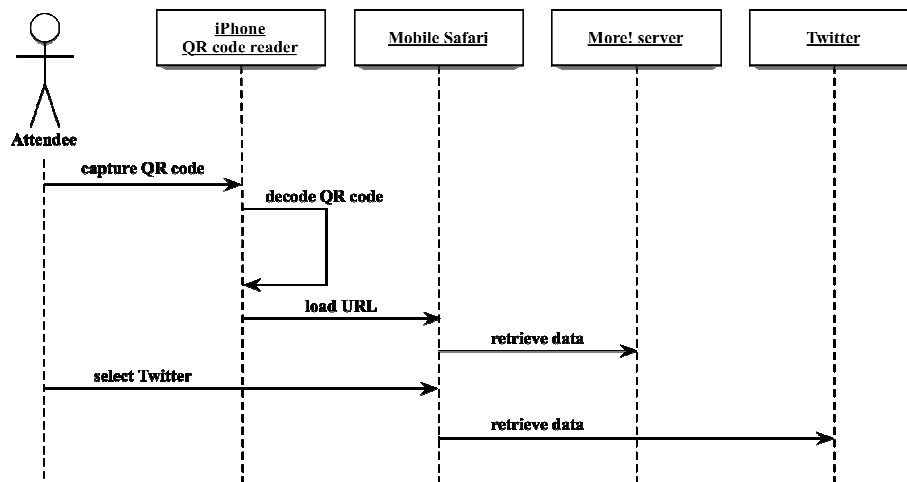


Figure 1. Interaction diagram of QR code usage

## Design process

An initial paper mockup of the application is presented (Fig. 2). There are four distinct areas in the application: general speaker information, current presentation documents, publication list and social tools.

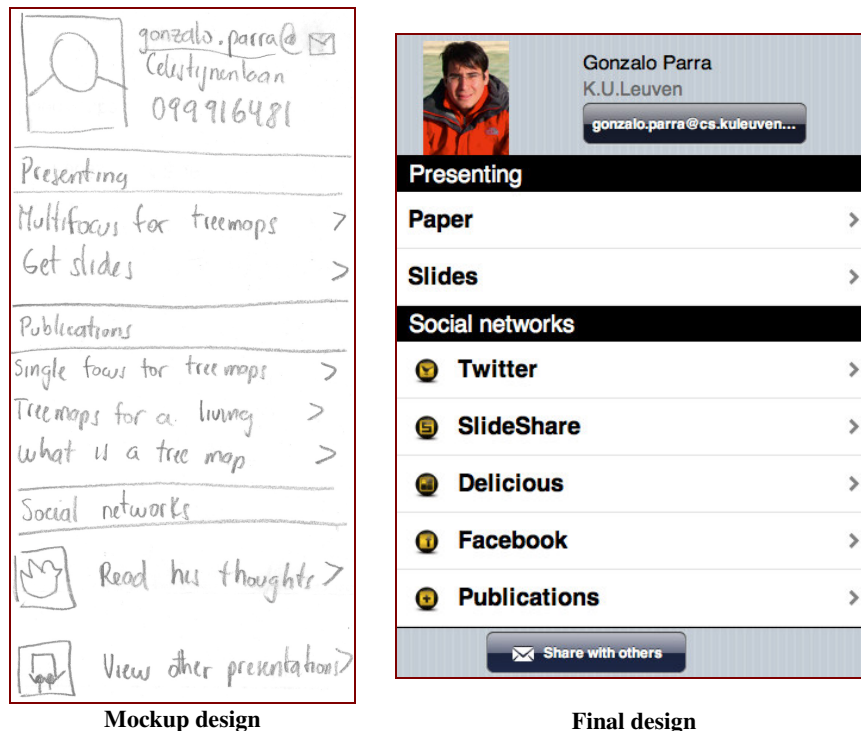


Figure 2. Comparison of the application mockup and final design

This paper mockup was evaluated with three potential end users (researchers) in order to measure the user subjective satisfaction (Rangel De Queiroz et al. 2009). For this initial proof of concept evaluation, we applied the think aloud

protocol (Lewis et al. 1993) in order to capture information about the interface design and the usability of the tool. Categories where the users experienced some difficulties or proposed alternatives to our initial design were summarized in (Tab. 1). These categories include: feature requests, presentation of options, application feedback and missing information. The last category refers to the contents that are currently presented by the application and the additions requested by the test users.

All of the participants requested more features. These requests were: sound feedback, contacts and calendar integration (in the iPhone case), and a integrated chat. All of the participants also requested to expose the information options in a different way. As presented in the figure (Fig. 2), initially the publications were listed before the social tools. This became a problem, as the list of publications could be quite long for a very active speaker. Finally, one participant experienced some minor difficulties regarding the user interaction and feedback of the application.

Table 1. Summary of evaluations			
Category	User 1	User 2	User 3
Feature requests	X	X	X
Presentation of options	X	X	X
Application feedback	X		
Missing information	X		

Regarding the contents of the application, one user suggested to include more information in the application, specifically an extra social tool as (TripIt 2009).

Based on this feedback, we developed a final design that did not include any extra functionality but that did have the presentation of the information re-arranged. The publications were listed after the social tools, in order to avoid a long list that will “hide” the social tools from the user.

The decision to not include extra features was taken in order to not introduce complexity in a tool that should be simple, straightforward and fast to use. The finally implemented design can be seen in (Fig. 2).

The workflow of the application is as follows:

1. The speaker exposes a QR code (that encodes a URL link to the “More!” application, such as <http://ariadne.cs.kuleuven.be/more/gonzalo> for Fig. 3), either on the first slide or on all the slides of his presentation.
2. Attendants capture and decode the QR code with their smart phones and are redirected to the “More!” application. As an alternative, the attendee can also use the aforementioned URL and a regular web browser to load the application
3. “More!” presents the data on the client tool.



Figure 3. QR code example and the encoded “More!” application

## Evaluation

In order to evaluate the impact of the tool on the research community, we have considered two aspects: usability and functionality of the application. The user subjective satisfaction in mobile devices is applied as an evaluation measurement (Rangel De Queiroz et al. 2009). This measurement gives us some feedback regarding the satisfaction level of the user, and highlights problems and its impacts. The data was gathered using 2 methods: questionnaires and unstructured interviews or discussions.

For this purpose, two evaluations took place using the implemented version of the application. The two metrics were evaluated in two different scenarios. The first scenario focuses more on the usability, while the second scenario focus more on the functionality metric.

### Initial evaluation

The usability of the tool was evaluated with 20 university students of Computer Science at the K.U.Leuven. The students were presented with a description of the application and a typical scenario of how it could be applied. Smart phones were provided for the evaluation. The group of students had prior knowledge of social tools, but only basic or no experience with mobile devices. After the students tested the tools, they were requested to fill in an evaluation form covering four aspects: two related to usability and two related to functionality. The results are presented in (Tab. 2). In the form, the values that represent the agreement with a statement ranged from 1 to 5, with 5 the highest agreement value.

Table 2. Results of the initial evaluation		
Evaluated aspect	Agreement Average	Standard deviation
<b>Know more of the speaker's work</b>	4,44	0,51
<b>Easily follow speaker's presentation</b>	4,06	0,64
<b>Easy to use</b>	4,44	0,51
<b>Satisfaction with the tool design</b>	4,28	0,46

As observed, all the participants clearly agreed that “More!” is simple and easy to use, with an average of 4,44 and a standard deviation of 0,51. The participants were a bit less satisfied with the design of the tool and the user experience: the agreement average was only 4,28.

Regarding the functionality, the students were presented with a real world scenario where the tool would be used. The results show that the participants agreed, with an average of 4,44 points, that the tool enabled them to obtain more information about the speaker. Regarding the ability to retrieve more information about the current presentation, the average was the lowest, with a 4,06 points. This was due to the fact that some students were not completely aware of what a conference is in real life, and misunderstood the scenario.

The participants were requested to list functionalities that they think the application should have. The most relevant responses were:

- Bookmarking of speakers for future reference.
- Possibility to synchronize the tool with current slide presented by the speaker.
- Tagging of interesting slides.
- As second option, avoid QR capturing and directly type the speaker's name.

As a final part of the evaluation, the students were requested to list the strengths and weaknesses of the application. The most important strengths they found were:

- use of the graphical code (QR) as the initial fingerprint,
- possibility to explore the slides,
- simple and easy to use.

The major weakness encountered by the participants was the incompatibility of some social tool web sites with the smart phone web browser. To conclude, the major problem identified in this evaluation was related to the devices. These were not used regularly by the students, which made them afraid of the device and the test. Also, one of the devices had no properly configured e-mail account, causing the participant to not experience the full functionality of the tool.

## First evaluation in practice

The second evaluation of the tool focused on the functionality and took place in a real world situation. The 15 participants of a workshop at the Alpine Rendez-Vous 2009 were introduced to the experiment and requested to use the tool (STELLAR 2009). After the workshop, the participants were requested to fill in a questionnaire regarding functionality, and some usability questions. The results are presented in (Tab. 3). As the previous evaluation, the values that represent the agreement with a statement ranged from 1 to 5, with 5 the highest agreement value.

<b>Table 3. Results of the first evaluation in practice</b>		
<b>Evaluated aspect</b>	<b>Agreement Average</b>	<b>Standard deviation</b>
<b>Know more of the speaker's work</b>	4,25	0,50
<b>Easily follow speaker's presentation</b>	3,00	1,15
<b>Easy to use</b>	4,75	0,50
<b>Satisfaction with the tool design</b>	4,00	0,82

As observed, the participants clearly agreed that the application is simple and easy to use, with an average agreement of 4,75 over 5. The participants were a bit less satisfied with the design of the tool: the agreement average was only 4,00. These results are similar to the ones obtained in the previous evaluation.

Regarding the functionality, the participants agreed, with an average of 4,25 points, that the tool enabled them to obtain more information about the speaker. Regarding the ability to retrieve more information about the current presentation, the average was the lowest, with a 3,00 points. This results behavior was also present in the previous evaluation.

The usage of the application was also tracked during the workshop. The preliminary results allow us to conclude that the tool was successfully accepted and used among the workshop participants. During the workshop and the subsequent day, there were 42 visits to the web application. Out of the 42 visits, 19 were unique visitors over the 3 days. In total, there were 97 page views with an average of 4 minutes per visit. Results are presented in (Tab 4.). In the table, the unique visitors count is larger than 19, due to the fact that a unique visitor is considered in a day span, not the previous three days span value.

<b>Table 4. Overview of visits from 29/11 until 02/12</b>				
<b>Date</b>	<b>Visits</b>	<b>Unique visitors</b>	<b>Page views</b>	<b>Time on site (min.)</b>
30-Nov	20	11	53	05:05
01-Dec	15	9	32	04:16
02-Dec	7	6	12	00:37

To evaluate the portability of the application, an overview of the visitors per browser and operating system is presented in Table 5. While Mobile Safari was the most used mobile web browser with almost 53%, regular desktop clients like Firefox, Opera and the desktop version of Safari, were also used. This indicates that avoiding the dependence on a smart phone technology was a correct decision, in order to encourage more participants to use the application.

<b>Table 5. Overview of visits per browser</b>			
<b>Browser</b>	<b>Operating System</b>	<b>Visits</b>	<b>Visits (%)</b>
<b>Safari</b>	iPhone	19	45.24%
	iPod	3	7.14%
	MacOS	10	23.81%
	Android	2	4.76%
<b>Firefox</b>	Windows	3	7.14%
	MacOS	2	4.76%
	Linux	2	4.76%
<b>Opera</b>	MacOS	1	2.38%

## Conclusions and Further Work

There is considerable interest in the research community for Science2.0 tools that can improve the way research is done. The “More!” application provides a mobile on-site discovery tool for researchers, where some personal information, research information and social tools are aggregated and presented to an interested conference attendee. The evaluations of the tool from the Human Computer Interaction point of view show that users agreed on the simplicity of the tool and are fairly satisfied with the components and design. Also in a real life setting evaluation, the users expressed that the tool is simple and easy to use and helps them to know more about the speaker. The web application approach of the tool proved to be useful as attendees used both smart phones browsers and regular web browsers.

Also, different unexpected use cases were obtained after the evaluation. These scenarios are:

- deciding if a presentation is worth attending by quickly checking the presentation slides and paper,
- catching up and following the presentation more easily, even if a part of the talk was missed by an attendee,
- elaborating more and better questions by having the presentation slides at hand.

Even though these are still early evaluations, we are convinced that “More!” can improve the connections between researchers. However, there is plenty of opportunity for further improvement.

- The current approach of manual data gathering is a big entry barrier for the tool. We have done some initial work to automatically extract the information from scientific papers in PDF format with ParsCit (Council et al. 2008). In this way, we obtain relations between authors, e-mails and affiliations. Later, a text distance metric introduced by (Levenshtein 1966) is applied to match duplicated results, as the main problems are related with duplicate authors, e-mail addresses, and the notation of affiliations. Also, the quality and scope of the data obtained by this method can be improved by using tools like DBLP and Linked Data approaches, or by direct contacts with the publishers in order to increase the coverage of the database. As a possible extension, the evaluation participants expressed the need to manipulate the exposed the application’s data. This modify/update feature could be considered in the near future.
- In a more general sense, there is a need to collect the type of information that “More!” builds on, in a scalable and reliable way. A specific researcher oriented web crawler or approaches like the ones discussed in (WEPS 2009) or used in person search engines like (Spock 2009) could help to automate this process. The more information about someone is consulted through “More!”, the more relevant it probably is to also include these people in his network.
- We need to better understand how mobile applications like “More!” can enhance the experience of researchers at events like seminars, workshops and conferences. Can such tools increase the awareness about related work, or even collaboration between researchers? And will that help the research community to perform in a more effective and efficient way? This is a vast area of questions and concerns that is at the core of “Science2.0”. We are currently planning the second evaluation in practice of “More!” to further explore these questions and translate them into measurable characteristics.

In any case, we are very excited about the potential of applications like “More!” and the evolution towards ‘Science2.0’.

## References

- Academia.edu. (2009). <http://academia.edu>. Last retrieved: December 2009.
- AcaWiki. (2009). <http://acawiki.org>. Last retrieved: December 2009.
- Aercs. (2009). <http://bosch.informatik.rwth-aachen.de:5080/AERCS/>. Last retrieved: December 2009.
- Arbogast, S. (2009). ConferenceGuide. <http://www.conferencegui.de/>. Last retrieved: December 2009.
- AuthorMapper. (2009). <http://www.authormapper.com>. Last retrieved: December 2009.
- Bass, L., Clements, P., and Kazman, R. (2003). Software Architecture in Practice, Second Edition. Addison Wesley.
- BibSonomy. (2009). <http://www.bibsonomy.org>. Last retrieved: December 2009.



Cabezas-Clavijo, A., Torres Salinas, D., and Delgado López-Cozar, E. (2009). Ciencia 2.0: catálogo de herramientas e implicaciones para la actividad investigadora. *Profesional de la información*, El, 18(1):72-80+.

Chen, Y. (2009). Academic Conference Finder. <http://itunes.com/apps/academicconferencefinder>. Last Retrieved: December 2009.

Citeulike. (2009). Last retrieved: <http://www.citeulike.org>. December 2009.

Codina, L. (2009). Ciencia 2.0: Redes sociales y aplicaciones en línea para académicos. *Hipertext.net*, (7). 295.

Councill, I. G., Lee Giles, C. and Kan M. (2008). ParsCit: An open-source CRF reference string parsing package. *Language Resources and Evaluation*, 2008.

Drachsler, H., Pecceu, D., Arts, T., Hutten, E., Rutledge, L., Van Rosmalen, P., Hummel, H. G. K., & Koper, R. (2009). ReMashed - Recommendations for Mash-Up Personal Learning Environments. In U. Cress, V. Dimitrova, & M. Specht (Eds.), *Learning in the Synergy of Multiple Disciplines, EC-TEL 2009* (Vol. LNCS 5794). Nice, France: Springer.

FriendFeed. (2009). <http://friendfeed.com>. Last retrieved: December 2009.

Glaser, H., Millard, I. and Carr, L. (2009) RKBExplorer: Repositories, Linked Data and Research Support. In: Eprints User Group, Open Repositories 2009, 20/05/2009, Atlanta, GA, USA.

Infochimps. (2009). <http://infochimps.org>. Last retrieved: December 2009.

ISO/IEC. (2006). ISO/IEC 18004:2006, Information technology - Automatic identification and data capture techniques - QR Code 2005 bar code symbology specification.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*. Technical Report 8.

Lewis, C, Rieman, J., and Blustein, A. (1993). *Task-Centered User Interface Design: A Practical Introduction*.

Mendeley. (2009). <http://www.mendeley.com>. Last retrieved: December 2009.

Myexperiment. (2009). <http://www.myexperiment.org>. Last retrieved: December 2009.

Noovo. (2009). <http://www.noovo.com/>. Last retrieved: December 2009.

O'Reilly, T. (2005). What is web 2.0 – design patterns and business models for the next generation of software. <http://oreilly.com/web2/archive/what-is-web-20.html>. Last retrieved December 2009.

PAWS. (2009). <http://ir.exp.sis.pitt.edu/ks2/DL/AnnotatEd/ELearn/index.cgi>. Last Retrieved: December 2009.

Rangel De Queiroz, J. E., and Sousa Ferreira, D. (2009). A Multidimensional Approach for the Evaluation of Mobile Application User Interfaces. *Human-Computer Interaction International 2009*, 242-251.

Reinhardt, W., Ebner, M., Beham, G., and Costa, C. (2009). How people are using twitter during conferences. *Education and Media*. 145–156.

ResearchGATE. (2009). <http://www.researchgate.net>. Last retrieved: December 2009.

Shneiderman, B. (2008). Computer science: Science 2.0. *Science*, 319(5868). 1349-1350.

Socialthing!. (2009). <http://socialthing.com>. Last retrieved: December 2009.

Spock. (2009). <http://www.spock.com/>. Last retrieved: December 2009.

STELLAR. (2009). <http://www.stellarnet.eu/programme/wp3/rendez-vous/>. Last retrieved: December 2009.

Tripit. (2009). <http://www.tripit.com/>. Last Retrieved: December 2009.

VideoLectures.net. (2009). <http://videlectures.net/>. Last retrieved: December 2009.

Waldrop, M. M. (2008). Science 2.0 - is open access science the future?. Scientific american.  
<http://www.sciam.com/article.cfm?id=science-2-point-0>. Last retrieved: December 2009.

WEPS. (2009). Searching Information about entities in the Web. <http://nlp.uned.es/weps/>. Last Retrieved: December 2009.

### **Acknowledgements**

We gratefully acknowledge the support of the STELLAR Network of Excellence on Technology-Enhanced Learning and the feedback and comments from the STELLAR participants as well as our students.

---

# Research at the table

Bram Vandeputte and Erik Duval

Katholieke Universiteit Leuven, Department of Computer Science  
Celestijnenlaan 200a, 3001 Leuven, Belgium  
{bram.vandeputte,erik.duval}@cs.kuleuven.be

**Abstract** In this paper we describe how we want to take advantage of the rapid developments in technology to assist researchers in doing research. More specifically in exploring the publication space. For this purpose we have designed and developed a prototype application to take advantage of large displays with multi touch enabled input. We describe the current state, the next steps and how we will to evaluate it. To conclude we give an outlook on further possibilities and challenges that lay ahead.

**Key words:** research2.0, information visualization, multi touch, large display, research.fm

## 1 Introduction

How great would it be to integrate the process of exploring publications, finding them and reading them in an almost seamless way ? This sort of idea was already described in *As we may think* in 1945 by Vannevar Bush [2]. The Memex was described as the perfect desk of a researcher, having all the knowledge of the world readily available. At that time personal computers were not even invented, but since then technology has advanced tremendously and become very common. Using current state of the art technologies, we want to find out how we can ease the process of exploring publications. This process is an important part of a researcher's job, as he wants to know what is going on in his field of research.

To be able to get this kind of understanding, Russell et al [10] have pointed out that it is imperative that the right representation is found for exploring a network of (publication) data.

The idea of visualizing publication networks has been inspired by the work of Klerkx et al, where they explore learning object repositories [8] and social bookmarks [7] in a visual manner.

In this paper we first introduce and describe the problem. We then motivate our hardware platform, describe the origin of the data and we explain the detailed workings of the application. In the next section we compare our work with existing studies. Then we describe how we evaluate this and finally we propose the next steps to be taken. To conclude we summarize our findings and discuss further possibilities.

## 2 Problem statement

An important part of a researcher's job is reading scientific papers. This ensures that the researcher is up to speed of what is going on in his research field. It is also a prerequisite for writing scientific papers, as handbooks such as the one from Robert A. Day [3] emphasizes.

There are three basic ways of dealing with scientific papers. There is active search, where you search for a particular paper or a 'good' paper on a specific topic you have in mind. There are dozens of websites that serve this purpose really well, such as Google Scholar<sup>1</sup>, ISI web of Knowledge<sup>2</sup>, DBLP<sup>3</sup> etc. There is also what we can call passive search, where you get alerted whenever new publication material is available. Google Scholar has recently added a feature where you can be alerted whenever something new comes up that matches certain keywords. Also many of the journal magazines let you subscribe to a list to send you the table of contents when a new issue is available. Finally one can focus on relations between papers and authors. There are existing tools where this is possible, but we think that there is not enough technical support available for exploring these networks.

To explain the problem we want to solve, we will briefly describe the use cases we want to tackle with this work. The use cases can be grouped into two categories. In the first category the use cases have a mainly *top down* approach, while the second category holds the use cases that typically need a *bottom up* approach.

### 2.1 World overview

Typically, in this use case a user would like to start with a complete overview of all nodes laid out in a graph. The user then wants to zoom in on parts of the graph that draw her attention. This can be used to find out patterns or clusters. In this case the user usually is already an expert in the field, trying to understand or improve his knowledge about the field.

### 2.2 Explore your neighbors

In this case you might want to start from a view with a focus on yourself, or the author or paper that you want to start from. Then you want to browse to nodes in your 'neighborhood', which are likely to be related and/or interesting. Here you can try to find answers to questions like : Where am I in the research publication space ? Who should I talk or connect to ?

---

<sup>1</sup> <http://scholar.google.com>

<sup>2</sup> <http://apps.isiknowledge.com>

<sup>3</sup> <http://www.informatik.uni-trier.de/~ley/db/>

### 3 The application

#### 3.1 The hardware

**The input modalities** We chose for supporting a multi-touch setting, as we want to explore direct and multi touch capabilities. This to find out whether these relatively new input methods can help to make it easier for researchers to interact with the fairly complex graph like structures.

**The display** The application will entail a visualization of a deeply connected network containing up to hundreds (maybe thousands) of nodes. This property feeds the need for using a large display. These large displays, with increasingly higher resolutions, are also rapidly becoming cheaper and more common, which makes it easier to include them in our study and makes this study more relevant.

A problem that sometimes arises on multi touch input devices is when one touches the screen to give input, the finger or hand occludes information one wants to see at that moment. This can be solved in two ways, either we make the information appear next to the touch point, or we make the information bigger so it is less likely to be occluded. Both solutions can benefit from a larger display, as you have simply more space to put the information.

Studies by Forlines et al [5] and Kin et al [6] have already shown that on tabletop displays multi touch input has performance and spatial awareness advantages over the traditional mouse, which reinforces our choice of hardware. From a research perspective, we want to explore if and how a large screen estate can influence the possibilities of this kind of visualization.

#### 3.2 The data

**EC-TEL conference** Our first scope was to visualize all the publications from all editions of one conference. We extracted metadata from papers and put them in a database. Unfortunately this extraction process is still very error prone and a lot of semi-manual cleaning up needed to be done. The approach took quite a bit of effort and is not very scalable.

To try and make access to these publication data easier, we propose an open architecture for exchanging these publication metadata. This architecture is currently being discussed and developed in the STELLAR project<sup>4</sup>, with both suggestions for collecting these data using BuRST feeds<sup>5</sup> and a webservice API, called research.fm<sup>6</sup>, to make them available for tools and widgets like the one we are describing in this paper.

