D1.2 Trends in Connecting Learners

Citation for published version (APA):


Document status and date:
Published: 25/10/2010

Document Version:
Publisher's PDF, also known as Version of record

Document license:
CC BY

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
https://www.ou.nl/taverne-agreement

Take down policy
If you believe that this document breaches copyright please contact us at:
pure-support@ou.nl
providing details and we will investigate your claim.

Downloaded from https://research.ou.nl/ on date: 23 Sep. 2020
Trends in Connecting Learners (Annex)
First Research & Technology Scouting Report

Edited by
Eelco Herder
Ivana Marenzi
## Amendment History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Contributor(s)</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>19/01/10</td>
<td>Eelco Herder</td>
<td>Initial Structure</td>
</tr>
<tr>
<td>0.2</td>
<td>12/02/10</td>
<td>Eelco Herder, Ivana Marenzi</td>
<td>Small-Scale Studies inserted</td>
</tr>
<tr>
<td>0.3</td>
<td>19/02/10</td>
<td>Eelco Herder, Ivana Marenzi</td>
<td>Co-Author Analysis inserted</td>
</tr>
<tr>
<td>0.4</td>
<td>23/02/10</td>
<td>Eelco Herder</td>
<td>Survey of Trends</td>
</tr>
<tr>
<td>0.5</td>
<td>26/02/10</td>
<td>Eelco Herder</td>
<td>Major edit of existing material</td>
</tr>
<tr>
<td>0.6</td>
<td>03/03/10</td>
<td>Eelco Herder</td>
<td>Introduction, conclusions</td>
</tr>
<tr>
<td>0.8</td>
<td>15/03/10</td>
<td>Ros Sutherland, Denis Gillet</td>
<td>Reviews and corrections</td>
</tr>
<tr>
<td>1.0</td>
<td>17/03/10</td>
<td>Eelco Herder, Ivana Marenzi</td>
<td>Revision according to reviewer remarks and results of proofreading</td>
</tr>
</tbody>
</table>

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.
Table of Contents

Executive Summary ........................................................................................................... 5
Annex 1: Who are you working with? Visualizing TEL Research Communities ... 6
Annex 2: Creation and Use of Annotations in Different Usage Contexts .......... 17
Annex 3: Methodologies for the Design of Structured CSCL..................................... 32
Annex 4: Identifying knowledge development - a context-aware approach for analyzing knowledge processes................................................................. 40
Annex 5: Phone, Email and Video Interactions with Characters in an Epidemiology Game: towards Authenticity................................................................. 59
Annex 6: Technology for classroom orchestration....................................................... 71
Annex 7: Location-based and contextual mobile learning .......................................... 92
Annex 8: Supporting students to self-direct intentional learning projects with social media................................................................. 113
Executive Summary

This document is an annex to the D1.2 RTST Trend Report and contains the white papers of the small-scale studies, as presented and discussed in Chapter 3 of the deliverable.

Small-scale studies are short projects, carried out by the Stellar partners - in some cases in cooperation with external partners. The aim of these studies is to provide insight in the current state-of-the-art in a specific field.

The studies have been selected after internal discussions among the partners during the Stellar Roadmapping Meeting in Bristol, several FlashMeetings and one-to-one conversations. Further, the setup of these studies and how they will succeed in providing a coherent view of the state-of-the-art has been analyzed by an internal reviewer.

In an online meeting, the study results were presented to one another, followed by a discussion on to what extent the conclusions and challenges derived from the study can be related to each other, to the trends as reported in the earlier chapter and to the Stellar Grand Challenges. The outcomes of this discussion and a reflection on the studies can be found in Chapter 3.
1. 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\textsuperscript{1} and CiteseerX\textsuperscript{2} - 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. 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.

2. 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 EDMedia 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 EDMEDIA 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 sub-areas such as design models, hypertext writing, open hypermedia and information visualization), while our factors represent broader research

\textsuperscript{1} http://www.informatik.uni-trier.de/~ley/db/
\textsuperscript{2} http://citeseerx.ist.psu.edu/
communities, centred around one or a few community-centred conferences such as Adaptive Hypermedia or AIED.