---

<sup>4</sup> <http://www.stellarnet.eu/>

<sup>5</sup> [http://stellarnet.eu/d/6/3/BuRST\\_format\\_adaption\\_discussion](http://stellarnet.eu/d/6/3/BuRST_format_adaption_discussion)

<sup>6</sup> <http://www.stellarnet.eu/d/6/3/KULDocumentation>

### 3.3 The network and the visualization

The obvious relations to visualize are the paper-author relations, and also co-authorship. To build up this network, we want to have a self-organizing and self-declustering algorithm. We chose to use Traer Physics<sup>7</sup>, an implementation of a simple particle system physics engine, which allows to combine a spring-graph algorithm with physical forces. This combination will take care of the organizing and declustering of the network, so we don't have to care about where to put the nodes. After experimenting with the parameters such as force, drag, mass of the particles, spring length and strength, ... We could see a clear network-like graph appearing when the network is stabilizing after a few seconds.

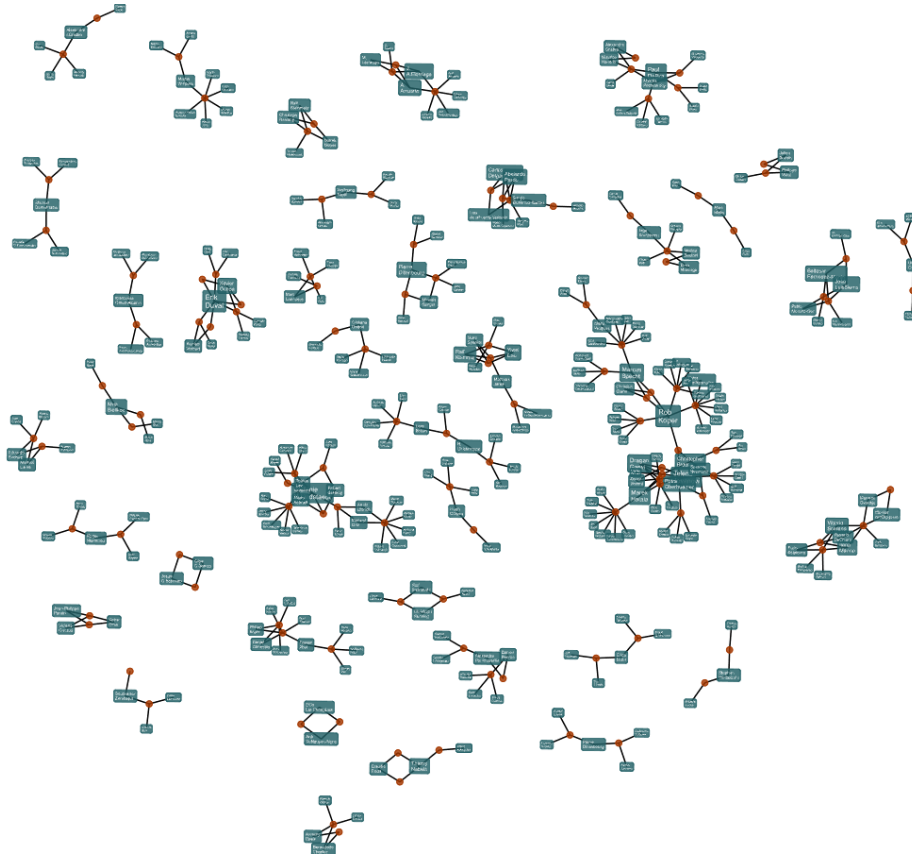


Figure 1: Overview of the whole publication network. The green nodes are authors, while the red ones are papers.

<sup>7</sup> <http://www.cs.princeton.edu/~traer/physics/>

Figure 1 shows a screenshot of the visualization in the overview state. All nodes present in the network are shown. This state addresses the first use case we described in section 2.1. It can help a researcher to find out whether there is a lot of collaboration going on in this field, where the biggest clusters can be found or who the most active authors are.

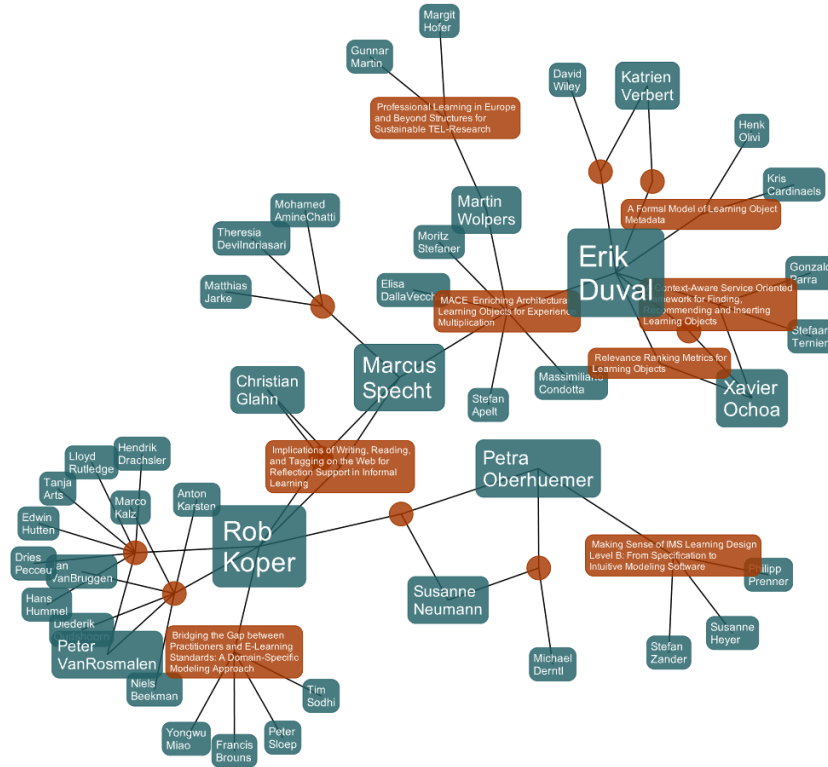


Figure 2: A detailed view of related authors. The green nodes are authors, the bigger they are, the more papers they have published. The red nodes are papers, where some of them have been expanded to show the title of the paper.

The second use case described in section 2.2 benefits from the view as shown in Figure 2. Here the visualization is zoomed in on a specific target. All the author names become clearly visible, so you can find an author very relevant for your work. One can also click on some paper nodes to get more information on the paper itself, so to find papers that are interesting, for example because they are closely related to your work. As you can see we are already experimenting with varying the node size of the author, based on his number of publications, to denote importance of this author.

## 4 Related Work

There are numerous other visualizations of publication data existing already. In this section we will highlight some visualizations that try to solve similar problems, and we will shortly describe how each of them differ from our approach.

### 4.1 Papercube

When this web application<sup>8</sup> first opens up, it immediately shows you a search box. This is useful when you are looking for something more specific, but it does not help when you want to explore the publication space and don't have a specific entry point in mind. There are quite some possibilities both in terms of relations and type of visualizations, so it can take a while for someone to get used to the interface and find what one actually wants. In our approach on the other hand, we want to make it easy for starting the exploration phase by directly showing the data. In this visualization the data is shown in a spring graph with a good lay-out. When you hover over a paper, the relations to other papers are highlighted, which is very helpful. One can also directly click through to the paper itself, so if you have found an interesting publication you can directly retrieve it online. Bergström et al [1] evaluated this application, and found that the users unanimously said that this kind of visualization can usefully augment existing digital libraries.

### 4.2 Ed-Media Relation Browser

The Ed-Media Relation Browser<sup>9</sup> is also an interactive, browser based, author visualization. In this approach they focus on one person and its direct relations, assisted with a strong filtering mechanism. The visualization only starts after you have entered a name. This emphasizes their focus on solving the problem of getting to know closely related authors. It does not allow one to study the field nor to discover the indirect relations between authors and papers. In our approach we try to solve this problem by allowing to zoom in on a specific person, but with a global navigation strategy so that the overview does not get lost. This visualization does not allow to rearrange the graph. To help the spatial memory we allow the user to organize the papers and authors however he likes. The authors, Ochoa et al [9], have also studied the complete publication space of a conference, but only with non interactive visualizations, where we allow to do so with a highly interactive visualization.

### 4.3 Microsoft Academia Search Visual Explorer

The Microsoft Academia Explorer<sup>10</sup> is similar to the Edmedia Relation Browser. Here you can drag the authors around to get a better view if something is not

---

<sup>8</sup> <http://papercube.peterbergstrom.com>

<sup>9</sup> <http://ariadne.cs.kuleuven.be/edmedia/>

<sup>10</sup> <http://academic.research.microsoft.com/VisualExplorer.aspx>



clear. Once you click on an other author, the graph keeps the link with the previous author but unfortunately all not directly related authors get thrown away. Thus also this visualization only displays direct relations. This application is also only targeted at visualizing authors. One can click through to see all the details of an author, but it is not possible to see the publication which make authors related. Our approach makes the transition from exploration to reading papers easier by bringing the papers visually in the network. If a paper draws attention, one can immediately retrieve more information from it.

## 5 How to evaluate ?

Due to the early stages of this work, there has not been any evaluation yet, but we are planning to do a complete evaluation and here we outline how we will approach this. The evaluation would be done on two levels :

**Macro level** We will introduce the test subjects to the application, explain them the purpose, how it works and what are its functionalities. On this level we want to get answers to questions like : Is this application useful ? Does it address an actual need ? And if so, are the people aware of the existing need ?

**Micro level** In another evaluation, we focus more on the micro level. We want to know if the application is usable, which functionalities and features work well and which do not. In this evaluation the subjects would get specific tasks and we would then record how and how fast these tasks are completed. The specific tasks are not defined yet, but one example could be : Find the most interesting paper written by author x.

**Public spaces** In order to get more feedback, we also plan to deploy this visualization at one or more conferences, where we can observe the people discovering the tool and see what the initial thoughts are.

## 6 Future work

At the time of writing, a first working version of the application has been developed with some basic functionalities. But before we can do a real evaluation of this visualization, we need to improve the functionality of the application. In this section we describe the next steps that will be taken to achieve this.

An important feature that is missing at the moment, is being able to search for a certain author or paper to use as a starting point for the visual exploration. At a first stage we will add a keyboard like possibility to enter part of an author or a paper. To show the results there are several options that can be tried out. The found results can be highlighted in some way, or once a single result is found the visualization can center the result and zoom in on it.

At the moment it is not very visually clear yet which papers or authors are the most important or the most relevant. We are already exploring the possibilities to improve this by trying out filtering mechanisms and visual improvements. These visual improvements can be highlighting certain nodes or areas, varying the size of the nodes based on these factors, varying the strengths of the connections, etc.

## 7 Conclusion

In general, the fundamental issue is to understand in a deeper way how we can support the work of researchers with the technology that is available and how we can evaluate that our efforts make a difference. The design based research presented in this paper tries to move that agenda forward.

A major problem we face is getting clean data. At the moment this is too hard: we had to invest considerable effort in extracting the bibliographical data from the PDF version of the papers and in manually cleaning up the result. Initiatives like DBLP<sup>11</sup>, Citeseer<sup>12</sup>, bibsonomy<sup>13</sup>, citeUlike<sup>14</sup> and others are targeting the same issue and we need to leverage their results in the context of our research.fm framework (see section 3.2) to create sustainable and scalable services for basic bibliographical data provision.

Assisting the user with navigation through the publication space is crucial. It is hard to figure out the correct way to combine navigation and search for manipulation of this information space. Currently, we only provide navigational access and we need to augment this with search facilities to locate relevant locations in this space: these can be papers or authors or relationships between them. We also need to add filtering facilities to reduce the complexity and size of this space to only that part that is relevant to the information need of an author.

We only use a fraction of the available metadata at the moment: our current visualization focuses on (co-)authorship relations between authors and papers. There is plenty of opportunity to also include other kinds of metadata in our scope: this could include forward and backward citations, geospatial information about the affiliations of the authors, textual relationships based on concept extraction techniques, etc. Assessing which kinds of such data help to address which kinds of problems researchers face and how we can exploit the data to make them useful and usable to that audience is a deep design challenge.

Finally, we do not exploit time information yet. However, especially as we start adding more of the metadata to our visualization, this will become an important concern. If we are able to integrate time information, then we can help users understand how a domain or publication outlet (conference, journal,

---

<sup>11</sup> <http://www.informatik.uni-trier.de/~ley/db/>

<sup>12</sup> <http://citeseer.ist.psu.edu/>

<sup>13</sup> <http://www.bibsonomy.org/>

<sup>14</sup> <http://www.citeulike.org/>

...) evolves, how a paper gains in influence, how the collaborative relationships between authors evolve, etc.

**Acknowledgements** We gratefully acknowledge the financial support of the European Commission through the STELLAR project and the STELLAR Network of Excellence.

## References

1. P. Bergström and D.C. Atkinson. Augmenting the Exploration of Digital Libraries with Web-Based Visualizations. In *Fourth International Conference on Digital Information Management, 2009. ICDIM 2009*, pages 1–7. Citeseer, 2009.
2. V. Bush and A.W.M. Think. The Atlantic Monthly. *As We May Think*, 176(1):101–108, 1945.
3. R.A. Day. How to write and publish scientific papers. *Memórias do Instituto Oswaldo Cruz*, 93, 1998.
4. T. Dwyer, B. Lee, D. Fisher, K.I. Quinn, P. Isenberg, G. Robertson, and C. North. A Comparison of User-Generated and Automatic Graph Layouts. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):961–968, 2009.
5. Clifton Forlines, Daniel Wigdor, Chia Shen, and Ravin Balakrishnan. Direct-touch vs. mouse input for tabletop displays. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 647–656, New York, NY, USA, 2007. ACM.
6. Kenrick Kin, Maneesh Agrawala, and Tony DeRose. Determining the benefits of direct-touch, bimanual, and multifinger input on a multitouch workstation. In *GI '09: Proceedings of Graphics Interface 2009*, pages 119–124, Toronto, Ont., Canada, Canada, 2009. Canadian Information Processing Society.
7. J. Klerkx and E. Duval. Visualizing Social Bookmarks. In *conference Workshop on Social Information Retrieval for Technology-Enhanced Learning (SIRTEL07) at the EC-TEL conference. September*, pages 17–20. Citeseer, 2007.
8. J. Klerkx, E. Duval, and M. Meire. Using information visualization for accessing learning object repositories. In *International Conference on Information Visualisation*, volume 8, pages 465–470, 2004.
9. Xavier Ochoa, Gonzalo Mendez, and Erik Duval. Who we are: Analysis of 10 years of the ed-media conference. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications ED-Media 2009*, pages 189–200, 2009.
10. Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. The cost structure of sensemaking. In *CHI '93: Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*, pages 269–276, New York, NY, USA, 1993. ACM.

---

# Muse: Visualizing the origins and connections of institutions based on co-authorship of publications

Till Nagel<sup>1,2</sup>, Erik Duval<sup>1</sup>

<sup>1</sup> Dept. Computerwetenschappen, Katholieke Universiteit Leuven, Celestijnenlaan 200A, 3001 Heverlee, Belgium

<sup>2</sup> Interaction Design Lab, University of Applied Sciences Potsdam, Pappelallee 8-9, 14469 Potsdam, Germany  
nagel@fh-potsdam.de, erik.duval@cs.kuleuven.be

**Abstract.** This paper introduces Muse, an interactive visualization of publications to explore the collaborations between institutions. For this, the data on co-authorship is utilized, as these signify an existing level of collaboration. The affiliations of authors are geo-located, resulting in relations not only among institutions, but also between regions and countries. We explain our ideas behind the visualization and the interactions, and briefly describe the data processing and the implementation of the working prototype. The prototype focuses on a visualization for large tabletop displays, enabling multiple users to explore their personal networks, as well as emerging patterns in shared networks within a collaborative public setting. For the prototype we used the publication data of the EC-TEL conference.

**Keywords:** geo-visualization, tabletop, research, human computer interaction

## 1 Introduction

There has been vast amount of research in the areas of bibliometry and scientrometry to extract and specify the metrics of scientific publication and citation networks. Several works used approaches to visualize these networks (e.g. [1], [2]). In the field of TEL, [3] analyzed and visualized ED-Media publications.

The objective of the presented visualization is not to study individuals and their personal co-authorship networks, but rather to enable analyzing the connection network of universities and research centers. The inter-institutional relationships are based on co-author data, as “co-authorship seems to reflect research collaboration between institutions, regions, and countries in an adequate manner” [4].

Our intention is to focus attention on the spatial relations by creating an easy-to-understand geo-visualization with an emphasis on affiliations and collaborations between these institutions. Studies have shown geographic proximity is important and does positively influence the intensity and frequency of scientific collaboration [5]. However, there has been little research on using geo-visualization for inter-institutional and inter-country collaboration based on publication data (e.g. [6]).

This work focuses on an interactive geo-visualization on large display, enabling multiple users to explore the networks of their affiliations, as well as emerging patterns in shared networks within a collaborative public setting.

We envision several use cases for the application, from which we briefly describe three, exemplarily. (1) A visitor wants to get an overview of the spatial characteristics of scientific collaboration. He starts exploring the institutions and their locations, with the application showing the number of co-authored publications over the years. This visualization supports him understanding whether there is a correlation between proximity and the amount of collaboration. (2) An attendee is interested in finding future partners for writing a proposal. She sees that a colleague from her institution once co-authored a paper with someone from a university department in her field. She writes down the author's name, to later ask her colleague to introduce her. (3) Two persons stand at the table and both are exploring their own affiliations. The application highlights the respective publications, thus enabling them to see shared publications of colleagues, by serendipity. They start talking about these former projects, and find out they have mutual research interests.

The paper introduces *Muse*<sup>1</sup>, a working prototype, whose main purpose is to ease the exploration of collaborations between institutions. In addition, the use of a large display tabletop, as well as the aimed-for simplicity of visualization and interaction intend to invite attendees to participate, and engage in discussions at a conference location. The following chapter gives a short overview on the data set. A description of the prototypes' visualizations and interactions follows. The paper closes with short conclusions and comments on future work.

## 2 Data Set

We are using the EC-TEL dataset as first illustration to show the connectivity in the scientific TEL community. With a young conference as EC-TEL we will not be able to show long-term transformations. Instead, here our aim is twofold: Showing how a striving conference evolved over recent years, and enabling attendees to explore their scientific neighborhoods in the TEL domain.

We harvested the publication data from the website of Springer, the proceedings publisher. We used Web-Harvest [7] to collect all titles, authors, and affiliations including their postal addresses (as well as further data). As the data originally is provided by the authors, using various languages, formats, and accuracies of data, we needed to apply different aggregation and unification heuristics, trying to reduce unintentional duplicates or other skewed data entries. First, the affiliation line is split up into the affiliation's name and its address, to allow a better unification of affiliations, and to display a shorter and more readable name in the visualization. The simplistic, language agnostic approach was to concatenate all text segments up to and including the last segment containing one of a set of specific keywords, selected for high probability of matching institutional name segments (e.g. "universi",

---

<sup>1</sup> The name of the application was chosen to reflect the meaning of "to look thoughtfully at". Secondly, Muse, the greek goddess, presides over literature and science.

“a[clk]adem”). Second, the affiliations were to be unified based on the similarity of the name<sup>2</sup>. After geo-coding the addresses, we also incorporated the spatial proximity to ensure not unifying institutions with very similar names but different locations, e.g. “Dept. of Preventive Medicine, Korea University, South Korea” and “Dept. of Preventive Medicine, Konkuk University, South Korea”.

Generally, it is difficult to structure real-world objects in a way to map all possibilities and special cases, thus we utilized a good-enough approach. Before realizing the prototype we probed into the data and looked for patterns to establish the visualization will be able to reflect those inherent relationships. Some of our analysis for the EC-TEL conferences 2006-2009 can be found at [8].

### 3 Prototype

We designed two working prototypes, with an iterative development approach to refine the visualizations, and to increase the usability of the interactions. The first interactive visualization was presented at the Science2.0 for TEL workshop at EC-TEL 2009. The presentation, and the public display at the venue thereafter allowed us to gather informal responses of attendees. We tried incorporating the given feedback into the second version, and aimed for improving the clarity of the visualization and the overall user experience in an on-location conference setting.



**Fig. 1.** Screenshot of first prototype with Germany and 2009 as selected country and year.

The first application consists of a static world map showing institutions as colored circles with its overall publication number mapped as size (see Fig. 1). Several further visualizations are in juxtaposition: An overview list shows the names of the countries with contributing authors, with a small *sparkline* [9] signifying the absolute

<sup>2</sup> This simplistic approach results in some false positives (e.g. “Av. Universidad 30”), which are recognized as part of the name, and some false negatives (e.g. “ETH Zürich”), which are regarded as part of the address. Furthermore, some entries could not be unified automatically, such as “Lehrstuhl Informatik V” with “Informatik 5 (Information Systems)”.

publications over the years. The concentric rings represent the relative distribution of publications of every participating country over the years, starting from older (inner) up to the latest conference (outer).

These multiple displays are connected, and every user interaction is reflected in all other views. After selecting a country in one of the displays the application provides details-on-demand on that country, and its respective publications and institutions in simple bar diagrams. When the user selects a year the publications are filtered to highlight the data of that specific conference (i.e. as yellow circles and bars).

While the multiple displays allowed looking into the dataset from different perspectives, they also tended to clutter the screen. To effectively communicate the data in a concise visual manner some of the useful but distracting displays have been eliminated in the second prototype. The main improvements were to reduce the visual and interaction complexity by focusing on one main visualization, and the employment of an interactive tabletop with the aim to facilitate multi-user scenarios.

With the large interactive surface, the user not only views and manipulates data on a single user system, but operates in a collaboratively created and used information space (see Fig. 2). In this setting, co-located users, who may or may not be associated with each other, explore the visualization together. Users can arrive or leave at any time, and have the ability to interact as an individual, or as a member of a group with similar interests, goals or attitudes. Cooperative interaction can involve periods of tightly coupled activities by groups with similar but diverging goals, alternated with more loosely coupled individual work. Such collaborative threads can close, split off and merge repeatedly.



**Fig. 2.** Users exploring institutions with the tabletop prototype.

A single large world map showing all institutions and their relations based on co-authorship are displayed. The user is able to select the region she is interested in by panning and zooming the map (while in the first prototype a user only could switch between World and Europe). Even though more complex map manipulations are possible, we chose this interaction approach, as by reducing the prototype to a single visualization the user can concentrate on the map, thus lessening her efforts. The user can select a country she is interested in. That country is selected, and additional

information and diagrams are shown, similar as in the first prototype. These info-windows can be moved to any point on the table. When two countries are selected the prototype displays the diagrams besides each other, allowing the user to compare them.

## **4 Conclusion and Future Work**

Although we have utilized only a small dataset with a rather small significance for general scientific network analysis, we see the Muse prototype with the used data set as beneficial case study. Through interactive filtering the user is able to explore the temporal as well as spatial relations between institutions, and can gather insights into the conference. The collaborative usage of the interactive tabletop display fosters communication among participants.

We intend to broaden the data set to other conferences. Currently, we see two possibilities: Besides using the harvesting tool to scrape further publications from Springer and other official sources, we plan to integrate publication data services, such as pub.fm [10]. Second, querying Web2.0 applications such as Mendeley [11] to gather social network data of the authors.

Furthermore, we are planning an evaluation on intelligibility of the visualization, and usability of the interactions. As direct response from users in a real-world setting can be worthwhile, we intend to create a brief questionnaire to gather feedback from attendees at the EC-TEL 2010.

## **References**

1. Ponds, R., Van Oort, F. G., and Frenken, K. The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 86(3), 423–443. (2007)
2. Henry, N., Goodell, H., Elmqvist, N., Fekete, J.-D.: 20 Years of Four HCI Conferences: A Visual Exploration. *International Journal of Human-Computer Interaction*, 23:3, 239-285. (2007)
3. Ochoa, X., Méndez, G., Duval, E. Who We Are: Analysis of 10 Years of the ED-MEDIA Conference. In: G. Siemens and C. Fulford (eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009*, 189-200. Chesapeake, VA: AACE. (2009)
4. Glänzel, W., Schubert, A.: Analyzing Scientific Networks through Co-authorship. In: H.F. Moed et al. (eds.), *Handbook of Quantitative Science and Technology Research*, 257-276. Kluwer, Dordrecht (2004)
5. Katz, J. S. Geographical Proximity and Scientific Collaboration. *Scientometrics* 31 (1), 31–43. (1994)
6. Cyberinfrastructure for Network Science Center. Research Collaborations by the Chinese Academy of Sciences. [http://scimaps.org/maps/map/research\\_collaborati\\_110/](http://scimaps.org/maps/map/research_collaborati_110/) (2009)
7. Web-Harvest Project. <http://web-harvest.sourceforge.net/>
8. Nagel, T. <http://tillnagel.com/2010/09/ectel/>
9. Tufte, E. *Beautiful Evidence*. Graphics Press. (2006)
10. Stellar project. <http://www.stellarnet.eu/d/6/3/KULDocumentation>
11. Mendeley. <http://www.mendeley.com/oapi/>



# Who are you working with?

## - Visualizing TEL Research Communities -

Marco Fisichella Eelco Herder Ivana Marenzi Wolfgang Nejdl  
L3S Research Center  
Hannover, Germany  
{fisichella, marenzi, herder, nejdl}@L3S.de

**Abstract:** Author Co-Citation Analysis (ACA) provides a principled way of analyzing research communities based on how often authors are cited together in scientific publications. In this paper, we present preliminary results based on ACA to analyze and visualize research communities in the area of technology enhanced learning, focusing on publicly available citation and conference information provided through CiteseerX and DBLP. We describe our approach to collecting, organizing and analyzing appropriate data, as well as the problems which have to be solved in this process. We also provide a thorough interpretation of the TEL research clusters obtained, which provide insights into these research communities. The results are promising, and show the method's potential as regards mapping and visualizing TEL research communities, making researchers aware of the different research communities relevant for technology enhanced learning, and thus better able to bridge communities wherever needed.