3. 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.

3.1 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, Klimma, Nejdl, Balacheff, Sharples, Davis, Zimmermann, Wolpers, Sutherland). From these conferences and journals (as identified in DBLP3), 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 TEL4. 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,

3 The detailed procedure is described in the Stellar deliverable D7.1: http://www.stellarnet.eu/d/7/1/Investigating_two_silos
4 Other topics are computer science (27 venues), artificial intelligence (26), human-computer interaction (22) and databases (5).
User Modelling 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\(^5\) and ECTEL\(^6\), 15 new authors from the IEEE TLT Board and Steering Committee\(^7\), and 5 more authors from the Telearn archive\(^8\). We also included the top-15 cited papers or books from EDMedia 2005 – 2008 (Ochoa et al. 2009). Second, after merging these sets, we selected the authors with at least 20 publications in CiteceerX 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 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\(^9\), 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.

### 3.2 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).


\(^7\) [http://www.computer.org/portal/web/tlt/edboard](http://www.computer.org/portal/web/tlt/edboard)

\(^8\) [http://telearn.noe-kaleidoscope.org/](http://telearn.noe-kaleidoscope.org/)

\(^9\) [http://dblpvis.uni-trier.de/help/overview.html](http://dblpvis.uni-trier.de/help/overview.html)
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.

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.

3.4 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 <i,j>
Update matrix cell <i,j> and <j,i>

Listing 1. Pseudo code for the matrix computation.

4. 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.

4.1 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 your 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 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 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

10 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.
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.

5. 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.
6. 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 modelling 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 Rüsling (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 modelling (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 EDMEDIA, as identified by Wild et al (in press).

7. 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.\footnote{The CSCL conference, for example, is not indexed in DBLP and therefore missing in our analysis.} 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.

8. References


Acknowledgments

Work on this paper has been partially sponsored by the STELLAR Network of Excellence.
Annex 2: Creation and Use of Annotations in Different Usage Contexts

Ricardo Kawase, George Papadakis, Eelco Herder and Wolfgang Nejdl,
L3S Research Center,
Appelstraße 9, 30167 Hannover, Germany
{kawase, papadakis, herder, nejdl}@L3S.de

Abstract. It is virtually impossible to think of a world without annotations. People write comments in paper margins and highlight text passages; they write reminders on post-its and put short messages on colleagues' desks. Online annotations serve different goals than their pen-and-paper counterparts; they are mainly limited to social bookmarking and social networking. In this paper we present a series of independent yet related user studies on various aspects of paper-based and online annotations. We show how shared annotations are created and used and how annotations may support re-finding. We provide design recommendations on how to better support these differences in user contexts and goals.

Keywords: Annotations, social bookmarking, re-finding.

1 Introduction

Learning has become an integral part of many people's everyday working life. Due to a more knowledge-based society and rapid changes in technology, one often has to search for and read information in order to keep up-to-date. Each individual presents a set of cognitive strategies that involve the learning process: each person learns in her own way, style and pace. At the same time, the character of learning at the workplace has shifted from a solitary, paper-based activity to a Web-based activity, making use of various resources, including discussion forums and social networking sites [28]. As a result, one ends up with a large collection of scattered digital resources; due to limitations of the Web, annotations – if any – are typically made separately (in a word processor or on a paper sheet). By contrast, annotating paper documents is a natural activity that involves direct interaction with the document and that is known to support understanding and memorization [27].

Annotations have been the focus of a large body of research in the fields of personal information management, hypertext and Web 2.0. This is mainly due to the significant role annotations play in our daily lives. People regularly write comments in paper margins and highlight text passages, in order to support comprehension and interpretation [19]. At a later stage, these annotations also serve as reminders and help people to re-find relevant passages. Moreover, comments and references are known to stimulate associative thinking. Annotations and notes are also heavily used for communication: people leave short messages, requests and reminders at colleagues' desks and monitors, or put them on the refrigerator, where they are likely to be found.

Given the abundance of annotations and notes in our daily lives, a logical next step was to incorporate them in computer systems. Inspired by Vannevar Bush's vision of associative trails [5], annotations have been a central element of classic hypermedia systems, such as Xanadu. Whereas the World Wide Web surprisingly did not support annotations in the early days, the current interaction-centered Web 2.0 does. Furthermore, in the past decade several interesting annotation systems were developed and released. Empirical data shows, though, that annotation systems are hardly used in general [9] – a fact that is in sharp contrast with the prevalence of interactive applications and social networking.

A probable cause for this phenomenon is a number of drawbacks that digital annotations are known to have in comparison to paper-and-pen annotation. In particular, the lack of tangibility and direct interaction imposes a higher cognitive load and distraction from the primary task. To balance things, digital annotations are accompanied by some unique
features, among others the support for sharing, easy indexing, and the ability to order, rate and search these annotations.

Our research is motivated by the observation that current annotations tools do not sufficiently cater for the differences in user practices and goals that are inherent when moving from the paper world to the digital world. In this paper, we present the results of a series of user studies, each one of them focusing on a different aspect of current digital annotation practices: comprehension and interpretation, sharing and communication, and re-finding. We compare these practices with similar ones, applied not only on paper but also online. The setup of the studies and the subsequent analysis are guided by the following research questions:

• what are the differences between paper-based annotation practices and their digital counterparts?
• are shared annotations different from private annotations? Moreover, to which extent do users benefit from shared annotations that are written by friends or colleagues?
• how do digital annotations support re-finding, in comparison with search engines and bookmarks?

The paper is structured as follows. In the next section, we discuss theoretical insights in the forms and purposes of paper-based and digital annotations. We continue with an overview of selected relevant digital annotation systems, followed by the experimental setup and results. The paper ends with a discussion on design implications of the different usage contexts that annotations should support.

2 Paper-Based and Digital Annotations

As we encounter annotations on a regular basis, we all have a certain idea on what an annotation is. In order to frame our research, in this section we give a more formal definition of annotations, summarize the different forms in which they occur and the purposes they serve. Further, we discuss issues that are reported to occur when moving annotations from paper to the computer screen.

In line with MacMullen [17] and Marshall [19], we define annotations as any additional content that is directly attached to a resource and that adds some implicit or explicit information in many different forms. Annotations may serve different purposes, including signalling for future attention, memory aiding, interpretation and reflection. Correspondingly, annotations may occur in many different forms. As an example, by highlighting, encircling or underlining text, we emphasize the importance of a certain part of the document; a strikethrough indicates that something is wrong, misplaced or not relevant; arrows signal relations between two or more elements.

Interacting with a document is known to stimulate critical thinking and reflection, a process that can be called ‘active reading’ [1], as opposed to passive consumption of text. In particular, text in the margin of a document may support better understanding of the topic during later reading.

Annotations are informal, ad-hoc and take many forms. Based on extensive field research, Marshall [20] came up with a categorization of the main forms of annotations and the functions that they serve:

• Underlining or highlighting titles and section headings serve as signals for future attention. The marked section may contain some particularly relevant information, or the reader might want to explore it in more detail at a later stage.

• In a similar vein, highlighting and marking words or phrases and within-text markings aim to serve remembering and placemarking important passages.

• Notations in margins or near figures may appear in many forms, including pictures, diagrams and calculations. These kinds of annotations typically serve comprehension.

• Marginal notes, jottings and between-line texts usually contain summarizations and interpretations of the document’s contents. These kinds of annotations vary from keywords to elaborate comments.

In addition, some people highlight dense text passages to monitor their progress through the text – with no particular future communicative value for themselves or for others. People also
often draw cartoons or pictures in the margin, or write personal notes that are unrelated to the document. As these ‘decorations’ add no information to the document, we do not consider them as annotations \textit{per se}.