## Introduction and Motivation

Technology Enhanced Learning (TEL) is a fascinating field, with lots of different research questions and aspects to focus on. Researchers in TEL can focus on learning infrastructure to support the re-use of learning objects or personalization, on intelligent tutoring systems, on mobile learning, or on collaborative learning in teams. They can also focus on professional learning and knowledge management infrastructures, learning in universities (computer science, engineering or other disciplines) and on learning in schools, with a lot of interesting research questions and results. Many different conferences and journals are devoted to different aspects of technology enhanced learning, providing a variety of forums through which to publish TEL research results.

The downside of this variety is, however, that TEL is a much more fragmented area than most other research areas, making it difficult to gain an overview of recent advances in the field. Even for experienced TEL researchers answering the questions: “*What communities and sub-communities can be identified in TEL*”, “*what research topics/specialties can be identified in a field of studies*” and “*what conferences are the most relevant for what topic and for which community*” is a difficult task, and for beginners it is obviously an impossible one.

Being aware of this fragmentation and of the various sub-communities which make up the TEL area is an important pre-requisite towards overcoming this fragmentation, increasing synergies between different sub-areas and researchers, and, last but not least, providing funding agencies with evidence of new research results, innovative applications and promising new approaches for technology enhanced learning.

This paper provides a first step towards this goal, by employing the technique of Author Co-citation Analysis (ACA) on the large subset of TEL conferences related to computer science as indexed by DBLP<sup>1</sup> and CiteseerX<sup>2</sup> – the latter provides citation information for each indexed paper. ACA relies on the insight, that if two authors are cited together very often in scientific articles, their work must be related to the same research field.

---

<sup>1</sup> <http://www.informatik.uni-trier.de/~ley/db/>

<sup>2</sup> <http://citeseerx.ist.psu.edu/>

We will describe our methodology for data collection, solutions for problems that we encountered, and the techniques of author-co-citation and factor analysis for detecting communities in a given research area. We will further describe and discuss our results, which provide an interesting insight into some important TEL research clusters, and close with a summary and discussion of next steps and future work.

## **Related work**

Co-author analysis and citation analysis is an important method when analyzing scientific communities. Ochoa et al. (2009) provides a very nice example of how such analysis can help provide greater insight into TEL research communities and collaborations, through visualizing and intuitively describing research community structure, focusing on TEL publications presented at recent ED-MEDIA conferences. They focus on co-author analysis and visualization of these relations and provide interesting insights into collaboration networks in the TEL area. Wild et al (in press) used the same data corpus for a trend analysis in the ED-MEDIA conference. By applying clustering techniques to the paper titles, they showed how certain technologies and approaches gained importance – including, among others, mobile learning, blended learning, portfolios, podcasts, game-based learning and assessment.

Similar introspective analyses have been applied to other research fields in the past. Henry et al. (2007) provide an analysis of the area of human computer interaction, based on the four major HCI conferences, focusing on citation analysis that use data relating to these conferences (between conferences, articles and authors), word cloud visualizations to characterize the four conferences, and other visualizations that characterize collaboration and other networks. This paper does not rely on sophisticated mathematical network analysis modes but is a very good example of the power of visualization to make the structure of these networks explicit.

The approach we build upon in this paper, author co-citation analysis, has not yet been used widely despite its potential for detecting and clustering scientific communities based on the mathematical notion of factor analysis. One of the best papers and a good introduction to this approach is the paper by White et al. (White, H. D. and McCain, K. W. 1998). This study presents an extensive domain analysis of a discipline – information science – in terms of 120 top-cited authors, based on their papers from 1975 to 1995, with citations retrieved from Social Scisearch via DIALOG. Tables and graphics reveal the specialist nature of the discipline over 24 years, based on author co-citation analysis. The results show an interesting split of the field into two main specialties, which barely overlap, namely experimental retrieval/information retrieval and citation analysis. Included is also a dynamic analysis of the field, based on three 8-year-periods, which shows changes of authors and areas. The analysis is based on journal citations, but neglects important conferences such as the ACM SIGIR conference, the most relevant conference for the IR community. In contrast, the citation database used in our paper, CiteseerX, includes all important computer science conferences and workshops, providing a broad overview of computer science as it relates to TEL.

Using similar techniques, Chaomei Chen and Les Carr (1999) present an analysis of hypertext research based on the ACM Hypertext conference series, with papers included from 9 conferences over 10 years. About half of the citations in this series refer to papers from the same series, which points to a very homogeneous research community. Again, dynamic analysis using three time periods is included. Only citations within these conference series were considered, while we include citations from all conferences. Due to their restricted focus, the factors discovered represent a finely grained view of the hypertext research area (including subareas such as design models, hypertext writing, open hypermedia and information visualization), while our factors represent broader research communities, centered around one or a few community-centered conferences such as Adaptive Hypermedia or AIED.

## **Collecting Co-Citation Data**

Following White et al. (1998), we assume that citing practices in a research community reflect the judgments as to which works by which authors are the most influential – for the field in general and for specific sub-themes. Aggregated over time, a definite structure emerges that can be considered the current state of the field. *Co-citation* is a very good way of establishing relations between authors that correspond to specific sub-themes and research areas

in a research community – even though they do not directly reference each other. We consider author A and B to be co-cited, if they are both cited by an author C – that is, both names appear at least once in the reference section of C’s paper. The more co-citations, the stronger the relationship is.

Our data sets were obtained from CiteSeerX and DBLP. CiteSeerX is a digital library focusing on the literature in computer and information science, being fairly complete. The articles are crawled automatically from the Web and then metadata and citations are extracted from these articles, again automatically. The CiteSeerX dataset contains more than 1.4 million paper records correlated with about 28 million citations. Due to the automatic data collection process, metadata in CiteSeerX are not always perfect, which leads to considerable problems that have to be solved before analysis starts. We will describe these problems and our solutions in the following subsections. In addition, DBLP is a computer science bibliography database, which relies more on human input (the maintainer of DBLP is Michael Ley, from the University of Trier), which covers about the same field as CiteSeerX, and currently contains about 1.3 million bibliographical records. DBLP metadata does not include citations, but has been used in our project to contribute high-quality metadata, to cope with ambiguous author names and to provide reliable conference statistics.

## Data collection

While it was not the goal of our research to determine the most relevant authors in TEL – such a goal would involve a more elaborate discussion on how “most relevant authors” should be defined – a good sample of highly cited authors in TEL covering as many areas of TEL as possible was obviously necessary. Obtaining such a sample for a diverse area such as TEL is no trivial matter. The following paragraphs discuss our approach and the steps needed to gather such a sample. Our data collection focused on data available through the CiteSeerX and the DBLP databases, both covering all computer science related research, and will extend this through additional databases covering educational and psychological research for TEL in the future.

*Obtaining a first sample.* To obtain a first sample of TEL conferences, we collected the lists of TEL conferences and journals to which a small sample of 13 well-known researchers submit their papers (Duval, Scott, Brusilovsky, Koper, Kieslinger, Klamma, Nejd, Balacheff, Sharples, Davis, Zimmermann, Wolpers, Sutherland). From these conferences and journals (as identified in DBLP<sup>3</sup>), we extracted the 100 most prolific researchers. In a second iteration, we collected the list of top-100 conferences and journals to which these 100 most prolific authors submit their papers. Our final sample of authors represents the most prolific authors from the 20 conferences and journals in the latter list that have a specific focus on TEL<sup>4</sup>. These conferences and journals cover 13.557 publications in total.

For these authors we created a co-citation matrix. This first step resulted in a rather sparse matrix (with some authors not co-cited with any other authors) and consequently a set of clusters extracted through our SPSS factor analysis which was difficult to interpret. Thus, subsequent iterations were designed to extend and refine the set of authors, as discussed in what follows; in addition they included other conferences such as Adaptive Hypermedia, User Modeling or Artificial Intelligence, which provide techniques for TEL infrastructures and algorithms.

*Adding more authors, increasing co-citations.* As regards extending and refining the set of authors, in the second iteration we first included more authors: the 50 most prolific authors from ED-MEDIA<sup>5</sup> and ECTEL<sup>6</sup>, 15 new authors from the IEEE TLT Board and Steering Committee<sup>7</sup>, and 5 more authors from the Telearn archive<sup>8</sup>. We also included the top-15 cited papers or books from ED-MEDIA 2005 – 2008 (Ochoa et al. 2009). Second, after merging these sets, we selected the authors with at least 20 publications in CiteSeerX DB and with at least 10 co-citations in our co-citation matrix. We also experimented with a threshold of 20 and 30 co-citations, but finally kept the 10-co-citation threshold, as the clusters obtained were of similar quality.

*Disambiguating authors.* At this point we realized there was a problem of disambiguation for some names, so we decided to check the name occurrences in DBLP (where author names are manually disambiguated by the DBLP maintainer, Michael Ley) and to keep only the author strings that unambiguously identified the TEL authors we wanted to include. For example, we deleted John Cook because we found 269 occurrences of his surname in DBLP

<sup>3</sup> The detailed procedure is described in the Stellar deliverable D7.1: [http://www.stellarnet.eu/d/7/1/Investigating\\_two\\_silos](http://www.stellarnet.eu/d/7/1/Investigating_two_silos)

<sup>4</sup> Other topics are computer science (27 venues), artificial intelligence (26), human-computer interaction (22) and databases (5).

<sup>5</sup> <http://ariadne.cs.kuleuven.be/edmedia/rankings.html>

<sup>6</sup> <http://ariadne.cs.kuleuven.be/ectel/rankings.html>

<sup>7</sup> <http://www.computer.org/portal/web/tlt/edboard>

<sup>8</sup> <http://telearn.no-kaleidoscope.org/>

but, when queried by his full name we found only 12 publications in DBLP and 8 publications in CiteceerX. We deleted John Black as well, because the occurrences both in DBLP and CiteceerX were too ambiguous to correctly attribute publications or citations (John Black, John A. Black, John B. Black, John D. Black, John E. Black, John R. Black, John A. Black Jr). Based on this disambiguation, we kept the full name of each author, and the initials when this did not result in duplicates or ambiguity in DBLP. This left us with 77 authors for our analysis.

*Adding and checking more conferences.* To better characterize the clusters found through Component analysis, we checked the top 4 venues for each author. This had to be done using DBLP, as CiteceerX does not contain complete references for all papers, but sometimes only refers to them as technical reports. We then used DBLPVis<sup>9</sup>, to check for the five most prolific authors in all these TEL conferences covered DBLP and CiteceerX (AIED, CSCW, EC-TEL, Edutainment, ICALT, ICCE, ICWL, ITiCSE, ITS (Intelligent Tutoring Systems), SIGCSE, Wissensmanagement, WMTE), to make our final co-citation matrix more complete, in total 55 authors. Using a threshold of 50 DBLP publications, we kept 30 of them. 25 of them were already in our matrix, which was an encouraging sign that our previous iterations had already produced a good sample for these TEL conferences. We added 5 new authors to our matrix, for a final matrix of 82 highly cited and co-cited TEL authors.

## Data processing – Problems and Solutions

We conducted our analysis on CiteceerX dataset. The following paragraphs discuss our approach and give an overview about the relevant tables considered from the database, as well as the problems encountered during data processing and our solutions for these problems.

*Tables.* CiteceerX is organized in terms of three main tables: *Papers*, *Authors* and *Citations*. The *Papers* table contains all the papers, unequivocally retrieved through an identifier. Every paper can be a different version of the same publication, each associated to a single value of the attribute *cluster*, e.g. one cluster ID is coupled with several paper IDs. In addition, the papers are connected with their authors. A single author can have multiple occurrences in the *Authors* table, one for each paper s/he wrote. Thus, the data set contains duplicated author identifiers, a common problem when dealing with publication data. Finally, the references for each paper are stored in the table *Citations* with the following information: paper identifier *cited\_paperID* of the paper which the reference is cited by, *citation title*, *venue*, *year* and the *authors* of the cited paper (a string field, with all authors concatenated).

*Processing.* To compute the co-citation matrix, we collected the subset of the paper citations corresponding to the references to papers written by the relevant authors, selected for our analysis. The lack of a paper identifier of the citation made our mining task more complex: to retrieve the cited papers of our author list, we had to search for our authors within the value of the attribute *authors* in the *Citations* table. This was possible after processing the dataset in three steps: 1) drop all the foreign keys inside the *Citations* table; 2) change dataset engine from InnoDB to MyISAM to enable efficient full-text search; 3) create a full-text index for the attribute *authors*. All the results were stored within a new *citations\_TopAuthors* table so as to provide reasonable processing time for our queries (the size of the new table is about 50,000 records compared to the 28 million in the original *Citations* table). Finally, to further increase processing time, we built another full-text index on authors.

*Multiple author aliases.* Since a single author can have multiple occurrences in the *Authors* table, we had to cope with the problem that author names may be misspelled or use initials instead of full first names; authors may also change their names or use different combinations of formal and informal names and initials in different papers, producing multiple identifiers we call *aliases* for a single person. The author “Wolfgang Nejdl” appears more than two hundreds time with his complete name, for example, and about ten times as “W. Nejdl”.

*Unique author identifier.* We then collected all the paper citations which had at least one previously computed alias in the *authors* attribute. For each of these circa fifty thousand records, we added one *firstAuthor* attribute in the new table to describe a single author with aliases with one identifier, e.g. we put “Nejdl” as identifier of “Wolfgang Nejdl” and “W. Nejdl”. Thanks to the fact that *firstAuthor* contains only one identifier, we were able to solve the problem of keeping information about the identifiers of a possible second or third author who wrote the same cited paper. We therefore duplicated, for each author of interest, the corresponding citation in the new table *citations\_TopAuthors* with the identifier for a second and subsequent author.

---

<sup>9</sup> <http://dblpvis.uni-trier.de/help/overview.html>

*Paper multi versioning.* Another issue we encountered was paper multi-versioning. Because the same paper can have several versions each of which has been crawled from the Web and given that each of these publications keep information about their references in the *Citations* table, we had to remove from our table the duplicate citations related to different editions of the same paper. To achieve this goal, we exploited the attribute *cluster*, as described before, of the table *Papers*.

## Matrix creation

For subsequent analysis, we then created a quadratic, symmetric matrix containing the listing of our selected authors as rows and columns, to be filled by co-citation data: for the  $j$ -th row and the  $i$ -th column, the retrieved value in this cell refers to the number of times the  $j$ -th author was co-cited with the  $i$ -th one.

For  $i$  equal to  $j$  we included a null value because it corresponds to the cell representing the number of co-citations of one author with her/himself.

Our matrix construction process includes three main steps:

- Select the identifier of all cited papers we collected in our table *citations\_TopAuthors*.
- For each of these identifiers, gather distinct authors, i.e. the values of the attribute *firstAuthor*.
- Whenever this previously computed result set carried more than one author, for each possible author pair, we incremented the corresponding values  $\langle i,j \rangle$  and  $\langle j,i \rangle$  in the matrix.

These steps lead to the following algorithm, described in pseudo-code and relevant SQL statements:

```

Select distinct cited_paperID from citations_TopAuthors;
For each cited_paperID
    Select distinct firstAuthor from citations_TopAuthors where cited_paperID = current cited_paperID
    If more than 1 firstAuthor
        Compute all possible author pairs
        For each author pair  $\langle i,j \rangle$ 
            Update matrix cell  $\langle i,j \rangle$  and  $\langle j,i \rangle$ 

```

**Listing 1.** Pseudo code for the matrix computation.

## Cluster Analysis and Discussion

We then proceeded to analyze our data, using Principal Component Analysis, to detect appropriate clusters / areas in TEL research, and then visualize and interpret these clusters.

### Using Principal Component Analysis to Detect TEL Research Areas

*Principal Component Analysis.* “In the social sciences we are often trying to measure things that cannot directly be measured (so-called latent variables)”, as Andy Field states in his book (Field, 2009). In our case, the interest in different topics or research areas of different authors in TEL cannot easily be measured. We could not measure motivation and interest directly, but we tried to analyze a possible underlying variable (collaboration in the form of co-citations among the major authors), to detect different sub-communities and possible trends. To do so, we used the statistical application SPSS to perform the Principal Component Analysis (PCA): a technique for identifying groups or clusters of variables and reduce the data set to a more manageable size while retaining as much of the original information as possible. Often, its operation can be thought of as revealing the internal structure of the data in a way which best explains the variance in the data.

*PCA vs FA.* Principal Component Analysis is similar to Factor analysis, but merely has the goal of finding linear components within the data and how a variable might contribute to these components (which basically means, finding some meaningful clusters within the data). Factor analysis uses the same techniques, but the aim is to build a sound mathematical model from which factors are estimated. The choice of PCA vs. FA depends on what we hope

to do with the analysis: whether we want to generalize the findings from our sample to a population, or whether we want to explore our data or test specific hypotheses. In our specific research, we used PCA because we wanted to explore the data with a descriptive method and apply our findings to the collected sample.

*Correlation determinant.* When we measure several variables with the PCA, the correlation between each pair of variables can be arranged in what is known as an R-matrix: a table of correlation coefficients between variables. The existence of clusters of large correlation coefficients between subsets of variables, suggests that those variables could be measuring aspects of the same underlying dimensions. These underlying dimensions are known as factors (or latent variables). In Factor analysis we strive to reduce this R-matrix to its underlying dimensions by looking at which variables seem to cluster together in a meaningful way. This data reduction is achieved by looking for variables that correlate highly with a group of other variables, but do not correlate with variables outside that group. Because our main aim is PCA, we did not have to worry about the correlation matrix determinant. Strictly speaking, the determinant or correlation matrix should be checked only in factor analysis: in pure principal component analysis it is not relevant (Field 2009), so that we could safely leave all our authors in the sample.

*Defining factors.* Not all factors are retained in an analysis, but only the most relevant and meaningful one for the research. In our case, we used Varimax orthogonal rotation<sup>10</sup> to discriminate between factors (to rotate the factor axes such that variables are loaded maximally to only one factor and we could better calculate the loading of the variable on each factor). We sorted the variables by size ordering them by their factor loadings, to display all the variables which load highly onto the same factor together. As a result we obtained a Rotated Component Matrix which shows the variables listed in order of size of their factor loadings. For interpretation purposes, we also suppressed absolute values which were less than 0.4.

We obtained 15 factors in total, which explain 78% of the variance; for this paper we focus on the first six factors, explaining 59%. Compared to (White and McCain 1997), where the first eight factors alone explain 78% of the variance, our lower value reflects the different disciplines that come together in TEL, producing many more sub-communities, while Information Science has some well-established communities that focus on a particular topic.

To describe the meaning of each factor more precisely we also added information regarding the conferences where our sample authors usually publish. For this paper, we included the top 4 venues for each author, as well as the number of papers published. Figure 1 shows the first two clusters, with a (small) subset of conferences displayed, Figure 2 clusters 3-6.

cluster	author	Citiesexr	Agents	AH	AIED	AIEDU (I)	ANLP	ASCLITE	Advanced AVI	AWCC	BIS	CADE	Canadian CAISE	CDVE	CE (Compi CELDA	CG	CHI
1	Paul de Bra	71															
2	Alexandra I. Cristea	31															
3	Paloma Diaz	20															
4	Maria Grigoriadou	49															
5	Vittorio Sciarano	38															
6	Kishak	37															
7	Manuel A. Pérez-Quiriones	56															
8	Marcus Specht	43															
9	Franca Garzotto	23															
10	Wolfgang Nejdl	225															
11	Ralf Steinmetz	134															
12	Judy Kay	57															
13	Hermann Maurer	54															
14	Maria Vinou	58															
15	Sabine Graf	23															
16	Jim Greer	49															
17	Erica Melis	128															
18	Claude Frisson	22															
19	Erkki Sutinen	46															
20	Marek Hatala	30															
21	Hugh C. Davis	65															
22	SUM		0	40	18	0	1	0	1	5	0	0	7	3	0	0	4
23																	
24	cluster 2																
25	Cristina Conati	51															
26	Carolin Penstein Rose	91															
27	Neil T. Heffernan	29															
28	Bruce M. McLaren	70															
29	Kurt VanLehn	99															
30	Vincent Alaven	70															
31	Joseph E. Beck	54															
32	Jack Mostow	47															
33	Antonia Mitrovic	55															
34	Erica Melis	128															
35	Kenneth R. Koedinger	102															
36	Beverly Park Woolf	30															
37	Brent Martin	28															
38	Claude Frisson	22															
39	Pierre Dillenbourg	23															
40	Maria Vinou	58															
41	Jim Greer	49															
42	W. Lewis Johnson	67															
43	Elad Soloway	22															
44	SUM		5	5	158	6	0	0	0	0	0	0	7	3	0	0	0

**Figure 1:** Authors and top 4 venues for each author, for the first two clusters

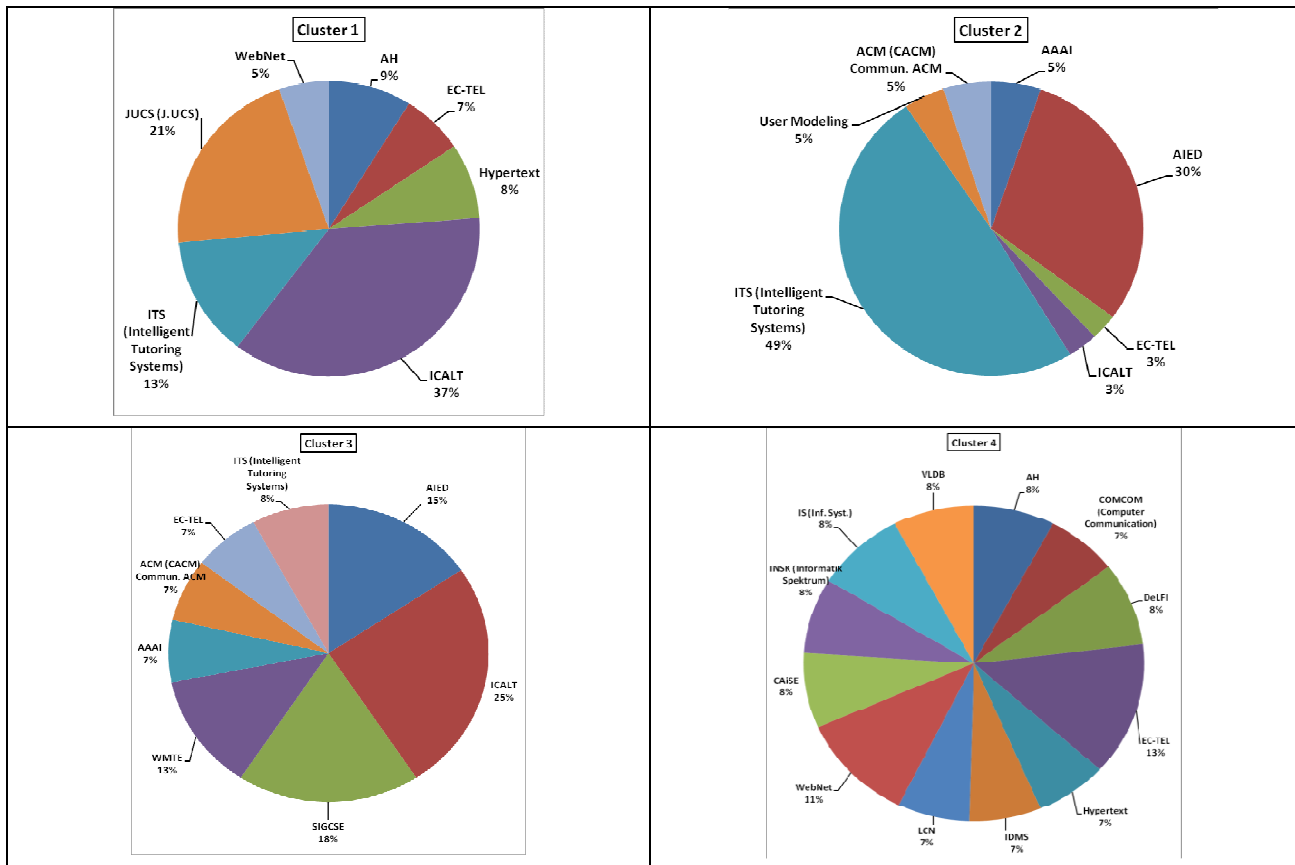
<sup>10</sup> The Varimax rotation attempts to maximize the dispersion of loadings within factors. It tries to load a smaller number of variables highly onto each factor resulting in more interpretable clusters of factors.