From the above discussion on paper-based annotations it can be observed that the different kinds of annotations mainly serve two purposes: to facilitate \textit{thinking} and comprehension, and to facilitate the \textit{re-finding} or reuse of relevant material. A further observation is that, due to the fact that annotations are attached to the original paper, one is easily prone to \textit{forget} potentially important annotations, or not to be able to find them, because that they are somewhere in a pile of papers. It was Vannevar Bush [5] who recognized the great advantage of digital annotations: they allow for the \textit{associative linking} of two fragments that may otherwise remain unconnected due to the physical constraints of paper. Further, at a later time users can retrace their \textit{trails} of annotations, for reflection on their activities and for re-finding and relating material.

Early hypertext systems followed Bush’ vision by implementing rich, typed and bidirectional links with various means for annotation. As pointed out by Obendorf et al [24], the read-only character of the Web and browser history mechanisms (including bookmarks, history lists and the back button) do not sufficiently cater for associative thinking or re-finding. In the past decade, a vast amount of (prototypic) systems have been developed to bring annotations to the Web – we discuss a selection of these systems in the next section –, but none of them has reached a big audience so far. On the other hand, social bookmarking sites such as Delicious\footnote{http://delicious.com/} show another advantage of digital annotations: \textit{sharing} information and comments online with friends or co-workers is far easier than distributing copies of paper. Nevertheless, users are still reported to resort to copying strategies, such as sending comments via e-mail [9, 27].

From the above it becomes clear that digital annotations have several potential advantages, in particular concerning the sharing of information and comments. Another advantage is that it is easier to \textit{index, order, rate and search} your repository of digital annotations, if appropriate indexing and querying mechanisms are provided. However, the transition from the paper-based world to the digital one comes with a number of disadvantages too, as discussed in the literature [7, 8, 10, 12, 20, 23 and 25].

- Lack of \textit{tangibility}: readers can hold paper as they like, to adjust perspective and distance. Electronic displays do not provide this flexibility. In combination with the lower resolution provided by electronic displays, this has an impact on \textit{legibility}.

- Lack of \textit{direct manipulation}: whereas readers can simply take a pencil to freely write and draw on a piece of paper, on electronic devices this requires indirect input via keyboard or mouse, often in combination with the selection and activation of specific tools. This explicit task switching adds to the cognitive burden of the user and makes annotating a distracting rather than a supporting activity.

- Lack of \textit{orientation}: paper documents give readers a better sense of location within the text, for example by the thickness on the sides of a book. These cues support text skimming, cross-reading and re-finding some text. Further, as digital annotations may not be directly placed \textit{in the context} of the original text fragment, it may cost more effort to relate the two of them at first sight.

- Lack of \textit{cooperative interaction}. As discussed earlier, digital annotations have the advantage that it takes only one click to share them with others. Still, circulating a piece of paper offers more intuitive ways to build upon each others’ comments and input than groupware facilities, such as revisions and versioning.

To summarize, digital annotations have several inherent disadvantages that are related to the nature of electronic devices. Digital annotations have several advantages too, in terms of sharing, searching and re-finding – which have been confirmed to be very desirable features in several studies [e.g. 11]. This suggests that digital annotations are not just the electronic counterpart of paper-based annotations, but that there are some intrinsic differences in the way they are conceived and used. This implies that designers of digital annotation systems need to find the right mix of mimicking practices from ‘old-fashioned’ paper-based annotation and promoting the slightly different practices of digital, online practices [14].
2.1 Annotations in Elearning

The benefits and opportunities of electronic and automatic annotations, elaborating on their paper-based counterparts, have long ago been envisioned by Vanevar Bush in the Memex [5]. Bush envisaged that by relating all documents that users have read and attaching their annotations to these documents, individuals could organize and re-find information resources in an associative manner, together with any earlier annotations. Whereas the original rich forms of annotations in Hypertext systems – with different categories, directions and even multi-links – allowed for these associative trails, in the Web as it is today this functionality is not totally fulfilled, as readers have limited possibilities for sharing comments or questions by writing back to the pages. As a result, users spend a lot of effort trying to comprehend the different formats of how people comment on-line resources using coping strategies such as sending comments via e-mail [29].