cluster 3	author	CiteseerX	cluster 5	author	CiteseerX
,923	Stefanie N. Lindstaedt	20	,911	Mordechai Ben-Ari	23
,919	Mark Guzdial	37	,823	Guido Rößling	32
,718	Mike Sharples	38	,555	Susan H. Rodger	21
,679	W. Lewis Johnson	67			
,594	Ron Oliver	39			
,571	Erkki Sutinen	46			
cluster 4	author	CiteseerX	cluster 6	author	CiteseerX
,929	Daniel Olmedilla	45	,667	José Luis Sierra	48
,781	Peter Brusilovsky	93	,585	Colin Tattersall	33
,735	Marek Hatala	30	,577	Rob Koper	91
,747	Ralf Steinmetz	134	,570	Baltasar Fernández-Manjón	46
			,585	Sabine Graf	23

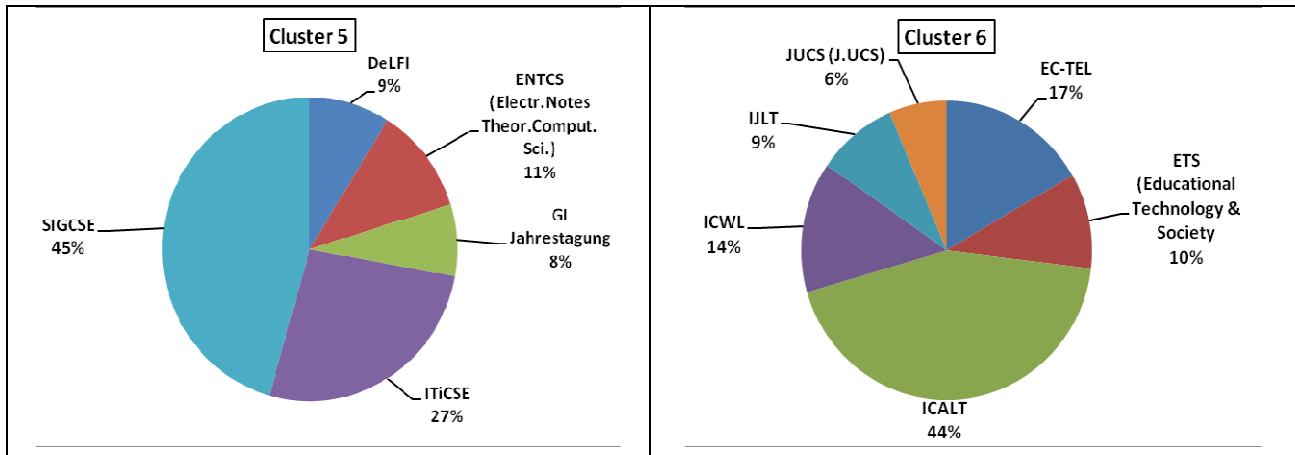
**Figure 2:** Authors, factor loadings and CiteseerX publications for cluster 3-6

### Visualizing TEL research clusters

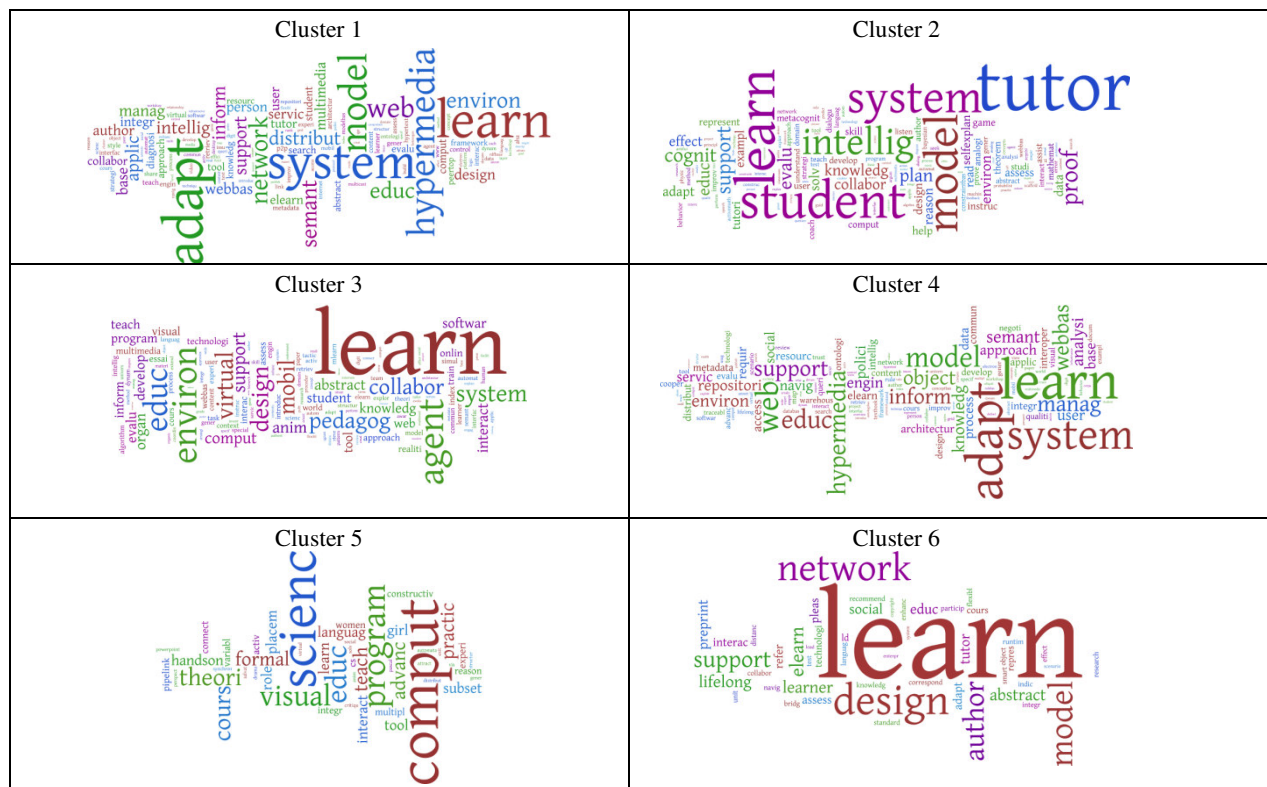
*Visualization based on conferences.* Based on this analysis, the following figures provide a visualization of the TEL research clusters obtained, first based on pie charts relating to the most relevant conferences for each cluster. To produce the conference-based charts, for each author we collected his/her four most frequented conferences according to DBLP (names of conferences as well as number of papers published by this author), added the number of papers for each conference and cluster, and then produced the following pie-charts including the most representative conferences for each cluster. For Clusters 1 and 2, conferences were selected if they included more than 20 publications (for Cluster 1) and 15 publications (for Cluster 2) from the cluster authors, for Clusters 3-6, we used a threshold of 5-7 publications to select the representative conferences.







*Visualization based on Tag Clouds.* Based on the clusters we retrieved, we selected from the CiteseerX dataset all the paper titles whose authors were in the cluster of interest. From the extracted paper titles we removed the words with less than 2 characters and the words consisting of numbers because these were not useful when determining the topic of a paper; for those words containing punctuation marks such as "\-\"?\" \%\" and \", we removed the punctuation marks and combined the remaining parts. We also removed stop words and applied stemming, as well as duplicate words inside a paper's title. We then assigned a counter to each distinct word, counting the number of occurrences of the word inside the titles. Last, we sorted all words in increasing order based on the counters and visualized the first 150 words.





## Discussion

The combined information from the clusters of researchers, the main conferences and journals that they address and the most often used keywords in their publications clearly show the differences in focus in the community – in terms of research as well as in terms of publications and connections. In this section, we discuss the main findings from the visualizations presented before.

The main publication venues (Figure 3) of the *first cluster* of researchers (Figure 1) include – besides main TEL conferences such as ITS and ICALT and the general journal JUCS – Adaptive Hypermedia, Hypertext and ECTEL. From the word cloud (Figure 4) of this cluster – with “Adapt”, “Model” and “Hypermedia” as distinctive words – a clear focus on *adaptive hypermedia systems* can be observed. This cluster contains authors like Paul de Bra (his four most frequent conferences are Hypertext, WebNet, AH and EC-TEL), Marcus Specht (EC-TEL, AH, WebNet), Hugh Davis (ICALT, Hypertext) and Wolfgang Nejdl (AH and many non-TEL conferences focusing on the Web and Information Systems). The cluster also includes personalization as represented in other relevant conferences listed (Judy Kay, for example, publishes most in ITS, AH and AIED).

Most authors in the *second cluster* have their roots in the field of artificial intelligence – as shown from the main publication venues AAAI and AIED. The conference on Intelligent Tutoring Systems is – in terms of quantity – the most important conference of this cluster. Authors in this cluster include Carolyn Penstein Rose (ITS and AIED), Bruce McLaren (ITS, AIED and EC-TEL) and Kurt Van Lehn (ITS and AIED). Jim Greer is included in the first two clusters, publishing most in ITS and AIED, but also in the EC-TEL and UM conferences, which are closer to the first cluster. Whereas the focus of the first cluster is on personalization and adaptation, the second cluster mainly focuses on understanding learners’ needs, by applying reasoning techniques to the models of the learner – this can also be observed from the word clouds – “Learn(-er/-ing)”, “Student”, “Model” and “Cogni(tion)” are the most significant words for this cluster.

The differences in terms of background and focus between the first two clusters are striking, given the similarity in research goals. Learner or user modeling is the first step in the process of adapting a system to the learner (Paramythis and Weibelzahl 2005). It is to be expected that these clusters will become more related with one another, as the targeted conferences AH (first cluster) and UM (second cluster) have merged into the UMAP conference in 2009.

Terms that show up in the *third cluster* are “Environment”, “Mobile”, “Pedagogy”, “Agent” and “Design”. Researchers in this cluster have more diverse backgrounds than in the first two clusters, but with the common denominator that they focus on the application of specific technologies to learning. These focuses include mobile technologies (Mike Sharples, Erkki Sukinen – WMTE), computer science education (SIGCSE, Mark Guzdial) and knowledge management.

The *fourth cluster* is an interesting cluster, related to Cluster 1 (“Personalization”), with Peter Brusilovsky as most prominent author. However, this cluster is more focused on learning objects than the first cluster, as witnessed by Erik Duval, as another prominent author. Apart from “Adaptation” and “Hypermedia”, the word clouds of this cluster include “Object”, “Semantic”, “Repository” and “Metadata”. As the first cluster, it also includes authors publishing not only in TEL, but in other areas (Ralf Steinmetz and Matthias Jarke), which (because of the smaller cluster size) has a bigger impact on the pie chart, which now includes several non-TEL related conferences relevant to information systems and communications as an explicit hint as to how other computer science related areas often influence TEL research.

The *fifth cluster* is a very application oriented cluster, with two TEL conferences mostly relating to computer science education (SIGCSE, ITiCSE, Mordechai Ben-Ari as prominent author), and an interesting non-TEL conference on Theoretical Computer Science showing the background of Guido Rößling (ENTCS, otherwise publishing mainly in ITiCSE and DeLFI, the German eLearning conference).

In terms of number of publications, Rob Koper is the most prominent researcher in the *sixth cluster*. An online search on these researchers shows that all of them have contributed to the theory of Learning Design (Koper and Tattersall 2005) and related technologies and standards, such as SCORM (Dodds 2007) – as exemplified by Baltasar Fernández-Manjón. Not surprisingly, “Learning Design” is the leading term of this cluster’s word cloud.

It is apparent that the lists of most popular conferences and journals for each cluster do not only contain TEL-specific conferences: they also contain conferences with a focus on artificial intelligence (AAAI) and human-computer interaction (AH, UM). On the one hand, this shows the importance of these areas to TEL – which matches the numbers of non-TEL venues that we identified during our data collection, as explained earlier in this paper – but

also shows that TEL-related work is presented at other venues. This can be interpreted as evidence for the multidisciplinary character of TEL research.

From these six clusters, the *building blocks* of the computer-science related research in TEL can be observed as:

- human-computer interaction, most prominently (adaptive) hypermedia systems (cluster 1)
- artificial intelligence and (reasoning techniques for) user modeling (cluster 2)
- semantics, repositories and metadata (cluster 4)

Cluster 3 and 6 represent the more TEL-specific innovative areas. The terms in their word clouds overlap to a large extent with the 'new terms' in ED-MEDIA, as identified by Wild et al (in press).

## Conclusions and Future Work

In this paper, we used author co-citation analysis to analyze and visualize research communities in the area of technology enhanced learning, focusing on publicly available citation information provided through CiteseerX and conference information available through DBLP. The results are visualized based on relevant conferences and themes for each cluster, providing a first important step to provide a structured overview over research in technology enhanced learning and make TEL researchers aware of the different research communities relevant for their work.

As an important next step, we will extend our dataset with additional publication and citation data relevant for TEL, most importantly education and psychology, as relevant for example for computer supported collaborative learning<sup>11</sup>. These steps are currently performed, together with other project partners, in the context of the STELLAR Network of Excellence.

We hope that this work as well as future work building on it, will help overcome TEL research fragmentation, by making TEL researchers aware of the different research communities relevant for technology enhanced learning, and thus more able to bridge communities wherever needed.

## References

- Chen, C., Les Carr (1999). *Trailblazing the Literature of Hypertext: Author Co-Citation Analysis* (1989-1998). Hypertext 1999.
- Dodds, P. *SCORM Primer*. Retrieved from <http://adlcommunity.net/mod/resource/view.php?id=458>. Last modified: April, 2007.
- Henry, N., Howard Goodell, H., Elmqvist N., Fekete J-D. *20 Years of Four HCI Conferences: A Visual Exploration*. Intl. Journal of Human-Computer Interaction, Volume 23/3, December 2007.
- Ochoa, X., Mendez, G., Duval, E. (2009). *Who we are: Analysis of 10 years of the ED-MEDIA Conference*. ED-MEDIA 2009.
- Field A., (2009). *Discovering Statistics Using SPSS*. SAGE Publications
- Koper, R. & Tattersall, C. *Learning Design – A Handbook on Modelling and Delivering Networked Education and Training*. Springer.
- Paramythis, A., & Weibelzahl, S. (2005). *A decomposition model for the Layered Evaluation of Interactive Adaptive Systems*. 10th International Conference on User Modeling, 2005, Springer.
- White, H.D. and McCain, K.W. (1998). *Visualizing a discipline: an author co-citation analysis of information science, 1972–1995*. J. Am. Soc. Inf. Sci. 49, 4, Apr. 1998.
- Wild, F., Valentine, C. and Scott, P. *Shifting Interests: Changes in the Lexical Semantics of ED-MEDIA*. Intl. Journal of Elearning, in press.

## Acknowledgments

Work on this paper has been partially sponsored by the STELLAR Network of Excellence.

---

<sup>11</sup> The CSCL conference, for example, is not indexed in DBLP and therefore missing in our analysis.

---

## Research 2.0: Drawing on the Wisdom of the Crowds to Develop a Research Vision

Marie Joubert and Rosamund Sutherland

University of Bristol, UK  
{marie.joubert, ros.sutherland}@bristol.ac.uk

**Abstract:** This paper describes and reflects upon taking a ‘Research 2.0’ approach to developing a ‘vision and strategy statement’ for a network of researchers involved in researching Technology Enhanced Learning (TEL). It relates how the statement was developed first by collecting content from colleagues within the network through face to face meetings and contributions to a wiki and then by creating a coherent linear text document which further developed the content on the wiki. It discusses the risks inherent in the approach and outlines the strategies taken to address the risks. It suggests that, although the approach taken was successful, the success was limited owing to factors including a) limited engagement by the community with other people’s contributions, b) a reluctance to amend other people’s contributions and c) the difficulty of aggregating the multiple voices within the community while retaining faithfulness to the philosophies underpinning a ‘Research 2.0’ approach.

**Keywords:** deliverables, wiki, collaboration, analysis, #stellarnet

### 1 Introduction

This paper describes, and reflects on, the approach taken to developing a research ‘vision and strategy’ statement for the European Network of Excellence, STELLAR. The statement needed to reflect the views of a diverse community of researchers in Technology Enhanced Learning (TEL), represented by individuals from a variety of backgrounds such as computer science, engineering, education and psychology. The representatives of the community work in sixteen different labs in nine different countries in Europe and work within a wide range of research and cultural traditions. Given the diversity of backgrounds of the individuals within the community, producing a joint vision and strategy was a significant challenge.

This paper reflects on how the deliverable was produced using a ‘Research 2.0’ approach, critically examining the process and the products. The paper develops a use case scenario, discusses the influence of Research 2.0 on the scientific practice of developing the statement and evaluates the use of Research 2.0 tools. The paper describes the novel approach adopted and the successes and failures of the endeavour.

## **2 Background**

STELLAR is a multi-disciplinary consortium (Network of Excellence) which aims to bring together the different research traditions and disciplines within TEL. The cornerstone of the work of STELLAR is the Description of Work (DoW), which was developed by drawing on knowledge and expertise of members of previous Networks of Excellence, Kaleidoscope and Pro-Learn.

The DoW identified three themes (called ‘Grand Challenges’) intended to be a starting point for providing a framework to identify and formalise the visions and strategies for TEL: 1) Connecting learners 2) Orchestrating learning 3) Contextualizing virtual learning environments and instrumentalising learning contexts. For each theme, the DoW also posed a number of related research questions.

One of the early deliverables for the consortium was to produce a document outlining the vision and strategy of the whole STELLAR consortium, by developing the themes in the DoW. One partner of STELLAR (University of Bristol) had ultimate responsibility for the document, but considered the vision and strategy for the consortium to be the responsibility of all partners, and wanted to find a way for the whole consortium to contribute to the joint vision and strategy. As such, the enterprise could be seen as successful if all partners were actively engaged of in the construction of the vision and strategy.

As a Network, STELLAR subscribes to the idea of ‘Science 2.0’ as a way of working; this approach draws on ‘Web 2.0’ and can broadly be described as being underpinned by the democratic principle in which members of a community have the opportunity to contribute to a collaborative project and the contributions of all individuals are valued and become aggregated to represent the ‘wisdom of the crowds’ [1].

‘... Web 2.0 has been ushered in by what might be a thought of as rhetoric of ‘democratisation’. This is defined by stories and images of ‘the people’ reclaiming the Internet and taking control of its content; a kind of ‘people’s internet’ ... This, we are led to believe, has led to a new collaborative, participatory or open culture, where anyone can get involved, and everyone has the potential to be seen or heard.’ [2]

‘ The Internet is enabling an unprecedented number and variety of individuals to contribute knowledge, by authoring content individually or collaboratively and by helping one another directly in online forums. [3]

We argue that, because the research approach parallels the ‘Web 2.0’ approach, it could be called ‘Research 2.0’. Research 2.0 uses tools and technologies as appropriate for the tasks involved in the research process, and these may include Web 2.0 tools, such as wikis, blogs, micro-blogs, podcasts, reference management and sharing (e.g. Delicious and Mendeley), photograph sharing (e.g. Flickr\*) and social networks. (For example, see [4] and [5]). However, we argue that Research 2.0 can also use more traditional non-digital research tools to generate content such as face to face discussion, focus groups and interviews. Our key concern was knowledge creation using appropriate methods and tools.

### 3 Quantity and Quality of Knowledge Produced in Wikis

We suggest that we have much to learn about knowledge creation within a 2.0 approach from the use of Web 2.0 tools and hence we draw on literature relating to Web 2.0, and in particular wikis, to inform us. We focus on the literature concerning wikis for two reasons: first because Wikipedia is generally agreed to be a successful example of knowledge creation (e.g. see [6], [7]) and second because we chose to use a wiki for our knowledge creation project. This literature falls into two key areas: the first is concerned with the processes of collaborating to produce knowledge and the second with the nature and extent of knowledge itself. The literature review below is framed within these two key areas.

**Processes of collaborating:** Producing knowledge collaboratively using Web 2.0 technologies (wikis) is still relatively new and the concern of much literature in the area is about ‘what works’. We argue that understanding online collaboration is at the heart of the ‘what works’ question. Coleman and Levin, 2008, put forward their view on collaboration:

Collaboration is, we believe, primarily about people, about trust, and about the willingness to share information and work in a coordinated manner to achieve a common goal [8] (p 25).

We agree; collaboration is between people, who coordinate to achieve a common goal; in the context of this paper, this coordinated working involves sharing knowledge and building knowledge together. Those concerned are willing to share knowledge and want to share knowledge.

Contributors’ motivations seem to be critical for sustaining Wikipedia and other collaborative user-generated content outlets. [9], (p1)

As Coleman and Levine (ibid) point out, it is important to establish trust between the collaborators. This seems to be particularly important in online collaboration:

Web 2.0 is built upon Trust, whether that be trust placed in individuals, in assertions, or in the uses and reuses of data. [9].

... in and of themselves, these technologies cannot ensure productive online interactions. Leading enterprises that are experimenting with social networks and online communities are already discovering this fact and along with it, the importance of establishing trust as the foundation for online collaboration [10]

A further point made by Coleman and Levine is that in successful collaboration the goal is shared and that members of the collaboration have the same (or similar) end point in mind. This point was also made by Wagner and Majchrzak [11], who developed a set of enabling characteristics for successfully engaging ‘customers’ in a wiki through a detailed study of three cases: “Boomtown Times” (a pseudonym) wiki editorial experiment, Novell’s Cool Solutions wiki, and Wikipedia. They found that if users’ goals were aligned, the endeavour was more likely to succeed.

A factor that is sometimes reported in the literature as contributing to successful online collaboration concerns explicit rules related to contributing content. Wikipedia includes a page of ‘rules’ and ‘guidelines’ which are described as a ‘policy, a widely accepted standard that all editors should normally follow. Changes made to it should reflect consensus.’ (see [http://en.wikipedia.org/wiki/What\\_Wikipedia\\_is\\_not](http://en.wikipedia.org/wiki/What_Wikipedia_is_not)). Wagner and Majchrzak (ibid) suggest that these guidelines ensure quality:

Wikipedia has strong editing guidelines that are motivated by the refactoring rules of software development and principles of objectivity. This ensures that articles, which might have suffered in readability from the disjointed work of multiple contributors and commentator, ultimately becomes very readable again. [11]

However, while there are some who consider that rules encourage contribution to the wiki, such as Wagner and Majchrzak (ibid), others have found that the presence of rules makes little difference, (e.g. [12]).

Finally, it seems that constructive engagement could be encouraged by allowing different levels of participation; ‘lurking’, commenting on others’ contributions, making original contributions, editing and asking for explanations of others’ ideas and organisation of content for better structure. [11,12]

**Quality of knowledge:** Wikis can be successful tools for collecting and aggregating knowledge. As pointed out above, Wikipedia, probably the best known wiki, is generally seen as a success. At the time of writing this paper (July 2010) it had over 3 million articles in the English version, and it is in the top ten web sites accessed anywhere. This demonstrates that it is possible to create a wiki that ‘works’ in terms of community engagement. There is debate, however, about the quality of the knowledge on wikis.

Whereas wikis sometimes have rules of engagement, the knowledge produced on wikis is usually not subject to editorial control which leads to concerns over the provenance of information posted. Concerns relate to various aspects of knowledge, largely to do with the accuracy of knowledge. For example, Don Fallis (2008) suggests that:

serious concerns have been raised about the quality (e.g., accuracy, completeness, comprehensibility, etc.) of the information on Wikipedia [13] (p 1663)

Fallis’ article suggests that Wikipedia has been dismissed by much of the library and information science communities because it is seen as unreliable. He presents a thorough analysis of potential different types of inaccurate information in terms of factual accuracy, completeness, currency and comprehensibility and he demonstrates that Wikipedia fails rigorous tests of accuracy in these respects. However, he continues by arguing that Wikipedia is ‘quite reliable’ and ‘quite verifiable’ and that it contains ‘quite a lot of high-quality accurate information’ (p 1669). He makes the point that ‘it is probably epistemically better ... that people have access to this information source’. (p 1669). He argues that there are ways in which the reliability of information on Wikipedia can be improved, but points out that the cost of this would undermine some of the values on which the project is based, such as the number of contributions and the speed with which entries are added and updated. His key point is that ultimately it is the responsibility of readers ‘to decide whether to believe what they read on Wikipedia’ (p 1671) and he concludes by suggesting ways in which to help readers in this respect (e.g. signaling evidence of the quality of articles, directing readers to further reading, flagging omissions).

Concerns over the accuracy of information on wikis and Wikipedia in particular frequently relate to factual content (and this is to be expected in the case of Wikipedia which collects ‘facts’). However, there are other concerns which relate to the quality of knowledge built using online collaboration. For example, Anderson [5] argues that the ‘Web of Content’ (WoC) discourages ‘a deep level of critical thinking’ because development of content is influenced by a ‘powerful zeitgeist’. The computer

scientist, Jaron Lanier, in an essay about the dangers of elevating collectivism above merit and thus lowering standards, describes a similar concern:

What I've seen is a loss of insight and subtlety, a disregard for the nuances of considered opinions, and an increased tendency to enshrine the official or normative beliefs of an organization. [14]

This section has outlined some of the key issues relating to the collaborative production of knowledge within an online environment, with a focus on the use of wikis. It demonstrates the keys risks associated with using a wiki in terms of the amount of knowledge produced and the quality of the knowledge. In terms of the former, the main risk seems to be non-participation in the process of knowledge building and we recognised within our project that we may need to take steps to encourage our colleagues in STELLAR to contribute to the wiki. In terms of the latter, the risk for us was less clear. Our project was not essentially about collecting facts, as Wikipedia is, and we did not consider that we risked inaccurate contributions. Our project was more about developing arguments, debate, insight and vision and did, perhaps, run the risks described by Anderson and Lanier above. These risks were less clear to us at the beginning of the project but as it developed we put strategies in place to encourage high quality debate.

## **4 Developing the Vision and Strategy Statement**

### **4.1 Starting Points**

The text from the DoW was used as a starting point to create a 'Grand Challenges' wiki. The text was pasted into three main pages, one for each of the three Grand Challenge themes. At the same time, the wider STELLAR community was asked to recommend reading related to producing a TEL vision and strategy statement. The recommended readings and were put together and distributed to the STELLAR network and posted onto the STELLAR web site. Members of STELLAR were asked to engage with the readings prior to the face to face meeting described below.

### **4.2 Face to Face Meeting**

A day-long face-to-face meeting was set up in Bristol in May 2009 (month 4 of STELLAR). 33 members of STELLAR participated and worked in three groups, each with a chair and a note-taker. The groups were constructed to include individuals who represented the diverse research interests and perspectives within STELLAR.

In the morning there were two discussion sessions. Participants remained in the same groups for both these sessions although the chairs and note takers were different.

In the first session groups discussed questions relating to the Grand Challenge theme 'connecting learners'. Each group was given one of three questions to discuss:

- What are key enabling and success factors for learner networks?

- What impact could web 2.0 technologies have on learning in educational institutions and what are the implications for a) professional development b) design and organisation of learning spaces c) policy makers?
- What are the changing demands for workplace knowledge and skills and what are the implications for a) leaders and managers and b) the workforce?

In the second session groups discussed questions relating to the Grand Challenge theme ‘orchestrating learning’:

- What is the role of the teacher/more knowledgeable other in orchestrating learning and how does this relate to collaboration and the knowledge of students?
- What is the role of assessment and evaluation in learning and how can technology play a role?
- From the point of view of the learner what is the relationship between higher-order skills and learning of a particular knowledge domain and what is the role of technology in this respect?

For the third session (which took place in the afternoon), participants were put into new groups. These groups discussed questions relating to the Grand Challenge theme ‘Contextualising virtual learning environments and instrumentalising learning contexts’:

- How can new forms of technology-enhanced learning enable novel experiences for learners and for development of human competences and capabilities?
- How can the mobility of the learner in distributed and multi environment learning settings be supported, to include the transition between a) real and virtual contexts b) informal and formal learning contexts?
- Which standards are needed to achieve interoperability and reusability of learning resources in this field? How can we harmonise the existing learning standards?

The main purpose of the meeting was to expand the collective understanding of the community concerning the three research themes, through knowledge contributed by experts within the community and discussion and development of related research questions. The meeting was set up using an adaptation of the ‘knowledge café’ methodology (Firestone and McElroy, 2005). Within this methodology discussion is not driven by an agenda, and this is seen to encourage groups to develop discussion in line with the expertise and interests of the individuals in the group.

Note-takers were told that the notes would be added to the wiki but otherwise were not given any specific instructions or guidelines. They adopted different approaches but generally attempted to capture as many of the points being made as possible, not attempting to organise the points into coherent prose. The examples below are taken from discussion starting from the questions ‘What are key enabling and success factors for learner networks?’ and ‘What is the role of the teacher/more knowledgeable other in orchestrating learning and how does this relate to collaboration and the knowledge of students?’ The examples demonstrate different approaches taken to note taking.

#### **Example 1**

This is the first set of aspects created in the first grand challenge vision workshop on May 20th, 2009 in Bristol:

- Connections with people with whom you interact
- Merging of Formal & informal, Lifelong, Self-organised / self-constructed,
- One holistic network per person, not a private one, professional one...



- Medium used for communication is fundamental; Software can support maintenance and building of network
- Challenge: Integrate networks with learning processes
- Most prominently: Social network; but not only people: Networks of people, artefacts (e.g. paper), and tools (distributed cognition, actor-network theory)
- Sense of being in control essential (when to use, how to use, ...) / responsibility

### **Example 2**

What does a more knowledgeable other offer? A frame of reference/organised state of mind, knowledgeable other takes a scaffolding role - metalevel role - from research on expertise. Not just content knowledge - pedagogy as a whole - mediating content - children in school unlikely to have pedagogical expertise, but just more content knowledge. Teacher required to facilitate knowledge transfer/representation. Maybe there is a changing role of teacher within 21st century - but not necessarily to do with technology.

In one group, the notes were entered directly into the wiki and in the others they were written in a word-processed document and pasted into the wiki. These notes were seen as the starting point for extending the community's understanding of the Grand Challenges and the plan was to develop them into a more coherent whole over a period of weeks to form a substantial part of the vision statement. Importantly they were faithful to the spirit of the Research 2.0 approach in that contributions from all individuals were valued and the notes represent the collective responses of the community to the nine Grand Challenge questions chosen as the starting point.

### **4.3 Online Collaboration**

After the Bristol meeting STELLAR partners were invited to join a small team to coordinate the ongoing contributions to the wiki (to be called the D1.1 team). Apart from the Bristol team (UB), five partners volunteered: Istituto Tecnologie Didattiche in Italy (ITD), Ludwig-Maximilians-Universität München in Germany (LMU), Centre for Social Innovation in Austria (ZSI), Know Centre in Austria (KC) and Université Joseph Fourier in France (UJF). UB took a leadership role, with other team members taking responsibility for provoking STELLAR members to contribute to a particular subsection of the wiki related to:

- connecting learners (ITD and ZSI)
- orchestrating learning (LMU and KC)
- contextualising virtual learning environments and instrumentalising learning contexts (UJF and UB)

In the first half of June 2009, the D1.1 team met once online (using FlashMeeting, see <http://flashmeeting.open.ac.uk/home.html>) to discuss how to proceed. Following this, UB put together a written plan which outlined a tight time-frame for the development of the wiki:

- 22/6/09 to 6/7/09 – intensive work by all D1.1 team to get contributions from the whole STELLAR community.

- 6/7/09 to 30/7/09 UB will take responsibility for developing the wiki into a deliverable. Other D1.1 team members will be asked to contribute by a) writing sections b) reviewing sections and c) clarifying sections where necessary

UB also suggested strategies for the D1.1 team to use to provoke colleagues to contribute to the wiki. For example, written suggestions included:

For example, there might be a part of the wiki which you think requires further development; you could use this as a basis to develop a question for people to answer. You might make a sub page with this one question and invite people you know have expertise in the area to contribute a paragraph.

You might find that two people are making similar points, or two people are disagreeing, it might be worthwhile pointing out the synergies and encouraging further debate. However it could be important to find a way of keeping the ‘disagreements’ in the document.

The team met online again in the third week of June to discuss progress and to kick-start the phase during which the D1.1 worked intensively with colleagues to encourage them to contribute. Towards the end of this phase, one member of the UJF team came to work intensively on the wiki with the UB team for three days in the final week of July 2009.

This section has described the ways in which the online collaboration was organised. The next two sections reflect on the results of the online collaborations in terms of a) the extent of engagement of the STELLAR community and b) the nature of the contributions.

## 5 Reflections

### 5.1 Extent of Engagement with the Wiki

The wiki includes functionality to record the editing history of pages; an example covering the editing history of one page over the period of eight days is provided below:

► (cur   prev) ○	14:33, 21 July 2009	MurielNey (Talk   contribs)	(2,134 bytes) (undo)
► (cur   prev) ○	11:17, 17 July 2009	MikeSharples (Talk   contribs)	m (2,112 bytes) (→Contextualising virtual learning environments and instrumentalising learning contexts) (undo)
► (cur   prev) ○	11:16, 17 July 2009	MikeSharples (Talk   contribs)	(2,109 bytes) (→Contextualising virtual learning environments and instrumentalising learning contexts) (undo)
► (cur   prev) ○	07:37, 17 July 2009	MurielNey (Talk   contribs)	(2,219 bytes) (undo)
► (cur   prev) ○	13:43, 13 July 2009	MarieJoubert (Talk   contribs)	(1,870 bytes) (Created page with "Contextualising virtual learning environments and instrumentalising learning contexts " == As learning has become an integrative part of our life, and as it takes...)

Figure 1: Editing history of a wiki page

This information allows us to analyse the extent of engagement. Overall about 20 people from STELLAR contributed to the wiki in the period of development from 22nd June to 6th July 2009. However, sometimes a contribution under one name represented a collation of several contributions from an institution so it could be

argued that there were more contributors. The majority of the contributions were made by a small number of people, usually within a short time frame. For example, the three main pages: ‘Connecting Learners’, ‘Orchestrating Learning’ and ‘Contextualising Virtual Learning Environments and instrumentalising learning contexts’ pages had the following contributions:

Table 1: Contributions to the three main pages of the wiki

Page	Name	Date (in 2009) and Number of Contributions
Connecting Learners	Marie Joubert	13 July (1)
		16 July (1)
	Rosamund Sutherland	14 July (1)
		28 July (2)
		30 July (2)
	Nicolas Balacheff	26 July (5)
		30 July (2)
	Stefanie Lindstaedt	28 July (3)
Orchestrating Learning	Marie Joubert	13 July (1)
		16 July (1)
	Rosamund Sutherland	14 July (1)
		28 July (1)
		30 July (4)
	Nicolas Balacheff	26 July (5)
		30 July (2)
	Stefanie Lindstaedt	28 July (3)
Contextualising Virtual Learning Environments and instrumentalising learning contexts	Marie Joubert	13 July (1)
	Muriel Ney	17 July (1)
		21 July (1)
	Mike Sharples	17 July (2)

When individuals were asked to contribute by adding content, explanation or examples, generally they were very willing to do so. For example, when UB approached the Open University of the Netherlands (OUNL) asking for a clarification of what is meant by ‘interoperability’, the response was immediate and detailed.

Most people who contributed used the ‘Edit’ function to enter text directly into the wiki, either by adding in new text or amending text already present. A few used the ‘Discussion’ function.

The D1.1 team made concerted efforts to encourage contributions, but as their comments suggest, this was not always easy:

‘We have done really our best to obtain inputs and feedback, but it has been a hard task’ (email communication).

They went on to suggest that it had been difficult because people were not motivated to contribute because they did not understand the origins of the wiki and did not know what its purpose was.

Authorship was also seen as an issue for a number of reasons. There were conflicting ideas about whether or not to acknowledge individual contributions,

I am working on the wiki this week (until Friday). Although everything will appear under my name, I am integrating contributions from different people of my group. Thus I would like to let you know that VL and JP should also be mentioned in case there is a list of authors in the end (email communication).

Others were concerned about the extent to which it was appropriate to edit/modify/add to/ delete the contributions of other people. There seemed to be a tension between valuing and respecting other people's contributions (and not vandalising the wiki) but at the same time building the best possible document. As one contributor suggested, he was happy as an academic to use a word processor and the 'track changes' tool to write collaboratively. He suggested that using track changes can be seen as a way of checking with the original author that changes are acceptable; in other words track changes points out the suggested changes (which can then be accepted or rejected). In a wiki, however, the changes are not so obvious and anyone interested in the changes made would have to make a small effort to access the trail of development.

Many of those who did make changes seemed to need to check the changes they had made with the original authors. For example:

'I have done a bit of re-organisation, tell me if I am barking up the wrong tree' (email communication).

There was some debate about writing IN the wiki as opposed to writing in a word processor. There were some who thought that it was much easier to do the latter, but others who argued that this meant that the full authoring trail would be lost. Again, there was some debate about the authoring trail and about how important it is to retain the trail. On a similar note, there was a comment that sometimes people try to be the 'last author' in a wiki that is going to be frozen at a given time, because then their voice will be heard.

Finally, a possible barrier to contributing to the wiki may have been the technical difficulty of logging in to the wiki. We do not consider it to be very difficult, but it seems that some people found it confusing. For example, one STELLAR emailed to say:

'Unfortunately, it appears that I can't log in to edit it despite I can log in <http://www.stellarnet.eu/>'. (email communication)

## **5.2 Nature of Contributions**

The contributions varied in style and length. In general, they tended to take the form of paragraphs setting out the perspective of an individual. The first example, below, takes the form of an explanation about the meaning of 'interoperability', provided in response to a direct request from the D1.1 team (mentioned above). This response was sent by email.

Essentially this is about sharing resources and tools and system spanning. Within the community several specifications/standards are used. Basically there are several standards of content exchange that allow for exchange of learning content between different platforms. Furthermore interoperability is an important topic that considers more the functional integration of different learning services.

The D1.1 team found this sort of explanation to be very helpful as a starting point but found that contributions were seldom expanded, by either the original authors or other colleagues, with arguments, examples or references.

The second example below starts from ‘taken as read’ assumptions (contexts are more fluid) to suggest a change in focus for educational theory. It goes on to wrap up the paragraph by arguing against polarisation of educational theories.

When the context was relatively stable (in the case of fixed classrooms) educational theory tended to focus on content. However now that contexts are more fluid there is a shift from a focus on ‘content’ to a focus on ‘context’. However such a polarisation of ‘content’ and ‘context’ might be unhelpful in terms of understanding issues related to learning and knowledge construction.

The D1.1 team found this paragraph helpful and interesting, but again noticed that there were no further contributions to the paragraph.

In general, the D1.1 team found that the contributions on the wiki were individually valuable but that the levels of engagement with other people’s contributions was disappointing. There was little evidence of individuals challenging other people’s contributions or questioning what they had said, but typically were more concerned with phrasing and style. This is demonstrated by the example below, in Figure 2, which was taken from the editing history of the page on Orchestrating Learning. The text on the left is the earlier version, and the text on the right is an edited version.

The development of digital <del>teehnology</del> , their <del>interfae</del> -and association <del>te</del> -communication technology, has <del>open</del> -the possibility <del>te-aecees-te</del> -a large diversity of learning tools and all <del>kind</del> -of resources, as well as new infrastructures to support interactions and communications among learners and teachers or trainers – or in more general terms, among learners and knowledgeable others. This - evolution is supported by the <del>mergenee</del> -of theoretical frameworks which <del>provides</del> -new means to understand learning and to design more efficient and more relevant environments to support it. Situated cognition and situated learning theories, collaborative learning, exploratory learning as well as mobile learning theories are stimulating new approaches to learning, pedagogy, didactics and assessment.	The development of digital <u>technologies</u> , their <u>interfaces</u> and association <u>with</u> communication technology, has <u>opened up</u> the possibility <u>of accessing</u> a large diversity of learning tools and all <u>kinds</u> of resources, as well as new infrastructures to support interactions and communications among learners and teachers or trainers – or in more general terms, among learners and + knowledgeable others. This evolution is supported by the <u>emergence</u> of theoretical frameworks which <u>provide</u> new means to understand learning and to design more efficient and more relevant environments to support it. Situated cognition and situated learning theories, collaborative learning, exploratory learning as well as mobile learning theories are stimulating new approaches to learning, pedagogy, didactics and assessment.
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 2: Example of edited text

## 6 Producing the Deliverable

In order to produce the final document – a linear text document – the text was copied from the wiki into a word processor document. A UB team of two took responsibility for editing it. This involved forming it into a coherent narrative, removing repetition,

adding references, examples and explanations and amending text to achieve consistency in language and style.

A draft final document was completed. Once again, the UB team felt that it was important, even at this late stage, to work within a Research 2.0 approach and so the document was distributed to the whole STELLAR community with a request for feedback. In particular, the community was asked to check that any contributions they had made had been represented in the way they wanted.

Two members of the community were asked to provide internal peer reviews and a final version was produced, taking into account the feedback from the community and from the internal peer reviewers.

## **7 Conclusions**

The aim of the project described in this paper was to use a Research 2.0 approach to develop a vision and strategy statement for the STELLAR network. This paper described the processes and reported on the outcomes. This concluding section reflects on the project and ends with some recommendations.

We claim that the project was successful in many respects; members of the community did make contributions and the D1.1 editors were able to produce a deliverable based on the contents of the wiki. We suggest that the success of this way of gathering the views of the community can be explained by the existing ‘pre-conditions’ for a successful online collaborative venture, as outlined in the ‘Quantity and quality of knowledge produced in wikis’ section above. In particular the members of the community were willing and able to share knowledge and had, by the end of the Bristol meeting, developed a level of trust. On the whole, we could claim also that the community had a common goal, although – as reported above – perhaps this was not clear to all colleagues.

However, we were slightly disappointed that the D1.1 team had to work so hard to encourage the community to engage more deeply with the wiki and that many of the contributions were less well developed than we had hoped. As described above, the D1.1 team realised, as the project unfolded, that there was a risk that contributions may be less well formed and debated than hoped for, and made efforts to encourage deeper engagement.

Finally, we reflect on the Research 2.0 approach we took. This approach aimed to draw on the wisdom of the crowds (in this case STELLAR) and to aggregate the multiple voices of the individuals in the community in order to develop a coherent and unified vision and strategy for the community. However, the crowd had many voices and the spirit of 2.0 suggests that each should be valued and heard; the problem for us was that we could not aggregate all the voices while remaining faithful to the Research 2.0 philosophy underpinning our project. It may be that listening to the multiple voices of the crowds is at odds with forming an aggregation and it may be that we have to re-think how we conceptualise an ‘aggregation’ (particularly an aggregation of visions).

As pointed out above, the use of the wiki was perhaps not as successful as we hoped. We suggest that this was the case despite the will and technical ability of the

community to contribute. We do not fully understand why we were not as successful as we hoped, but we have some speculative suggestions:

1) Although it seemed that a good level of trust was present at the beginning of the project, STELLAR was a very new community and relationships within the community were still at an early stage. People did not know one another well and may have felt timid about making contributions. This paper has been written almost a year since the D1.1 project came to an end and in the intervening months the community has developed and grown, and (crucially) may be more willing to take the risk of publicly contributing to a growing wiki because of developing trust.

2) The construction of the wiki meant that it was difficult to engage with. There was too much text on each page, often well crafted, which did not seem to encourage discussion.

3) Members of the community did not seem to be clear about the goals of the wiki and how it would contribute to the vision and strategy of STELLAR. They therefore did not know what they should and should not be posting onto the wiki. Importantly, the project was not a research project; it was something different and therefore difficult to engage with.

4) Individuals were reluctant to change text that others had posted and others were reluctant to have their text changed.

In further work on developing STELLAR's vision and strategy, we intend to continue with the approach we used to produce this deliverable, and to experiment in the following ways:

- reduce the amount of text on each page and include prompts to encourage discussion
- make the hopes and intentions of the wiki (and the project) clear
- encourage the use of the 'Discussion' feature of the wiki to overcome the reluctance to change other people's entries
- make it clear that the wiki is a collaborative effort which is based on a Research 2.0 approach and is therefore about building knowledge together in a way that combines the voices of all the community.

## References

1. Surowiecki, J., and M. P. Silverman: The wisdom of crowds. *American Journal of Physics* 75:190 (2007)
2. Beer, D., and R. Burrows: Sociology and, of and in Web 2.0: Some initial considerations. *Sociological Research Online* 12 (5) (2007)
3. Adamic, L. A., Wei, X., Yang, J., Gerrish, S., Nam, K. K., Clarkson, G. S., et al: Individual focus and knowledge contribution. *First Monday*; Volume 15, Number 3 - 1 March 2010. Retrieved from <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2841/2475> (2010)
4. Ochoa, X., & Duval, E.: Quantitative Analysis of User-Generated Content on the Web. <http://journal.webscience.org>, 34(1). Retrieved from [journal.webscience.org/34/1/WebEvolve2008-03.pdf](http://journal.webscience.org/34/1/WebEvolve2008-03.pdf) (2008)
5. Andersen, P. : What is Web 2.0?: ideas, technologies and implications for education: JISC (2007)

6. Kittur, A., Suh, B., Pendleton, B. A., & Chi, E. H.: He says, she says: Conflict and Coordination in Wikipedia. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07, 453. New York, New York, USA: ACM Press. doi: 10.1145/1240624.1240698 (2007)
7. Ebner, M., Zechner, J., & Holzinger, A.: Why is Wikipedia so Successful ? Experiences in Establishing the Principles in Higher Education. In Proceedings of I-KNOW '06 (2006)
8. Coleman, D and Levine, S: Collaboration 2.0: Technology and Best Practices for Successful Collaboration in a Web 2.0 World Happy About; 2008:320 (2008)
9. Nov, O.: What motivates Wikipedians. In Communications of the ACM 50 (11) (pp. 60-64) (2007).
10. Miller, P.: Web 2.0: Building the New Library. Ariadane. Retrieved from 43131693cd2b7525be068ce24c2596cb031e9884 @ [www.ariadne.ac.uk](http://www.ariadne.ac.uk) (2005)
11. Kwan, M., & Ramachandran, D.: Trust and Online Reputation Systems. In J. Goldbeck, Computing with Social Trust - Human Computer Interaction Series (pp. 287-311). London: Springer-Verlag. Retrieved from [http://dx.doi.org/10.1007/978-1-84800-356-9\\_11](http://dx.doi.org/10.1007/978-1-84800-356-9_11)(2009)
12. Wagner, C and Majchrzak, A: Enabling Customer-Centricity Using Wikis and the Wiki Way Journal of Management Information Systems. 23(3):17-43 (2007)
13. Goldspink, C., Edmonds, B. and Gilbert, N.: Normative Behaviour in Wikipedia. 4th International Conference on e-Social Science, Manchester, June 2008. <http://epubs.surrey.ac.uk/cress/24/> (2008)
14. Fallis, D. Toward an Epistemology of Wikipedia Journal of the American Society for Information Science. 59(May):1662-1674 (2008)
15. Lanier, J. "Digital Maoism: The Hazards of the New Online Collectivism," Edge 183 (May 30, 2006), [http://www.edge.org/3rd\\_culture/lanier06/lanier06\\_index.html](http://www.edge.org/3rd_culture/lanier06/lanier06_index.html) (accessed October 25, 2010).