Recent Web 2.0 technologies provide an open resource environment where individuals can freely collaborate. Nevertheless, these technologies typically only cover just a slight portion of the Web or one specific kind of annotation. These technologies are typically implemented as Web servers or browser enhancements.

The basic idea of a Web annotation system is that the user has the ability to change, add or attach any type of content to any online resource, similar as she would do it with a paper document. An application (usually a browser plug-in) enables the user to modify the Web pages, highlight parts of it and add tags or comments, while the back-end of the system just need to check these annotations and associate them with the specific user and the specific URL.

As discussed in the previous section, by actively being involved with the text, users can better memorize and understand it. By contrast, annotating on a computer-screen is an activity that competes with the reading itself, due to the lack of direct manipulation. However, users will do so when the benefits are higher than the costs in terms of effort. These benefits may include the saving of time needed for re-finding, summarizing, organizing, sharing and contributing online annotations. A rather economical view on the balance between the drawbacks and benefits has been given by [30]'s information foraging theory, in which he described the above activities as information enrichment.

Today, both companies and academia institutions train learners to complete tasks and solve problems through project-centred learning. Since it may not be feasible for all participants involved in the projects to meet on a regular basis, they must be assisted by information and communication technology. To support this collaboration there are specific methods for Computer Supported Collaborative Learning (CSCL) provided by learning environments and other platforms can be adapted to fit this need. For the best results of the learning process, the methods should help each learner to act individually to reach her own goals and to cooperate by sharing and discussing ideas to accomplish an assignment.

As discussed in the previous section, in the same way annotations contribute for memory aiding, text interpretation and information re-finding, Web annotations provide the same functionality in the online environment. Web annotations are accessible anytime and anywhere, with diverse sharing possibilities, clearly enhancing group collaboration for cooperative tasks and learning processes. However it is important to remark that the full richness of paper annotations will only be achieved if the digital annotations hold the same beneficial feature of being 'in-context'. ‘In-context’ annotations are visible within the original resource, enhancing it with the observations and remarks of the annotator, which are likely to help in individual tasks in similar ways as is the case with paper documents [10].

Despite the limitations in terms of usability and tangibility, advantages of Web annotation tools go far beyond the advantages of regular paper annotations. In addition to the sharing capabilities within online communities, digital annotations can be indexed, ordered, rated and searched. These benefits are confirmed by several studies on annotations tools [e.g. 18], in which participants have remarked that search the annotations is a very desirable feature.

Even though there are currently systems that support annotations, studies have shown that users often resort to different strategies for simulating annotation tools, making use of e-mails and messages to self and separated text documents. The main reason for this phenomenon lies mainly in the necessary effort required for creating and organizing annotations: “If it takes three clicks to get it down, it’s easier to e-mail” [26]. As users will inevitably resort to other strategies if annotation tools require too much effort, it is necessary
to have a lightweight capture tool, with flexible organizational capacity, visibility and practical reminding. In particular if one takes into account that many annotations are primarily meant as temporary storage, or a means for cognitive support or as reminders, it becomes clear that these factors need to be better taken into account in annotation tools for personal information management and learning systems.

2.2 Digital Annotation Systems

In the past decade, a plethora of tools and services has emerged for storing, sharing and commenting upon Web content. Popular browser extensions include ScrapBook – which allows users to save and edit an offline version of a page – and StumbleUpon – which provides users with recommendations for sites, based upon sites visited and reviewed by other users. Social bookmarking is a popular activity, as can be observed from the popularity of Delicious. Further, many popular Websites allow users to leave comments. These tools and services provide valuable means for keeping track of, interacting with, and communicating about interesting content. However, annotations – as discussed in the previous section – are not the primary goal of these tools.