---

# The Afterlife of ‘Living Deliverables’: Angels or Zombies?

Fridolin Wild, Thomas Ullmann

Knowledge Media Institute, The Open University, UK  
{f.wild, t.ullmann}@open.ac.uk

**Abstract:** Within the STELLAR project, we provide the possibility to use living documents for the collaborative writing work on deliverables. Compared to ‘normal’ deliverables, ‘living’ deliverables come into existence much earlier than their delivery deadline and are expected to ‘live on’ after their official delivery to the European Commission. They are expected to foster collaboration. Within this contribution we investigate, how these deliverables have been used over the first 16 months of the project. We therefore propose a set of new analysis methods facilitating social network analysis on publicly available revision history data. With this instrumentarium, we critically look at whether the living deliverables have been successfully used for collaboration and whether their ‘afterlife’ beyond the contractual deadline had turned them into ‘zombies’ (still visible, but no or little live editing activities). The results show that the observed deliverables show signs of life, but often in connection with a topical change and in conjunction with changes in the pattern of collaboration.

**Keywords:** deliverables, wiki, collaboration, analysis, visualisation, #stellarnet

## 1 Introduction

In standard project management jargon, a ‘deliverable’ refers to a pre-defined, tangible, and verifiable work product such as a feasibility study or a prototype [1]. In research projects, deliverables often document process and outcomes of (more or less) systematic knowledge creation. They report on the progress against the tasks expected to be ‘delivered’ during a defined phase of the project. These documents sum up the focused work of a group or single person.

Within the STELLAR project, we provide the possibility to use living documents for the collaborative writing work on deliverables. They can be continuously updated and revised by all authors, even in parallel, using the popular wiki software MediaWiki (the software on which Wikipedia is based). Compared to ‘normal’ deliverables, ‘living’ deliverables come into existence much earlier than their delivery deadline and are expected to ‘live on’ after their official delivery to the European Commission. They are expected to foster collaboration in writing. Within this contribution we investigate, how these deliverables have been used over the first 16 months of the project. We will critically look at whether they have been successfully

used for collaboration and whether their 'afterlife' beyond the contractual deadline had turned them into a 'zombie' (arguably still some sort of life, but not a really welcome one). A zombie can still be seen, but does not show any signs of vital activity, whereas an angel cheerfully continues editing activities – but with the difference of being relieved from the duty of the mortal to deliver. It is clear that deadlines are typically drivers of activity, so also for angels, afterlife activity should be visibly less hectic and might focus on new or different areas of editing activity.

The analysis of the dynamics of wikis and their flagship Wikipedia is naturally a relatively young research field, since Wikipedia was created only back in 2001 – thereby making available a large public data-set of revision histories. Viegas et al. propose a method called 'history flows' for analysing the social dynamics expressed in the editing of Wikipedia articles [4]. They analyse the relationship between document revisions revealing cooperation and conflict patterns. Nunes et al. [3] use the revision history to visualize revision activity through sparklines in a timeline plot within their system 'WikiChanges', additionally supported by a 'tag-cloud'-like visualisation of term changes in the time frame selected (the font size is scaled by their changed frequency within the time window inspected). Arazy et al. [2] develop a series of glyphs to visualise contribution scores of authors in pages in order to ease the recognition of their work. Suh et al. [5] focus on identifying patterns of conflict with the help of so-called 'revert graphs', visualising the relation between authors of Wikipedia established through revisions that void previous edits. Baumgrass et al. [6] apply social network analysis in order to investigate corporate knowledge exchange processes in wikis. Closely related is also the work of Jesus et al. [7], within which network analysis is applied to study cluster-level collaboration between authors grouped by their work on related articles. Whereas [2,3,4,5] focus on the analysis of collaboration in individual pages, [6] and [7] deploy the same analytical technique – (social) network analysis –, but with a different focus of analysis [7] and in a different cultural and application setting [6].

All of them, however, share with our work the interest to shed light on the authorship relations documented in the revision histories. The user interface of the wikis is designed in a way, which centres the article and not so much the contributions of the single authors: its focus is on content and not authorship [2]. Making the authorship relation visible means extracting the relevant data from the revision histories of the pages and providing an easy to understand view of this data.

While a deliverable is the result of the edits of all authors, the revision history retains information about the contribution of each individual. This makes it easy to spot latest edits or compare changes with previous ones. It helps to keep track of the development of the pages contained in the living deliverable and, for example, make it easy to revert edits.

There are many ways of how to represent writing activity and collaboration of wiki pages. Within the rest of this paper, we first elaborate on our method of analysis used to make the collaborative writing process of living deliverables visible. With this, we analyse the data gathered within the STELLAR project so far: we visualize the overall co-authorship network; we outline the revision frequency over time to investigate if the living deliverables are indeed living; and we show how the collaboration network of authors and their contributions changes before and after a deadline. Finally, we conclude the paper with a summary and an outlook.

### 3 The data: Stellar's 'living deliverables'

The observed dataset consists of five living deliverables. They have been selected from the set of 14 wikis created so far for 19 project deliverables by excluding 'obvious zombies' and 'small group wikis' such as the coordination manual. Obvious zombies thereby relate to those wikis for which the group of collaborators did not use the offered wiki or abandoned it early in the writing process favouring different solutions to organise collaborative writing: these were mainly google docs and in several cases the exchange of word and excel files via mail with one or several editors consolidating tracked changes. The latter thereby being the main method used for the five management and evaluation deliverables that are much more clerical in nature and contain a lot of spreadsheet data – a task for which MediaWikis are hard to use.

Each living deliverable resides in its own MediaWiki instance. All wikis were initialized at the beginning of each deliverable writing period. While observing the process of the living deliverable evolution, we have to consider the fact that these documents served as input for the 'normal' deliverables (the type-set word or PDF file delivered to the European Commission), and the latter could then again feed back into the living deliverables.

The following Table 1 gives an overview of each of the investigated living deliverables. Among others, it outlines the number of authors, the number of pages contained in the wiki (and their number of page views), and – most notably – the number of edits these pages have received. All in all, the deliverables had an average number of 22.7 users, with a varying number of page views (in average 3,820). Some of them have received a substantial number of edits (such as the grand challenge document d1.1 and the science 2.0 mash-up deliverable d6.3, both earlier deliverables).

	Users	Total Views	Total Pages	Total Edits	Total Images	Pages/ Users	Edits/ Users
d1.1	78	14813	78	533	4	1	6.83
d1.2	9	1338	86	137	1	9.56	15.22
d6.1	4	677	39	152	28	9.75	38
d6.2	11	712	14	79	10	1.27	7.18
d6.3	21	2818	65	333	1	3.1	15.86
d7.1	13	2563	84	354	48	6.46	27.23

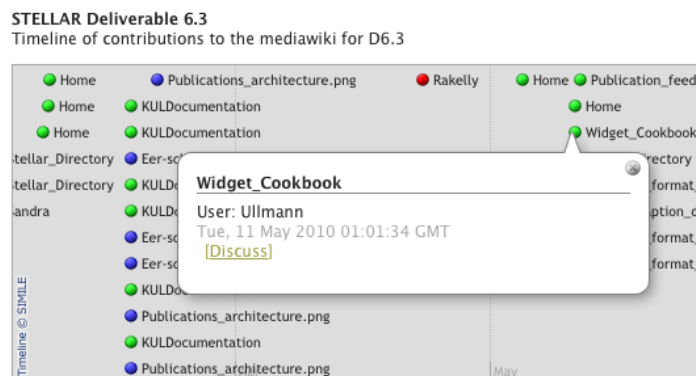
**Table 1.** Basic statistics of the investigated wikis.

#### 4 Method of analysis: SNA of the collaboration networks

The revision history of the living deliverables is a chronologically sorted list of changes of pages, listing – amongst others – the editing user, the page, the amount of characters changed with the revision, and a timestamp expressing when the revision was applied. One example of this revision history can be found in the snapshot of a revision history visualisation widget we have created to support the work in the deliverables (Figure 1): it shows the revision of one living deliverable in a scrollable timeline, listing the title of the changed page, the date of the change, and the name of the editor (pop-up bubble).

While this way of exploring the revision data has its benefit for following latest changes or browsing through the history of all changes, it does not provide much insight into the nature and vitality of the underlying collaboration, nor much insight into the focus of collaboration.

Collaboration is expressed in the co-authorship relations and can be extracted from the revision history. Co-authorship relations in living deliverables, however, can be investigated in many ways. The simplest form would be a list of authors of the deliverable or a page in it. List-like representations, however, do not show the structure of collaboration between the authors of the living deliverable. This extra dimension of information can provide insights into the collaboration network structure. We used a co-authorship social network analysis, which shows the relations established between authors by editing the same page. Therefore, an incident matrix was constructed listing the pages as incidents in the rows, the authors in the columns, and their number of edits of the respective page in the matrix cells. By multiplying the matrix with its transpose, an undirected affiliation matrix can be constructed and visualised as a network (see **Figure 2**).



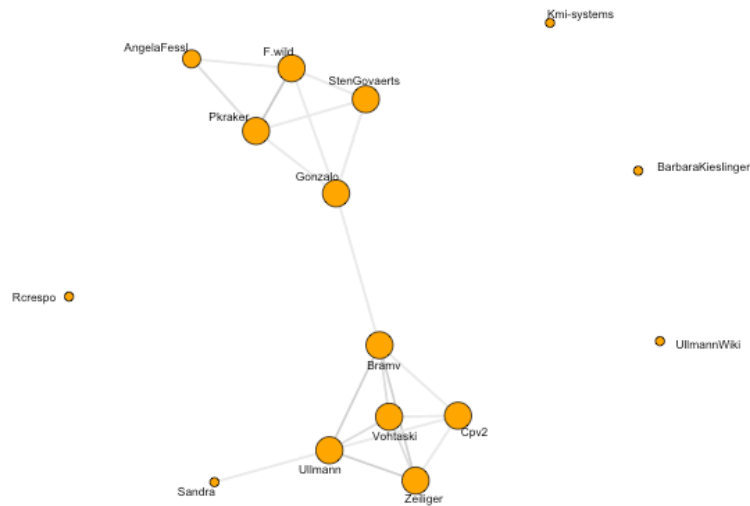
**Figure 1.** Timeline widget (visualizing the revision history of D6.3).

Since the central jump page ('home') of wikis is edited very often and by almost everyone (to, e.g., add links to new sub pages), it may be excluded from analysis in order to expose the clusters of collaborating authors more clearly (see Figure 3).



**Figure 2.** Collaboration network including edits of the central home page (D6.3).

The graph shows, a cluster of authors who contribute to a shared article. On the periphery of the cluster, the less connected authors are shown. By removing the central home page, two clusters can be seen, which are connected only through shared contributions of two authors. On the periphery there are four authors, who only wrote contributions to the main page or only on pages not edited by others, but not on any of the pages co-edited by the authors in the two clusters.



**Figure 3.** Collaboration network excluding the central home page.

This co-authorship visualisation has its benefit in showing who collaborated with whom. It does not, however, show the evolution of the living deliverable over time and it lacks information about the content on which the authors collaborated. This can

be extended by adding pages as nodes to the network and introducing directed editing relationships pointing from the authors to the pages they have changed. With that, authoring relations on particular pages become more salient.

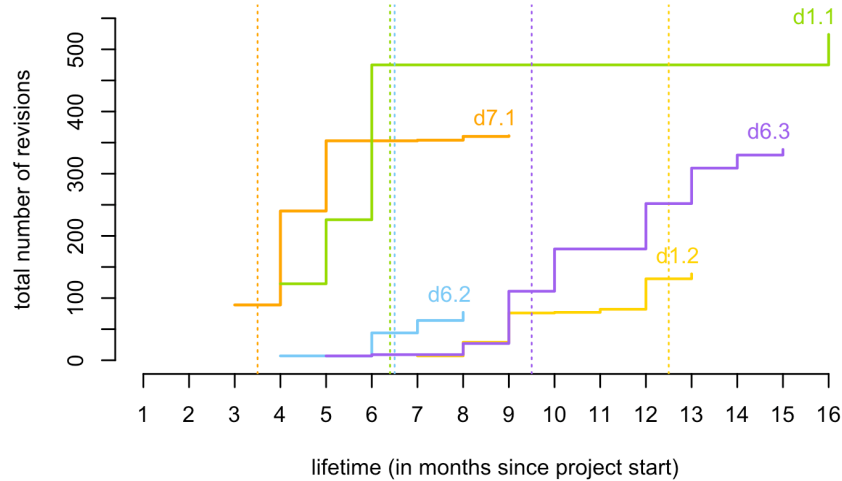
Additionally, the development of the overall number of non-minor edits over time provides information on the vitality of the wiki and complements the analysis.

## **5 Discussion: Is there an afterlife after the deadline?**

The deadline of regular deliverables marks the end of the writing process. After the deadline, the official writing process ends and there is no formal requirement to modify them anymore. As mentioned above, the purpose of living deliverables is to allow for more continuous collaboration beyond delivery deadlines. The assumption behind living documents is that knowledge construction processes are continuous and deliverables are artefacts of an underlying, continuous collaboration process. By turning these artefacts into living documents, they better reflect the dynamic structure of project work, which is somewhat artificially subjected to a project framework in order to allow for efficient and effective management. Not only in networks of excellence, where a consortium faces additionally the challenge to re-organise an open research network beyond the partnership, but also in other research project types, interdependencies of tasks naturally create feedback loops that should inform already 'delivered' work (such as from validation to conceptual design), thus creating an opportunity to update them.

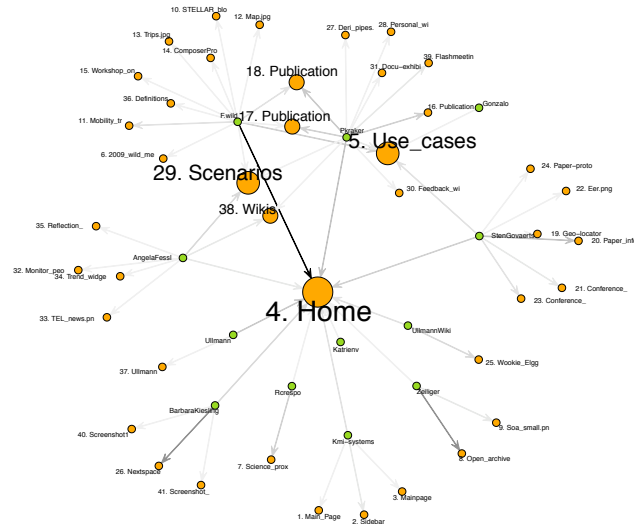
To test whether or not the documents were subject to editing activity also after the submission deadline, we gathered the revisions of each deliverable and cumulated the amount of revisions for each deliverable for each project month. The following line chart shows on the y-axis the amount of revisions and on the x-axis the time frames (16 project months). One deliverable already exists since 13 months, while others are in use for shorter periods of time. The vertical lines at month 3, 6, 9, and 12 represent the submission deadlines.

All deliverables continue their life also after their formal deadline. Even when considering a phase of two months after the deadlines (taking into account possible delays in delivery), still three of the deliverables show lively activity. According to the revision counts, the official deadline raised the number of revisions, while after a deadline the amount of revisions increases mostly less steep. The three deliverables d6.2 (blue), d6.3 (purple), and d1.2 (yellow) show a very steady increase over time, whereas particularly the early deliverables d7.1 (orange) and d1.1 (green) experience their most busy editing processes around the time of their deadline.



**Figure 4.** Total number of edits (cumulated) for each living deliverable.

While the line chart visualisation only shows the frequencies of the revisions over time, it does not provide information about the themes of collaboration and the collaboration network created in the co-editing activity – and how they have changed from before to after the deadline.



**Figure 5.** Authors (green) and their contributions to pages (orange): before the submission deadline.

Figures Figure 5 and Figure 6 show the network of authors and their contributions to pages in d6.3 before and after the submission deadline. While the focus before the deadline is clearly on ‘use cases’, ‘scenarios’ and the main page of the deliverable, the

figure for the network after the deadline shows a change towards more technical topics, like 'Tools', 'Services', and 'Widgets'.



**Figure 6.** Authors (orange) and their contributions to pages (green): after the submission deadline.

The other deliverables show similar patterns of activity: d7.1 again exposes a larger network of pages (but with a smaller number of contributors), where as d1.1 is significantly reduced in the number of contributors (but still showing a larger number of edits). The deliverable d6.2 shows a star pattern of authors editing the main page and d1.2 ceased its activity with its delivery deadline.

## 6 Conclusion and outlook

With the analysis presented, the conclusion can be drawn that there definitely is an afterlife for most of the living deliverables. With only one zombie exception, this afterlife is more like a blitheful continuation of activities – relieved of the duty of having a deadline. At least for the one deliverables we have analysed this in more depth and collaboration beyond the deadline exposes a large co-authorship network, accompanied by shift in focus.

As stated the data are extracted from the public revision histories of the living deliverables, made available by MediaWiki. They can be used to show whether wikis show any signs of editing activity and to further investigate the collaboration network structure expressed in these revisions. It is possible to inspect who is collaborating on particular pages. In large projects, like STELLAR, these visualisations can help to make activities more transparent which can create more awareness and accountability – and ultimately offers triggers for new activity.

For living deliverables as such, it provides a way to check for signs of life, especially when their delivery deadline has passed.



There are several limitations this study has. Most notably, collaboration in co-authoring wiki pages cannot be mistaken for the overall collaboration on the (printed) report delivered to the European Commission. All wikis had phases close to the deadline, where an export of the Wikipages into a Word-file served the final polishing and further elaboration. All the deliverables were embedded into collaborative activities of other nature, such as presence and virtual meetings (flashmeetings), reviews (with separate reports), and other forms of collaboration that left no traces in the wikis. Still they are part of the process of creating their content.

Moreover, we have so far looked at only a small number of living deliverables in a limited time period. It will be very interesting to see, whether our findings will be confirmed when repeated in the future with more data and a longer time frame. Not to mention that it will be interesting to see, whether there is an afterlife of the deliverables beyond the runtime of the project.

It is an open question, whether the analysis method used can be matured into a self-explaining visualisation that does not require any insider knowledge about the collaboration in order to correctly read it. Or in other words: an evaluation of usability and accuracy is pending. This might also be helpful further what (wiki-wise) the difference between a living and living dead deliverable is. And it might help to identify driving factors: is it the medium, the collaborators, or the content?

In its current form, the co-editing network plots depict only a holistic view of all contributions. A more flexible approach would be to let the user interactively choose time windows, thereby providing means to investigate collaboration patterns before and after significant events. An animation of the graph change over time would additionally help to understand the development of a living deliverable, emphasizing the process dimension further.

A more fine-grain distinction of the types of contributions and their drivers would serve further analysis: writing passages, proofreading, enhancing with links and media, discussing, altering, and deleting text are all important for the quality of an article, but possibly not all of them trigger further activity by collaborators. This would be equally interesting for life and afterlife of the deliverables.

Additional evidence sources are available to further investigate collaboration among the researchers outside the living deliverable. It would be very interesting to see whether collaboration patterns differ when looking at the accompanying virtual meetings, e-mail exchange, or presence meetings. Does the medium foster certain styles of collaborations or do they converge?

From a project oriented view the proposed type of analysis could serve as a feedback mechanism making achievements visible. This could help to activate discussion about research collaboration.

## **Acknowledgment**

The work presented in this paper was carried out as part of the STELLAR network of excellence, which is funded by the European Commission under the grant agreement number 231913.

## Reference

- [1] Duncan, W.: Developing a project-management body-of-knowledge document: the US Project Management Institute's approach, 1983-94, In: *International Journal of Project Management* Vol. 13, No. 2, pp. 89-94, 1995
- [2] Arazy, O., Stroulia, E., et al.: Recognizing contributions in wikis: Authorship categories, algorithms, and visualizations. *Journal of the American Society for Information Science and Technology*. 61, 6, 1166-1179 (2010).
- [3] Nunes, S., Ribeiro, C., David, G.: Wikichanges-exposing wikipedia revision activity. *Proceedings of the 2008 International Symposium on Wikis (WikiSym)* Porto, Portugal. (2008).
- [4] Viégas, F.B., Wattenberg, M., Dave, K.: Studying cooperation and conflict between authors with history flow visualizations. *Proceedings of the CHI 2004*, pp. 575-582, ACM, Vienna, Austria (2004).
- [5] Suh, B., Chi, E., Pendleton, B.A., Kittur, A.: Us vs. Them: Understanding Social Dynamics in Wikipedia with Revert Graph Visualizations, In: *IEEE Symposium on Visual Analytics Science and Technology 2007*, IEEE, Sacramento/CA, USA, 2007.
- [6] Baumgrass, A., Mueller, C., Meurath, B.: Analyzing Wiki-based Networks to Improve Knowledge Processes in Organizations, In: *Journal of Universal Computer Science*, Vol. 14, No. 5, pp. 526-545, ISSN 0948-695x, 2008.
- [7] Jesus, R., Schwartz, M., Lehmann, S.: Bipartite Networks of Wikipedia's Articles and Authors: a Meso-level Approach, In: *WikiSum'09*, ACM, Orlando/Florida, USA, 2009

## **Science 2.0: Supporting a Doctoral Community of Practice in Technology Enhanced Learning using Social Software**

*Denis Gillet<sup>1</sup>, Sandy El Helou<sup>1</sup>, Marie Joubert<sup>2</sup>, and Rosamund Sutherland<sup>2</sup>*

<sup>1</sup>École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland,  
E-mail: denis.gillet@epfl.ch;

<sup>2</sup>University of Bristol, Bristol, BS8 1JA, United Kingdom,  
E-mail: Ros.Sutherland@bristol.ac.uk

### **1. Introduction**

The STELLAR European network of Excellence (NoE) (<http://www.stellarnet.eu>) represents the effort of leading European institutions and projects in Technology Enhanced Learning (TEL) to unify their diverse and sometimes fragmented community. The Network is executed via a series of integration instruments designed to increase the research capacity of European TEL at all levels. One level is that of early career researchers, and in particular doctoral candidates, who are seen as key in establishing sustainable research capacity in TEL.

This paper focuses on the ways in which STELLAR supports doctoral candidates through the establishment of a Doctoral Community of Practice (CoP) in Technology Enhanced Learning as a STELLAR doctoral integration instrument for the doctoral stakeholder community. This TEL Doctoral CoP (DoCoP), officially established in Autumn 2009, is also instrumental in bringing together actors of engineering education research in academic institutions, as proposed in the USA by Streveler [1]. The paper discusses possible ways in which the DoCoP could be developed through the innovative use of Web 2.0 technologies, by outlining the characteristics of one such technology and describing the ways in which an imaginary PhD candidate might use the technology in their PhD journey.

### **2. Doctoral Community of Practice in TEL**

In STELLAR a high priority is given to PhD candidates, i.e. early career researchers conducting their PhD research in higher education institutions and typically being enrolled in a doctoral program. PhD education is a key issue for strengthening TEL research, shaping the domain and preparing the next generation of researchers and TEL entrepreneurs. This is especially true in European countries where the majority of PhD candidates, particularly those involved in the more technical aspects of TEL, move immediately to companies after graduation. As such, the integration of PhD candidates into STELLAR creates the conditions for the long-term sustainability of TEL research and TEL impact on the economy.

The first instrument to support the doctoral community in TEL is the STELLAR Doctoral Academy, which includes selected face-to-face events that bring the community together for doctoral courses, workshops and seminars. Such events typically last between a few days and one week. The organized workshops and lectures focus on theoretical, methodological and technological issues of relevance to TEL research (for example activity theory, embodied cognition, individual, social & organizational learning processes, human computer interaction, Web mash-up, adaptive and personal learning environments, as well as knowledge and competence management). Organization committees can apply to get an event sponsored by STELLAR. Upon acceptance, the sponsorship acts as a quality label (STELLAR accreditation) recognizing excellence and attracting participants who may be awarded STELLAR scholarships for participation. Sponsorships are granted to events gathering the top experts in the field and showing a high potential for the integration of the multidisciplinary and multicultural TEL community. Experts, PhD advisors and PhD candidates meet in such events that rely on Science 2.0 infrastructure to be established and for sustaining further interaction between the participants. Science 2.0 infrastructure stands for Web 2.0 solutions applied for scientific purposes and enabling interaction between peers, as well as sharing of resources offered by the community for the community.

The second instrument is the Doctoral Community of Practice (DoCoP). Its purpose is to bring or keep together PhD candidates working on TEL research, offering them opportunities to share, discuss, and receive feedback on their research by peers and experts. In this respect, as the doctoral candidates ('newcomers') interact with each other and with expert researchers ('oldtimers'), they are learning how to become TEL researchers [2]. As Wenger [3] has explained, a community of practice has three characteristics, firstly a shared domain of interest (in this case TEL research), secondly a community in which members interact and learn from each other, and thirdly the practice "members of a community of practice are practitioners). They develop a shared repertoire of resources: experiences, stories, tools, ways of addressing recurring problems—in short a shared practice. This takes time and sustained interaction". The DoCoP is supported by an on-line platform specifically designed for exchanges between doctoral candidates and researchers in the TEL field as part of the STELLAR Science 2.0 portal. The STELLAR Science 2.0 portal is presented in Section 3 below. A social software that could be integrated in the portal as an interaction space for the DoCoP is introduced in Section 4 and detailed in Section 5.

In creating and sustaining a doctoral community, the DoCoP has the potential to support doctoral candidates in three key ways. **First**, it can support researchers who may feel isolated, who work in small labs without a critical mass of PhD candidates or without staff with interdisciplinary skills, by integrating them into the STELLAR community in particular, and the TEL community in general. **Second**, the DoCoP could also serve as a platform for doctoral candidates and their advisors to find informal or formal co-advisors or jury members for the PhD thesis. **Third**, it is possible that, through engaging in the DoCoP, PhD candidates will be able to identify key research groups and companies whom they might want to visit as part of a mobility programme brokered by STELLAR. Such visits will help doctoral candidates to integrate collaborative research or testbed validation in their PhD thesis, and will also contribute to STELLAR's aim of developing capacity beyond institutional boundaries.

### 3. STELLAR Science 2.0 Framework

STELLAR adopts the perspective that in establishing effective communities of practice and stronger research communities in TEL, mutual learning of the participants is enhanced by supporting knowledge sharing and facilitating the advancement of one's own competences and knowledge while researching in TEL. In order to scaffold these processes, STELLAR uses the STELLAR "Science 2.0" concept which federates a variety of communication channels to ease internal exchanges within the scientific STELLAR network as well as beyond it to reach the general TEL community. These communication channels also aim to make research results known to the European and international TEL community.

STELLAR opens its framework and instruments to interlink with relevant scientific communities and people, and to support knowledge construction within the TEL research area. For this purpose, it aims at providing efficient access to the research outcomes of colleagues and peers in one's own and related disciplines, using Web 2.0 approaches that leverage the links between people as well as the links between documents, using a combination of pull and push techniques.

The STELLAR Science 2.0 portal is grounded within the context of the above objectives, integrating Web 2.0 tools, services and social software publically available or offered by the STELLAR community members. The Web of the second generation adds collaboration and communication mechanisms to scientific research resources by designing and implementing tools to support, follow and track discussions, argumentations and the whole history of a process which participates in the building of science. Blogs, forums, Wiki-pages and RSS feeds have become popular means for lightweight exchange, discussion and syndication of knowledge and opinions. Advanced TEL-oriented social networking or communication tools will be integrated together with analysis tools that will enable evaluation of the usage of the portal and the evolution of the community.

As mentioned in Section 2, the doctoral stakeholder community and its related doctoral community of practice are central to reducing fragmentation of research in technology enhanced learning in Europe. As a consequence, special components of the Science 2.0 portal have to be designed and dedicated to the DoCoP. These components aim at virtually complementing and sustaining interaction between face-to-face events such as TEL doctoral schools, workshops, and doctoral consortia.

An informal survey carried out during the 2009 Joint Summer School on TEL (a doctoral event sponsored by STELLAR) has demonstrated that the doctoral community members in TEL use a huge variety of tools, resources and repositories. It has also highlighted the needs for better integration and sharing in contexts and on purposes. STELLAR does not aim to create an additional tool to meet this need, but to adapt solutions developed in previous European projects to integrate in a space of the Science 2.0 portal all the digital and social ingredients of a successful PhD recipe in TEL. For example, the *Graaasp* social software developed for supporting Communities of Practice in the framework of the European PALETTE research project under the *eLogbook* name appears to fulfill simultaneously the need to have a single aggregation place for global resources and to have a unique personalization mean to organize such aggregated resources by contexts and by purposes. In addition, *Graaasp* also supports workflow management as requested when implementing and conducting doctoral school events or visits supported by mobility scholarships. It may be that STELLAR will adapt this software for use by the DoCoP. The remainder of this paper explores how *Graaasp* might fit with STELLAR's agenda.

#### 4. Constructing a Web 2.0 DoCoP interaction space

Successful Web 2.0 solutions result from the implementation of a proper participatory design approach and also from the recognition that the Darwin theory of evolution applies to them. In other words, only the Web 2.0 tools and services (species) that fit to the environment (user's contexts, expectations, and adoption thresholds) can survive and spread. They spread thanks to mass adoption, open source licensing, shared features or APIs (as the DNA of the best individuals).

Preliminary participatory design results can be derived from the analysis of the practice in the target community; in our case the community of PhD candidates and advisors in TEL. The objective is to define and construct an interaction space (component) of the Science 2.0 portal for supporting the DoCoP, as well as to improve the features of the *Graaasp* social software, which is a candidate solution as Web 2.0 DoCoP interaction space. The robustness of the Science 2.0 portal to the natural evolution of the Web 2.0 solutions is not considered here. It is supported practically by relying on standards and choosing solutions as open as possible, either for being integrated as component or feeds in other platforms, or for integrating other components or feeds. In effect, in the Web 2.0 realm, people and platforms are simultaneously consumers and providers.

For design purpose in our Science 2.0 context, it is important to reflect on the practice of PhD candidates in TEL, in terms of social networking, communication and online resources. As a seed for reflection, one can ask if PhD research in TEL is different from PhD research in general, and, if yes, in what sense and what are the consequences in terms of practice. There are at least two distinctive features of TEL research. First, its youth and second its interdisciplinary nature.

One impact of the youth of TEL research in terms of practice is that it is difficult to find and to assess relevant scholarly resources and experts. Workshops, conferences and journals are not yet fully established, indexed and recognized; thus putting more burden on the candidates' shoulders and their communities to point out relevant material of good quality. In view of this difficulty, the STELLAR Science 2.0 portal should direct PhD candidates towards the proper scholarly places to find references and also to publish their work. As a matter of fact, collaborative recommendation driven by the TEL community through the portal should give access to the most relevant scientific contributions in the field. Also, best practices on what a PhD thesis in TEL is have still to be fully established. The portal should point to all published PhD thesis in TEL and enable interaction and communication with TEL researchers, TEL experts, TEL PhD advisors, and TEL Research Alumni.

The impact of the interdisciplinary nature of TEL research is quite deep [4]. First of all, it is still difficult to carry out research tagged as TEL in most academic institutions. Moving from "Thesis in Computer Science with Application to TEL" or "Thesis in Education with Application to TEL" towards "Thesis in TEL with Application to ... Education in Computer Science (as example)" requires the full leveraging of the TEL research community. In that sense, supporting a strong visibility of the DoCoP senior and successful faculty members through the STELLAR portal is instrumental to enabling high-level PhD research in TEL. As TEL research requires multidisciplinary competences infrequently owned by a single individual, the role of the community as a coach to PhD advisors can be seen as stronger in TEL than in other research fields. This shows again the importance of supporting the DoCoP properly and highlights the importance of interaction in the community, as well as peer or

expert recommendation, possibly supported through the Science 2.0 portal.

Research in TEL is a niche that can be expanded considering the closeness between TEL and modern knowledge management practices. Currently, as a niche, in most institutions there is not the critical mass to establish a doctoral program in TEL. As a matter of fact, we claim that it should be avoided to fully build on the interdisciplinary nature of TEL. Hence, the DoCoP and the STELLAR science 2.0 portal should enable the operation of a virtual, distributed and informal doctoral program in TEL. This should integrate the existing face-to-face doctoral events only accessible to a subset of the DoCoP members and the associated resources. This program should be informal in the sense that it should not be built as a competitor of institutional programs, but as an additional resource for which the client institutions should grant ECTS credits themselves. Associated resources offered through the STELLAR Science 2.0 portal include pointers to *YouTube* or *FlashMeeting* videos of talks, seminars and lectures, as well as the associated slides and handouts available for example on *SlideShare*, where they can be tagged and rated.

## 5. The *Graaasp* Web 2.0 social software and an imaginary PhD journey through it

The *Graaasp* social software envisioned as an interaction space for the DoCoP can be described as a Web 2.0 application that can serve simultaneously as an aggregation, contextualization, discussion, and networking platform, a shared asset repository, as well as an activity management system.

*Graaasp* is built on the 3A interaction model [5] which is particularly focused on describing and designing social and collaborative environments. It was developed in the framework of the Palette European Project (<http://palette.ercim.org/>) following a participatory design approach: interviews and questionnaires with communities of practice such as *Learn-Net*, *Doctoral Program Lancaster*, *InCorPorate* and *Adira* helped identify their needs and translate them into design requirements. The 3A model accounts for three main constructs or entities: **Actors** are entities capable of initiating an event in a collaborative environment. They can be humans as well as virtual agents. Actors create collaboration spaces where they conduct **Group Activities** to reach specific objectives. In each of these activities, actors can take different roles, each of which consisting of a label and an associated set of rights. Furthermore, Actors produce, edit, share and annotate **Assets** in order to meet activities objectives. Assets can consist of simple text files, RSS feeds, wikis, videos or audio files. In addition, an activity can possibly have a well-defined planning of expected assets with concrete submission and evaluation deadlines, predefined evaluators and submitters. This is particularly useful in project-based learning communities and online educational environments. The model accounts for Web 2.0 features: entities can be tagged, shared, commented, linked together and rated. By design, *Graaasp* can serve not only as a networking platform, a repository of assets and an activity management system, but also as an aggregator bringing together content and services from other Web 2.0 applications. Internal tracking and notification features enable *Graaasp* to provide awareness to users on ongoing activities and participation.

To understand the potential role of *Graaasp* in enhancing the learning experience of PhD candidates in TEL, their interaction with one another as well as with senior researchers, we examine a scenario about an imaginary person, namely Pat, who has just started her PhD in the TEL field. Pat is invited by her PhD advisor and/or colleagues to register to *Graaasp*. Just like any other Web 2.0 social software, *Graaasp* has a low entry barrier; registration only requires a valid email and a password, or an OpenID. With time, Pat's profile can be gradually built up.

As initial action, Pat types in the search field “trust and reputation in Web 2.0 applications”, which will be the subject of her thesis. As a response, the system relies on content and linked-based analysis techniques to propose a list of relevant group activities, assets and actors. For example, *Graaasp* proposes a LinkedIn group entitled “Trust & reputation in Web 2.0”. Pat decides to join the group. Consequently, *Graaasp* uses the LinkedIn application programming interface (API) to send an adequate Web service request on her behalf. As assets, the system returns relevant documents, archived discussions threads saved and annotated in *Graaasp*, embedded *YouTube* videos such as conference talks on trust and reputation in Web 2.0, external papers from the IEEE and ACM digital libraries and the STELLAR open archive (<http://www.telearn.org>), as well as aggregated new feeds relevant to Pat's query. It also recommends a list of actors that are keen on the requested topic, ordered by relevancy and reputation. Actors include appropriate PhD candidates and senior researchers that are already

*Graaasp* users, as well as external people that have written relevant papers and/or participated in relevant group activities. For actors who are logged in to *Graaasp*, the system shows presence awareness to encourage interaction. Since actors do not only consist of people but also of agents and tools, *Graaasp* also suggests to Pat useful online systems related to TEL. The system also asks Pat if she wishes to be notified about the creation of any new activities and assets relevant to “Trust and Reputation in Web 2.0”. Pat responds in the affirmative.

In the mean time, a professor from Germany creates a group space called “Privacy, trust and reputation challenges in Web 2.0” in order to conduct collaborative activities related to the subject. The activity space is public and anyone can become a member. Pat is notified of the creation of the space, and decides to join it. She takes the role “PhD candidate”. Her membership is announced to other members. Another senior “PhD candidate” takes the initiative of opening a conversation with her and gives her hints on how to start learning about the field as well as references to assets (i.e. discussions, papers, online course notes) and group activities that can best introduce to the field. They also speak about the PhD process in general. Seeing that the discussion is interesting, Pat decides to save it as an asset, post it in the activity space and tag it “tips for beginners”, “TEL”, “PhD in TEL”. It will serve as a reference to her and other new PhD candidates. Afterwards, Pat is notified of the creation of a new sub-activity space within “Privacy, trust and reputation challenges in Web 2.0” dedicated to a summer school whose topics and application process are described in the space wiki. A plan for expected assets is created, specifying submission and evaluation deadlines. By a simple click, Pat downloads the submission deadline to her calendar. She intends to work hard to be able to develop her knowledge in the field, submit a position paper and eventually participate in the summer school.

After a discussion with Pat, her thesis advisor decides to create an activity in *Graaasp* that helps his candidate progress her PhD in TEL and facilitates follow-up. They define for this activity a plan of expected assets. First, Pat is expected to read material related to TEL, raise and discuss the challenges that she thinks are important to consider and solve. These discussions are to be submitted as assets for the thesis advisor. Once the submission is done, the advisor is notified by the system. He then reviews and comments the submitted asset before a scheduled face-to-face meeting. After the meeting, Pat uploads minutes of meeting report and links it to the asset that triggered the discussion.

It is worth mentioning that the system allows users to enforce an order in the submission and evaluation of expected assets. This means, for example, that the research plan cannot be submitted before the submission and the successful evaluation of the asset discussing TEL challenges. Some time later, and always with respect to her PhD progress, Pat types in the search field “PhD dissertation”. *Graaasp* proposes *AWSOME Dissertation*, an online project that other users have registered as a tool and tagged with the keywords “PhD dissertation” [6].

Last but not least, as Pat uses *Graaasp* more regularly and visits other actors’ profiles, she starts seeing the added value of completing her own profile information. She fills in information related not only to her academic background and current position, but also to the skills and competencies that she already has and the ones that she aims to acquire or develop. Pat also augments the profile of other actors she interacted with by giving personal feedback through comments and ratings. As is the case for tags, the system suggests skills and competencies that have already been added by other actors, to Pat. This helps her discover skills worth developing, build a common TEL vocabulary, and/or reach people with similar learning objectives. As Pat’s profile is gradually completed, and as she uses the collaboration platform more frequently, the recommendations rendered by the system become more and more personalized. As a matter of fact, as the system suggests resources, actors and group activities to Pat depending on her working context, it draws on information about Pat’s previous interactions to discover her preferences, her trusted network of actors, and the kind of actors, resources and group activities that are best fitted to her learning needs and goals.

Figure 1 below provides a mock up of the *Graaasp* interface, showing with a large orange banner and as current context the activity created by Pat. It is entitled “Pat’s PhD Research Follow up” (central left-hand side) and integrates a wiki with the description of the activity and a field for free comments that could be added by users having sufficient rights to access it. The current context is selected from the Favorites or the Clipboard area by clicking on the desired color rectangle. It could also be chosen among the results of a search query. Once an entity is selected as current context, related personal entities explicitly associated to the current context by the user are automatically displayed in three dedicated columns (central right-hand side). Hence, *Graaasp* can be seen as a contextual browser showing in a single screen all the relevant information aggregated by the user in the current context.

Feeds relevant for Pat's PhD research are only visible in this context, not in her other space dedicated to the theatric performance she is organizing with friends. In addition to these preselected entities, recommendation of external ones can be provided taking into account their existing relations with the current context and their relative importance in it (bottom left-hand side) [7]. By clicking on any rectangle, the user automatically trigger a change in context and all the interface components are updated. New relations can be created by dragging and dropping entities or by clicking on the relevant grey rectangles.

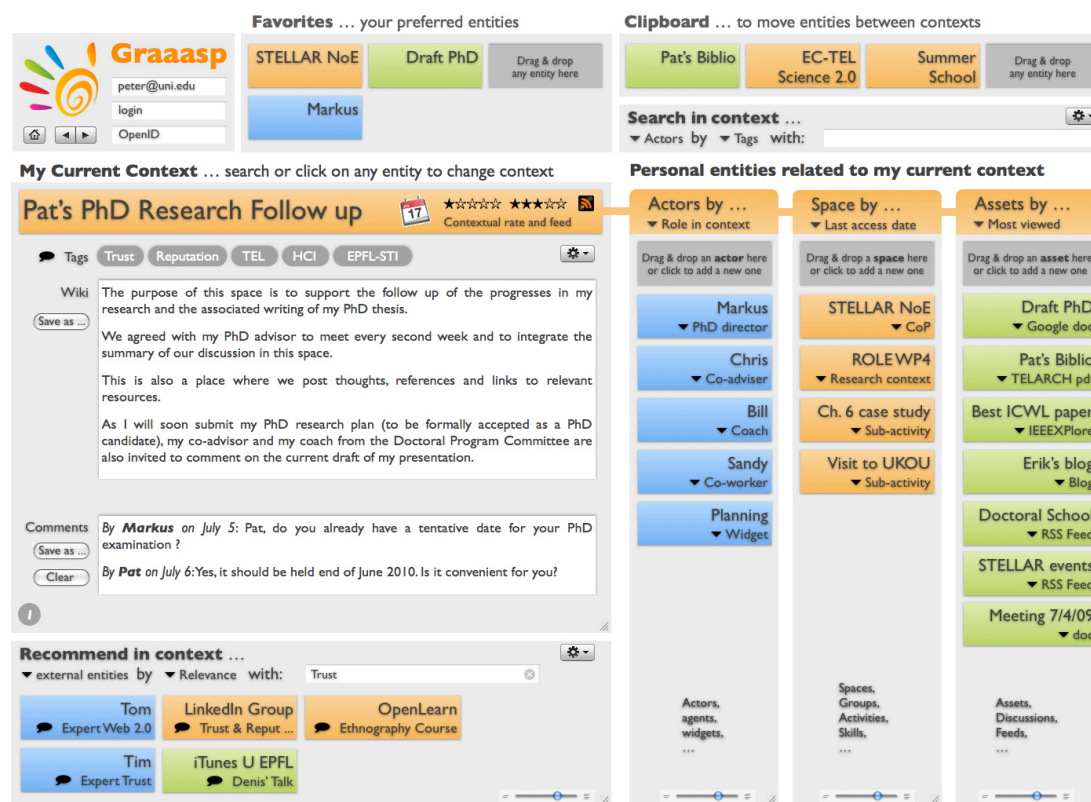


Figure 1. Recent Mock-up of the *Graaasp* interaction space.

#### 4. Concluding remarks

This paper has argued that TEL doctoral candidates have a particular need for support, and has described the structure of STELLAR's organizational commitments to provide this support in the form of doctoral events and a Doctoral Community of Practice (DoCoP). It further argued that the DoCoP should support its members through mutual learning and knowledge building, introducing the *Graaasp* social software, and explaining through an illustrative example how Web 2.0 features could be used to meet the aims of the DoCoP.

As a matter of fact, supporting a Doctoral Community of Practice (DoCoP) in Technology Enhanced Learning (TEL) is more a question of pointing to the right resources, services and people from a single place than a question of platform and technology. However, because of the large set and the variety of the digital assets and online communities involved, the way these entities are integrated plays a key role for adoption, appropriation and identification. In effect, building a strong community identity is also a must to position TEL in the worldwide research arena and to strengthen exchange in its multidisciplinary community. A single Web 2.0 interaction space bringing together the most relevant people and resources is instrumental in that perspective.

We believe that *Graaasp*, or similar social software, could provide an online environment that would support doctoral candidates and would provide the 'space' where the members of the DoCoP would interact, building on their relationships initiated and consolidated at face-to-face doctoral events. We believe it is important to work towards developing this software so that it becomes the first choice of



the members of the DoCoP, so that the space becomes well populated with group activities, actors and assets. We recognize, however, that populating such a tool depends on the commitment of the individuals involved and perhaps the greatest challenge for the STELLAR network is to encourage committed doctoral candidates and academics (such as their supervisors) to ‘get the ball rolling’. We suggest that once the ‘ball is rolling’, the value of the tool will grow and be recognized by all individuals in the DoCoP, in STELLAR and in the network of stakeholders.

### Acknowledgements

This paper integrates parts of the STELLAR project description with contributions from (in alphabetic order) of N. Balacheff, S. Barnes, R.M. Bottino, D. Burgos Solans, U. Cress, P. Dillenbourg, E. Duval, F. Fischer, D. Gillet, E. Herder, B. Kieslinger, M. Kravcik, S. Lindstaedt, W. Nejd, M. Ney, P. Scott, M. Sharples, H. Spada, M. Specht, R. Sutherland, P. Tchounikine, C. Windrum, and K. Wodzicki.

The STELLAR project is partially funded by the Seventh Framework Programme for Research and Technological Development (FP7) of the European Union in Information and Communication Technologies.

### References

- [1] Streveler, R.; Smith, K.; and Miller, R. 2005. Enhancing engineering education research capacity through building a community of practice. In Proceedings of the ASEE annual conference and exposition.
- [2] Lave, J.; and Wenger, E. 1991. *Situated Learning, Legitimate Peripheral Participation*. Cambridge University Press, Cambridge.
- [3] Wenger, E. 2006. *Communities of practice: a brief introduction* (<http://www.ewenger.com/theory>). Accessed September 6, 2009.
- [4] Pallas, A.M. 2001. Preparing Education Doctoral Students for Epistemological Diversity *Educational Researcher*, Vol. 30: pp. 1-6.
- [5] El Helou, S.; Gillet, D.; Salzmann, C.; and Rekik, Y. 2009. Software for Sustaining Interaction, Collaboration and Learning in Communities of Practice. In *Solutions and Innovations in Web-Based Technologies for Augmented Learning: Improved Platforms, Tools, and Applications*, *Advances in Web-based Learning (AWBL) Book Series*. pp. 300-316.
- [6] AWESOME Dissertation. *Academic Writing Empowered by Social Online Mediated Environments* (<http://awesome.leeds.ac.uk>). Accessed August 22, 2009.
- [7] El Helou, S.; Salzmann, C.; Sire, S.; and Gillet, D. 2009. The 3A Contextual Ranking System: Simultaneously Recommending Actors, Assets, and Group Activities. In *Proceedings of the 3rd ACM Conference on Recommender Systems*, New York, USA, October 22-25.

---

# Connecting Early Career Researchers: Investigating the Needs of Ph.D. Candidates in TEL Working with Web 2.0

Nina Heinze<sup>1</sup>, Marie Joubert<sup>1</sup>, Denis Gillet<sup>1</sup>

<sup>1</sup> STELLAR Network of Excellence

n.heinze@iwm-kmrc.de, marie.joubert@bristol.ac.uk, denis.gillet@epfl.ch

**Abstract.** This article describes the results of a case study conducted amongst 21 doctoral candidates and three senior researchers at the Joint European Summer School on Technology Enhanced Learning 2010. The study aims to analyse the needs of early career researchers working within the field of TEL in geographically distant communities, particularly with respect to online collaboration, communication and information exchange. This study can be seen as a needs analysis on support structures to enable research 2.0 in TEL among young researchers.

**Keywords:** communities of practice, Research 2.0, Social Media, Web 2.0, case study, awareness support

## 1 Introduction

Our personal experience suggests that collaboration and communication within the European TEL community usually looks like this: researchers use many offline and web-based tools to work and to share their findings and opinions, there is no standardised way of communicating, and various channels are used to disseminate information. It is difficult to keep up with who is doing what in the field, though many researchers are making a considerable effort to monitor the data that is being spread on the Web by colleagues [1], [2], [3]. Ph.D. candidates new to the field frequently have problems finding relevant information, people, events and platforms to help them in their research endeavours. Recent talks with a number of Ph.D. students we are in touch with have underlined these perceptions.

Some efforts have been undertaken to make it easier for doctoral candidates to stay up-to-date on current topics and events and to enable them to collaborate online. These include the establishment of inter- and transorganisational mailing lists,

newsgroups, social media groups or forums<sup>1</sup>. Despite these efforts, however, anecdotal evidence from our discussions with Ph.D. students indicates that doctoral candidates still feel that support in terms of information and collaboration could be improved. To address these concerns, the STELLAR Network of Excellence<sup>2</sup> supports doctoral events that aim to improve collaboration and communication between junior and senior researchers as well as enhance the flow of information. In addition, STELLAR also plans to create a virtual doctoral community of practice (DoCoP) to help Ph.D. candidates stay in touch, share and conduct research, help each other solve problems and get in touch with further junior and senior researchers by means of Web 2.0 technologies, the latter being nowadays referred as social media. We understand Communities of Practice (CoP) to be a group of people who share the same interests and passion for something they do and shape their identity by a shared domain of interest whilst engaging in activities around this domain with other members of the community. They thereby develop a shared repertoire of resources, a shared practice, as Wenger calls it in his explanation of a CoP [4]. For an overview of the implications of CoP's on learning and the possibilities of online CoP's see [4], [5], [6].