Annotea [15] is a W3C-sponsored standard for annotations. The standard is implemented in the Amaya Web editor and the Annozilla Firefox plugin. Annozilla allows users to create rich-text annotations that are stored on one of the annotation servers. The technical character of the installation instructions suggests that the system is not (yet) meant for the general audience, though. A more full-fledged annotation tool is Diigo. Using the Diigo toolbar, users can highlight text or attach 'inline sticky notes' to Web pages. From the Diigo website, users can make use of the collaborative features of this "social bookmarking and annotation service". Despite the wealth of features, Diigo cannot boast a big user population. According to the online user comments, this is due to both usability issues and the fact that all annotations are public.

In the research community, several annotation systems have been created, often with the focus on a specific feature or innovation. Below, we discuss a small selection. The browser plugin MADCOW [4] allows users to create a wide range of annotations and to attach it to text or even parts of images. The flexibility comes at a price, though: from the screenshots it can be observed that creating an annotation is a rather cumbersome process that involves many choices and selections. The Anchored Conversations system [5] follows a radically different approach: it provides a synchronous chat window that can be anchored to a specific point within a document. The chat provides the basis for the annotation. The authors elaborated on this idea in StickyChats, MapChat and KeyHoles. In the e-learning community we see systems such as OATS [3], which provides social annotation and tagging to learners. As can be observed from the above, an important goal of annotation is collaboration. Some projects, such as Quilt [16] and Comments [22], provide support for text and voice annotations in the context of collaborative writing. The presence of annotation and notification features in word processors is a further proof of the importance of communication and sharing comments in online collaborative work. However, we agree with [3] that these kinds of annotations, which support the creation of a document rather than adding information to existing documents, serve a different goal and should be treated as a different category.

So digital annotations – in one form or another – are manifold on the Web and often created using applications and services that are not specifically designed for annotating. We observe that social bookmarking is regarded as a natural extension for digital annotations; at the same time, users feel that some annotations should be private too. Finally, we notice that usability – in particular a balance between flexibility and minimizing user interaction – is of particular importance.

13 http://www.diigo.com/
2.3 A Straightforward Approach to Online Annotations

Our work is focused on experimental results using our own developed annotation tool. In this context, in order to better understand the real use of annotations and Web annotations, we implemented a straightforward online annotation system, SpreadCrumbs [13]. SpreadCrumbs provides a minimalist interface for adding post-it notes, crumbs, to any point within a Web page. After clicking the right mouse button, the user only has to provide a short title and the comment (Fig. 1).

By default, crumbs are private, but the user can decide to share the crumb with one or more contacts, who can add their own comments to the crumb, thus creating a discussion thread. If a user visits a page that contains an annotation – written by the user or shared by a contact – it is displayed on the location where it was created. In addition, users can scan their list of annotation in the browser side bar.

3 Experimental Setup

To address the research questions listed at the end of the introduction, we performed four user studies. Study I is a field study on how people annotate on paper. The results are compared with Study II, which concentrates on how people create online annotations. Study III was conducted in the same session as Study II and explores how people use shared annotations. Finally, in Study IV, we investigate the benefits of annotations for re-finding information after a longer period (making use of the results of Study II).

3.1 I - How People Annotate on Paper

The first of our studies aimed at exploring how people annotate on paper-based material, in order to understand its forms and purposes, its frequency, its benefits as well as its drawbacks. Our pool of participants consisted of 22 academics (15 male and 7 female, average age 29). We visited them at their regular working places and asked them to show us three research papers or articles that they recently printed and read. We collected 66 articles, covering a total of 591 pages of text. From the copies of these articles, we counted the number of annotations and we classified them according to Marshall's categorization [18]. During the process, we asked our participants for elaborations and explanations in a conversational style.