We saw it necessary to develop an understanding of the needs of Ph.D. candidates as the starting point for the development of the DoCoP planned in STELLAR. Our first step towards developing such an understanding was to consult with Ph.D. candidates.

An opportunity to do so arose at the 2010 Joint European Summer School on Technology Enhanced Learning, which took place in June 2010, gathering together about 50 Ph.D. candidates working in TEL. We conducted a workshop with focus on students' views on the creation of a doctoral community of practice in the field of TEL. 21 doctoral candidates as well as three senior researchers participated in the workshop. We asked them about what type of information may be of value to them to increase awareness in terms of collaboration, what type of awareness support would be of use to them, what tools they use when collaborating in dislocated research teams and how they believe a sustainable community of practice can be implemented. We report about our findings below.

## **2 Consulting on a DoCoP with Ph.D. Candidates in TEL – A Case Study**

During the workshop at the Summer School the doctoral candidates worked in groups of 5-6 people and were asked to discuss how they would wish to receive support for their doctoral work in terms of personal support, awareness support, tools for collaboration and the characteristics of a doctoral community of practice that would be of value to them. Each group then presented their findings, explained their results and engaged in discussions about their thoughts with the other participants of the

---

<sup>1</sup> Examples include JTEL Summer and Winter Schools, Doctoral Consortia at conferences like EC-TEL or Earli, the STELLAR Mobility Programme or DocNet from the University of St. Gallen, Switzerland

<sup>2</sup> <http://www.stellarnet.eu>

workshop. We recorded the entire session to be able to further analyse the results after the Summer School.

## 2.1 Results of analysis of needs of Ph.D. candidates in TEL

We analysed their reported needs and categorized them into two levels, each describing the personal involvement or gain of the individual researcher (see Table 1). The *individual level* of needs describes issues that occur on an individual level like review of one's own paper or managing one's own information. Support on this level aids the individual in her endeavour more than it does a larger peer-group. The *community level* is the actual community or peer-group level. Support on this level is useful for more than the individual researcher. A larger CoP would benefit from assistance on this level. Table 1, below, summarises the findings within each of these two categories.

**Table 1.** Needs of doctoral students on the individual and community levels.

Individual Level	Community Level
Peer-review of artefacts	Information modelling
Methodology	Researcher information
Problem solving	Futuregazing
General feedback	Networking
Jobs / internships / exchange programmes	Guidelines for community management
F2F meetings	Sharing testbeds / datasets
Information management	Peer groups
	Collaborative filtering

As we can see from Table 1 doctoral candidates would, on the one hand, appreciate support on a very individual level concerning the process of finishing their Ph.D. thesis like advice on the methodology they are planning to use, how to solve problems they encounter when doing their research as well as meeting face to face with a senior scientist to discuss their work to be able to better evaluate if they are on the right track. On the other hand, doctoral candidates see the need for a community of peers working in related fields to network, discuss their work, get a notion of where others in the field are, what their work is about and how they cope with writing a Ph.D.. In addition they would like to get feedback from a community of peers on their work and share research findings and data.

When we asked them about how they believe they can be supported in their endeavours and needs on a technical level we received answers related to information gathering like RSS feeds from relevant sites, collaboration tools like a semantic wiki with an ontology as well as information filtering tools like recommender systems and a reputation system to enable them to better match the information with their current needs. The proposed solutions Ph.D. candidates gave revolve around support issues that have a high technical (system) component. They require the provision of some sort of Web 2.0 tool or are in essence already a tool.

What we can see from the distinction we made is that the categorization of needs in two levels is not a sufficient distinction, since some issues on the individual and community levels are at the same time themes that fall into the area of proposed solutions like networking or sharing testbeds. This is not a surprise, though, since communication, collaboration and awareness of a community go hand in hand.

## 2.2 Results of awareness support of Ph.D. candidates in TEL

In addition, we asked the 4 groups to consider what kind of awareness support may be helpful in research communities with respect to contributing to increased productivity. With awareness we mean the state or quality of being aware of the current themes, projects, events and researchers including their background within the field of TEL and one's own position within it. Again they discussed within their groups and presented their findings in a plenary.

We analysed the plenary discussions and were able to place the findings into two areas. The first area, *personal*, pertains to information available on the personal/professional background of other researchers and contains topics like research background or projects that the person has worked on. The second area of interest in awareness support, *research*, concerns information on the actual output of researchers (artefacts like publications) as well as opinions of others about them. Table 2 sums up the awareness support results of the case-study participants.

**Table 2.** Awareness support

Personal level	Research level
Research background	Artefacts / publications
Expertise / Competencies	State-of-the-art of topic
Projects	Opinions from peers
Social media handles <sup>3</sup>	

Table 2 shows that doctoral candidates wish to have personal information on people within their area of research in terms of scientific background and expertise, as well as their online handles like Twitter and delicious user names or blogs. On the research level they suggest information on current artefacts and publications, as well as the state-of-the art of research in their field and opinions from peers on research, publications and other researchers.

When asked about technical solutions to make it possible to gather and filter information within the community to increase one's awareness of the field of TEL in terms of people, topics, and events, the Ph.D. students proposed open-source solutions to share datasets as well as reputation mechanisms to increase awareness of and within the TEL community. However, the results on the tool level were low which we believe is due to the fact that there are few good services available and the time we gave the doctoral candidates was too short to come up with productive and creative feedback.

---

<sup>3</sup> social media handels are usernames for social media services like Twitter, Delicious, Slideshare or URL's to blogs or wikis

### **2.3 Suggestions for the creation of a doctoral community of practice by Ph.D. candidates in TEL**

The last part of the workshop revolved around collecting ideas on how a sustainable virtual doctoral community of practice (DoCoP) amongst former and future Ph.D. candidates participating in STELLAR doctoral events could be established and maintained. We saw a key consideration within this discussion as the tools used to support the DoCoP. Further, participants were also asked which Web 2.0 tools they use in their own practice and for what purposes in order to inform our understanding of what they value. This discussion, again, took place amongst the whole group.

Our analysis of the discussions led to three main results. The first is that the participants in our case study find it unlikely that a larger doctoral community of practice can be sustained in a reasonable manner by itself. Their experience is that events such as, for example, the Summer School, function as an umbrella, or a macro-level of community, out of which several smaller, actual communities of practice arise with about 6 to 10 members. The Ph.D. candidates suggested that these smaller communities of practice should be supported not by a particular tool or service, since the community members would decide on those depending on their needs and habits, but rather by the provision of guidelines on collaboration, including the use of existing Web 2.0 tools for research and community management.

The second conclusion the participants drew was that the sustainability of a community of practice, based on the philosophy underpinning Research 2.0, would be highly dependent on individuals dedicated to it. They concluded that the community is independent of the tools in the sense that tools are used regardless of the community. Participants recommended a community facilitator to keep the flow of information going and the community members active in participating.

The third conclusion was that the tool or service needs to fulfil collaboration and communication functions and should be user-friendly in the sense that it is easy to use. The doctoral candidates already use a number of tools for these purposes as well as for research and the organisation of their projects, they did not see the pressing need for a “new” tool or platform.

Table 3, below, summarises the participants’ reported use of Web 2.0 tools for communication, collaboration, research instruments and organisation.

**Table 3.** Tools used by case-study participants

<b>Tool</b>	<b>Communication</b>	<b>Collaboration</b>	<b>Research</b>	<b>Organization</b>
E-mail	x	x		x
Google Docs		x		
Google Talk	x			
Google Scholar			x	
Google Analytics			x	
Google Forms			x	
Google Sites	x	x		x
Google Wave	x	x	x	x
BSCW	x	x		x
Dropbox		x		

Mendeley		x	x	
Group Wikis	x	x	x	x
FlashMeeting	x	x		
Skype	x			
MSN Messenger	x			
Doodle				x
Gigapedia			x	
Library			x	

### 3 Conclusions

The results show that Research 2.0 in a doctoral community takes place on many different levels and involves quite a few issues that need to be taken into account. For one, Ph.D. candidates spend time working alone, independently on their thesis and would value support on a very personal, face-to-face level from senior researchers. Further, doctoral candidates appreciate a community of peers they can discuss problems with, share results and remain up-to-date on what is happening in their field of research. They would like to have tools that make it easier for them to gather information on relevant researchers to their topic, important events, possibilities for scholarships and internships as well as collaboration tools like a semantic wiki to collaborate and share findings. In addition, doctoral candidates find an awareness support system useful that allows them to see how is doing what in the TEL community with whom.

In terms of creating a sustainable doctoral community of practice within the field of TEL we could distinguish two main findings: we have a large, fuzzy community of TEL researchers and Ph.D. candidates. Bringing them together in one virtual doctoral community of practice and having them all collaborate and communicate seems unlikely. However, this large community is in need of a virtual space that collects information, makes it available to others and has mechanisms to share that information to increase awareness of the community and bring it closer together. This type of umbrella-platform can enable the smaller communities within the field of TEL gather under the same roof, form and proliferate and share information within the smaller communities as well as the larger TEL community.

Our second conclusion is that there seems to be little need to develop a super-tool to fulfil the needs of Ph.D. students to work and collaborate in their community. What we could see is that doctoral candidates use tools for collaborating, communicating, conducting research and organizing their work flow processes and information. There is little need for yet another tool according to the workshop participants. In addition, the participants noted that preferences as well as needs differ, so tool choice should be left up to the Ph.D. candidates. Rather, there is a need for guidelines on existing tools and their use for research.

In summary, we can say that the findings from the workshop we conducted lead to the conclusion that Ph.D. candidates working within the field of TEL feel they have sufficient Web 2.0 tools at their disposal but would appreciate more support in terms of their use as well as finding and filtering information relevant to their research.

## **References**

1. Reinhardt, W., Moi, M., Varlemann, T.: Artefact-Actor-Networks as tie between social networks and artefact networks. In Proceedings of the 5th International Conference on Collaborative Computing (CollaborateCom'09) (2009)
2. Ochoa, X., Mendez, G., Duval, E.: Who we are: Analysis of 10 years of the ED-Media Conference. In: Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (ED-Media 2009), pp. 189-200 (2009)
3. Fisichella, M., Herder, E., Marenzi, I., Nejd, W.: Who are you working with? Visualizing TEL Research Communities. In: Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications (ED-Media 2010), pp. 2010 (2010)
4. Wenger, E.: Communities of Practice: Learning, Meaning, and Identity. Cambridge University Press, New York (1998)
5. Wenger, E., McDermott, R., Snyder, W.: Cultivating Communities of Practice: A Guide to Managing Knowledge. Harvard Business School Press, Boston (2001)
6. Wenger, E., White, N., Smith, J.D.: Digital Habitats: Stewarding Technology for Communities. CPSquare, Portland (2009)

## **Acknowledgement**

This case study was carried out within the STELLAR Network of Excellence under the European Seventh Framework Programme (FP7).



# MODELLING A STAKEHOLDER COMMUNITY VIA A SOCIAL PLATFORM: THE CASE OF TELEUROPE.EU

Noaa Barak<sup>1</sup>, Daniel Burgos<sup>1,2</sup>, Anthony Fisher Camilleri<sup>3</sup>, Fred de Vries<sup>4</sup>,  
Marcus Specht<sup>4</sup>, and Caroline Windrum<sup>5</sup>

<sup>1</sup> Atos Origin, Albarracín 25, Madrid 28037, Spain, <sup>2</sup> International University of La Rioja. G. V. Rey Juan Carlos I, 41, Logroño 26002, Spain, <sup>3</sup> Sciento, Via Val d'Aposa 3, Bologna, Italy, <sup>4</sup> Open University of the Netherlands, <sup>5</sup> University of Nottingham, UK  
{noaa.barak, daniel.burgos}@atosresearch.eu

**Abstract.** Past attempts at creating stakeholder networks for specific fields of research or industrial sectors have shown to be a resource-consuming and time-consuming process, which requires continuous monitoring and political efforts, as well as the trial-and-error deployment of technological tools. Still, these networks are thought to be an efficient and essential communication instrument for addressing challenges and building capacities. The EU FP7 STELLAR Network of Excellence has the mission of establishing a network for Technology Enhanced Learning (TEL) stakeholders, and has decided to do so via an online social community called TELEurope. In this paper we provide an overview of some relevant experience in establishing collaborative networks in the fields of business sciences, learning networks and communities of practice and reflect on our experience thus far with TELEurope.

**Keywords:** stakeholder network, social networks, STELLAR, TELEurope, Technology-Enhanced Learning, community of practice, virtual community

## 1 Introduction and background

Stakeholder networks are becoming a popular tool to facilitate dialogue in a thematic field or sector, with the aim of reaching a network-wide approach for resolving problems and addressing challenges. Svendsen and Laberge [1] define a stakeholder network as “a web of groups, organisations and/or individuals who come together to address a complex and shared cross-boundary problem, issue or opportunity”. Past experiences in establishing such networks, have fostered the development of theories and models to cope with three main challenges of establishing an effective stakeholder network: 1) identification of stakeholders, 2) network management, and 3) engagement. Nowadays, when reflecting on the establishment of a new stakeholder network, these three challenges have to be addressed, in view of the possibilities of using social software and web-based collaborative tools. In this regard, we draw out perspectives from three fields: business sciences [1], Learning Networks [2] [3] and communities of practice [4], [5].

With regards to the business sciences view, the literature lists cases in which corporations were forced to establish a stakeholder network by challenges which posed a barrier to their activity and had to be resolved through interactions with stakeholders [1]. A recent paradigm shift is described in which the approach towards stakeholder networks is becoming less *organisation-centric* and more *network-focused*. The *organisation-centric* approach, trying to manage and control stakeholders and inviting them to the network based on their potential influence, tends to result in short-term relationships, mostly focused on yielding benefits for the organiser, rather than to the overall network. In contrast, according to the *network-focused* approach, the initiating organisation is a symbiotic part of the environment and its sustainability depends on the well-being of its stakeholders. This approach yields a *multi-stakeholder network*, rather than a bilateral connection.

With regards to the learning network view [2][3], we highlight the key requirements to facilitate exchange and participation in a learning network: 1) facilities for members to create, search, get/access and study, 2) governance by community policies, 3) instruments to manage, change and apply policies, 4) high level of dialogue, interaction and collaboration, 5) an explicit exchange reward system which is consistent with self-organization principles, and 6) a right balance between usability for the participants and flexibility/complexity.

Finally, a Community of Practice (CoP) is describes as “a group of professionals who share a common interest for a domain or a specific topic. They meet on a regular basis, face-to-face or online... They share their daily practice... and generate new insights and understanding of their profession.” [4]. Although a somewhat homogenous group this can be considered as a special case of a stakeholder network.

## **2 The STELLAR Stakeholder Network**

STELLAR is a Network of Excellence, funded by the European Commission, with the aim of unifying the diverse community of TEL in Europe. STELLAR is motivated by “the need for European research on TEL to build upon, synergize and extend the valuable work we have started by significantly building capacity in TEL research within Europe” [6]. As preparation, the authors of this paper have researched and mapped the TEL research community in Europe [7], and then drafted the terms of reference and the theoretical structure of the stakeholder network [8]. December 2009 saw the official launch of the STELLAR stakeholder network via a stakeholder panel held at the Online Educa Berlin conference (OEB), and the initiation of an ELGG-based [9] social platform branded by the name TELeurope.eu [10].

### **2.1 Mapping TEL stakeholders in Europe**

Our findings in relation to the map of TEL stakeholder are out of the scope of this paper and are reported in [11]; however our methodology is relevant to the understanding of the convening process of the network. Rather than simply

categorising stakeholders, a framework was developed whereby each stakeholder is classified according to their membership within sub-groups of a set of groups. The characteristics of each sub-group were described in terms of their overall position and interest in TEL, as well as in terms of their needs and possible interest in the instruments making up the STELLAR project. The measurement of stakeholders' alliance potential was based upon work done by the World Bank [12]. The purpose of defining this indicator is to allow the network coordinators to identify how members might be most appropriately rallied in the pursuit of a particular initiative.

## **2.2 Description of the STELLAR stakeholder network**

Three key principles have shaped the development of the STELLAR network: 1) openness: to welcome all those with interests in technology enhanced learning whether as researchers, practitioners, designers or users. 2) Collaboration; the Network has to bring together different perspectives and interests in order to increase cohesion and reduce fragmentation in TEL at the European level. 3) Sustainability. To this end, the most important decision taken as a result of this strategy has been the branding of the online community as "TELeurope.eu": designed to be a visible and recognisable brand within the European TEL Community. Choosing to separate TELeurope from the main STELLAR brand is expected to strengthen the sustainability of the network and ease the connection with other initiatives in the field.

TELeurope provides a number of benefits to potential members: 1) networking opportunities and being part of a larger community of shared interests, 2) increasing the visibility of personal profiles, 3) receiving news of projects, events, promotional opportunities, leading developments in TEL and other communications, 4) access to resources such as reports and 'grey' literature, 5) access to expert discussions and opinions, and 6) opportunities for funding, collaboration and employment. At the more collective level, the community could potentially benefit from: 7) collective lobbying power, 8) access to test beds on a regional scale, 9) expert reviews, 10) EU-wide TEL research presence, and 11) a "neutral zone" to discuss field related matters; a way of reducing barriers between research, innovation, policy and practice; and contributing to the development of the research agenda related to TEL.

The first set of benefits might be viewed as offering immediate or direct forms of interaction between stakeholders; however, the second set of benefits will take time to evolve, emerging out of the growing sense of community identity.

Based on these principles and previous analysis, we have pre-structured TELeurope with the aim of containing three main groups: 1) Stakeholder advisory board, 2) Network of Networks, and 3) Research and innovation. In addition, members will be able to freely create their own groups. In practice, over a period of four months, over 300 individuals have created an account on the TELeurope platform and have mainly used the service to enable the creation of their own groups. In total, 17 groups were created. In most cases, these were linked to existing TEL research projects and conferences looking for the involvement of others in the community, or were linked to specific STELLAR activities. Some activity outside of

TELeurope has been taking place with regards to the Stakeholder Advisory Board; however the other two main groups are lagging behind.

### 3 Analysis

STELLAR aims at establishing a *multi-stakeholder network*, which will bridge the gaps between communities and disciplines in TEL. However, some of its actions while convening the network can be perceived as promoting a unidirectional flow of information, accommodating the needs of STELLAR, but less the needs and interests of TEL stakeholders. It has identified stakeholders and events based on their popularity and influence, as perceived by STELLAR, and these were targeted as the first invitees to join the network. Similarly, in a preliminary list of use cases drafted by STELLAR members [8], 8 out of 13 use cases (over 60%) involved STELLAR members; this is a bias towards the initiator of the network. Yet, examples to the contrary exist: an extended list of use-cases is currently underway, intentionally highlighting networking between non-STELLAR members. Additionally, at the Online Educa Berlin (Dec 2009) a group of stakeholders were invited to attend a stakeholder panel in which they were asked to reflect on the purpose of the network.

We have consciously chosen to make the TELeurope community open for all persons of interest, upon the completion of a simple registration process and confirmation of an email address. However, it seems that, despite this *network centric* approach, stakeholders are lacking a strong incentive to become engaged in the community, and instructions from the convenor's side are not readily acted upon. Furthermore, the establishment of TELeurope did not take into consideration existing processes and best practices of establishing a stakeholder network, from the three perspectives presented previously and so it has not been successful in clearly communicating some of the key features of the network to its members.

With regards to usability, complexity and flexibility, TELeurope has to be balanced. Members make extensive use of the group creating function, thus demonstrating the network's flexibility. However, several aspects need to be improved, such as the lack of support for specific use-cases, an unintuitive interface and the provision of basic features common to many social networks.

At the moment, an awareness service ("dashboard") presents recent activities on the platform, as well as site announcements, but it cannot yet be configured to comply with members' preferences. When compared to the main collaborative features in communities of practice [4], it is evident that TELeurope is still lacking some advanced services to foster social interactions and awareness.

Finally, table 1 presents TELeurope's compliance with the basic requirement for exchange and participation [2]. Although TELeurope offers some of the functionalities mentioned, it is still failing when it comes to the application of policies and a reward mechanism. The lack of statistics on user engagement makes it difficult to maintain a balance between usability and flexibility of the system.

Table 1: TELeurope's compliance with the requirements for participation and exchange of information

Requirement	Realisation in TELeurope
Facilities to create, search, get/access and study	Supported through: create and join groups, participate in discussion, upload and download files, post messages on “the wire”, search members and content
Governance by community policies	Monitoring is performed by the moderators and focuses on removing spammers; it does not reflect policy
Instruments to manage, change and apply the different policies	Unavailable as a community activity
High level of dialogue, interaction and collaboration	There is not much interaction yet; needs to be stimulated and supported by the STELLAR project
An explicit exchange reward system	A general plan has been outlined and different metrics have been discussed [8], not yet implemented.
Balance between usability and flexibility/complexity	Customisable homepage and profiles are available; however data on usability is scarce and it is hard to assess the balance. Basic information and tutorials are not available.

## 4 Discussion and Conclusions

In this paper, we relate to views and experiences coming from Business, Learning Networks and Communities of Practice, however none is entirely adequate to the special case of a web-based stakeholder network dedicated to research.

After one year of activity, and four months since the launch of TELeurope, the STELLAR stakeholder network is experiencing a tension between an organisation-centric and a system-whole approach. On the one hand, the network is defined as a *multi-stakeholder* network, aimed at providing benefits to the TEL field as a whole, rather than to STELLAR or a specific group of stakeholders inside the network. On the other hand, activities so far have had an *organisation-centric* feel – trying to demonstrate STELLAR’s achievements more than the actual progress on facilitating inter-disciplinary dialogue. In order to achieve its goals, TELeurope should take a step towards a more *network-focused* view, possibly by applying participatory design.

The authors of this paper intend to continue using TELeurope as a case-study through which they hope to draft a theoretical basis to support similar networks, highlighting both the organisational and the technical aspects.

**Acknowledgements.** This paper was supported by the STELLAR Network of Excellence (FP7, ref. 231913), and the Spanish national project TELMA ([www.ines.org.es/telma](http://www.ines.org.es/telma), TSI-020110-2009-85). The authors wish to thank their collaborators in STELLAR, with a special mention to Dr. Marie Joubert of the University of Bristol, UK.

## 5 References

- [1] Svendsen, A.C., Laberge, M. Convening Stakeholder Networks: A New Way of Thinking, Being and Engaging. JCC 19, 91--104 (2005)
- [2] Hummel, H., Tattersall, C., Burgos, D., Brouns, F., Kurvers, H., Koper, R. Critical Facilities for Active Participation in Learning Networks. IJWBC, vol. 2, issue 1 (2005)
- [3] Koper, R., Pannekeet, K., Hendriks, M., Hummel, H. Building communities for the exchange of learning objects: theoretical foundations and requirements. ALT-J, Vol. 12, No. 1, 21--35 (2004)
- [4] El Helou, S., Raffier Möller, MD., Daele, A., Gillet, D. Social Software for Supporting Interaction in a Community of Practice Dedicated to e-Learning. Workshop on Social Networking in Organizations, CSCW (2008)
- [5] Burgos, D. The structure and behavior of virtual communities engaged in informal learning about e-learning standards (Estudio de la estructura y del comportamiento de las comunidades virtuales de aprendizaje no formal sobre estandarización del e-learning). Doctoral dissertation, European University of Madrid, Villaviciosa de Odón, Madrid, Spain (2006)
- [6] STELLAR: The network of excellence for Technology Enhanced Learning (project reference: 231913), <http://www.stellarnet.eu/>
- [7] Camilleri, A., Barak, N., de Vries, F., Joubert, M., Specht, M., Windrum, C., Persico, D., Kraker, P. Report of the process of establishing a Stakeholder Panel. STELLAR Deliverable 5.1 (2009)
- [8] Barak, N., Burgos, D., Specht M., De Vries F., Windrum C., Camilleri A., Lindstaedt S., Persico D. Report on Guidelines and Monitoring tools-metrics. STELLAR Deliverable 5.2 (2009)
- [9] ELGG, <http://elgg.org/>
- [10] TELeurope, <http://teleurope.eu/>
- [11] Camilleri, A., Barak, N., Burgos, D., Ullmann, T. D. Engaging the community in multidisciplinary TEL research: a case-study on networking in Europe. EDEN (2010).
- [12] World Bank. Policy Toolkit for Strengthening Healthcare Reform. 2000, <http://info.worldbank.org/etools/docs/library/48545/RD1.PDF>.pdf last accessed July 2009.