3.2 II - How People Annotate Online

As a comparison with the first study, in our second study we investigated how people create annotations online. Our participant pool consisted of 24 males and 10 females, with an average age of 28 (no overlap with the participants of Study I). As the results will inevitably be biased by the features of the tools used, our participants were randomly and equally split...
into two groups: the first group created annotations using the Delicious social bookmarking service, the second group made use of the SpreadCrumbs annotation tool.

After a short introduction to the basic features of the tool (either SpreadCrumbs or Delicious), we asked the participant to find answers for ten random questions (which we presented as a ‘training task’ for the second half of the experiment, Study III). All questions were specific information-finding tasks that could be solved by a brief internet search with any popular search engine. We ensured that the questions were sufficiently obscure, to minimize the chance of participants knowing the answers themselves: most of the answers were numerical in nature – an example question is: “How many home runs did Ken Griffey Jr. hit in 2005 playing for Cincinnati?”. The participant was asked to annotate the Web page that contained the answer for later reference. In addition to the data gathered during this session, we also analyzed the annotations that were written by regular SpreadCrumbs users.

3.3 III - How People Use Shared Annotations

After answering the 10 questions in the previous study, the second part of the individual session – with the same participants and the same groups – started with a scenario on collaborative decision making. The participants were asked to plan a trip to London, by reviewing the options, as collected by their ‘partner’ (the experimenter). Via either SpreadCrumbs or Del.icio.us, the participant received a number of annotations/bookmarks on suitable hotels, restaurants, museums and musicals in London. The participants evaluated the given options – by visiting the bookmarked sites and/or by reading the annotations – and finally decided for one option in each category. After having finished both tasks, the participants were asked to fill out a short usability questionnaire and to evaluate the tools – not covered in this paper.

3.4 IV - How People Use Annotations for Re-finding

Five months after the initial round of the studies, the participants were invited to participate again. This time, their task was to re-find the answers that they had previously found during the first ‘training’ task (Study II). The long time interval ensured that the participants remembered neither the answers they had provided nor the resources they had used to find the answers. In total, 30 out of the initial 34 participants were involved in this phase of our study.

The participants were divided into three equivalent groups of 10 people, each one corresponding to a specific re-finding methodology and corresponding tool. As a base line, the first group had to search again for the same information using a search engine. This group was formed by randomly choosing 5 participants from the bookmark group together with another 5 from the annotation group of the first session. The second group used bookmarks to re-find the information. This group consisted of those participants that used Delicious in Studies II and III. They had the URLs of the visited resources at their disposal, along with the comments that some participants had added before. The third group consisted of the SpreadCrumbs users of Study II and III (minus the participants assigned to the search engine condition). These participants had the in-context annotations at their disposal that they had created in Study II – in several cases located near the relevant text fragment of the page.

The participants of all three groups were presented with one question at a time, chosen in random order. We ensured that all subjects accomplished all of their tasks under the same conditions and that their performance is compared on an equal basis.

After the appropriate Web resource was found, thus completing the searching stage, the participant had to locate the answer in the page and highlight it using the mouse – the browsing stage. The participants were allowed to perform this task the way they would in a non-controlled environment.

The necessary data for estimating and evaluating the average and overall browsing time per individual were collected with the help of screen capture and data-logging software that recorded all participants’ actions. Further, the participants were asked to answer two questionnaires, one regarding the information re-finding experience and another one investigating their opinion on the tool they used.
5 Results

In this section we discuss the results of the four studies described above. Our first focus is on the differences between paper-based and digital annotations in terms of format and usage. Second, we focus on what makes shared annotations different from private annotations and how they are used. Finally, we explore to what extent annotations help in re-finding information.

5.1 How People Annotate on Paper

To compare annotations in the online context with paper-based annotations, we visited the working place of 22 PhDs students and pos-Docs. We asked each one of them to take a look at the last 3 research papers or articles that they have printed and read. In total we have collected 66 articles, covering a total of 591 pages of text. We found 1778 annotations and an average of 3.08 annotations per page. The table below shows the average of each type of annotation per page.

<table>
<thead>
<tr>
<th>Annotation types</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highlighting/Mark sections headings</td>
<td>153</td>
<td>8.6%</td>
</tr>
<tr>
<td>Highlighting/Mark text</td>
<td>1297</td>
<td>73%</td>
</tr>
<tr>
<td>Problem solving</td>
<td>2</td>
<td>0.1%</td>
</tr>
<tr>
<td>General notes (Notes in the margins)</td>
<td>326</td>
<td>18.3%</td>
</tr>
</tbody>
</table>

The far majority of the annotations (73%) involved the highlighting and marking of text. Some participants had the tendency to only highlight main words within a sentence or paragraph. In these cases we counted the collection of highlighted words belonging to a continuous block of text as one piece of annotation.

9% of the documents discussed with the participants turned out to be part of collaborative work in which two or more people were involved. All except two participants reported that they shared their comments via email or some online communication tool; only two participants shared the same sheet of paper, which contained annotations from both parties. Participants who share annotations reported that they do annotate in a different (more careful) way when they annotate concerning another reader.
To examine in more detail the annotation strategies, we asked our participants to classify the goal of reading the paper. We distinguished between the following categories: reading for writing, reading for learning, reviewing and other. Reading for writing is the common activity of reading related articles to extract ideas and references specifically for propose of writing. Reading for learning includes the act of getting updated in some particular field, read about new publications or learning some new approaches to apply in some other activity, such as solving math problems or implementing algorithms. Reviewing consist exclusively of reading papers to give feedback to the author. Finally, any other type of reading was categorized as other. The table below shows some statistics.
Table 3. Results by reading goal.

<table>
<thead>
<tr>
<th></th>
<th>Writing</th>
<th>Learning</th>
<th>Review</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Articles</td>
<td>31</td>
<td>23</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Articles annotated</td>
<td>28</td>
<td>16</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Annotations/Page</td>
<td>2.36</td>
<td>4.7</td>
<td>1.11</td>
<td>6.3</td>
</tr>
</tbody>
</table>

**Annotation types**

<table>
<thead>
<tr>
<th>Annotation type</th>
<th>Highlighting/Mark sections</th>
<th>Highlighting/Mark text</th>
<th>Problem solving</th>
<th>General notes (Notes in the margins)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highlighting/Mark sections headings</td>
<td>10.5%</td>
<td>7.5%</td>
<td>9.4%</td>
<td>4.8%</td>
</tr>
<tr>
<td>Highlighting/Mark text</td>
<td>66.0%</td>
<td>82.9%</td>
<td>40.6%</td>
<td>72.2%</td>
</tr>
<tr>
<td>Problem solving</td>
<td>0.1%</td>
<td>0.9%</td>
<td></td>
<td>0.9%</td>
</tr>
<tr>
<td>General notes (Notes in the margins)</td>
<td>23.3%</td>
<td>9.6%</td>
<td>49.1%</td>
<td>23.0%</td>
</tr>
</tbody>
</table>

In addition to comments directly put on paper, three participants also used the technique of attaching annotations to the original document with post-its that were attached to the paper. From the 66 articles analyzed, 10 (15%) did not contain any annotation. One participant that did not have any annotation in any printed paper said that she keeps her annotations in a separate file in her computer for each digital article. Two other participants said that they first do a very quick reading on the computer to check the relevance of the text, and if it is relevant than they print it. In their own words: “First I read on the computer to see if I really need to print”.

From the annotations we identified many different ways of signalling important parts on the text. As an example, one participant created her own symbols for annotating: squares around the terms means new terminology, underline means definitions and circles means open question or issues over some topic. Those annotations symbols were used combined with highlighting (importance) and many times they even overlapped. We noticed that in many cases participants also used different marking colours for highlighting with the purpose of attributing different levels of importance.

One last interesting observation was the behaviour of one of the participants, who keeps two printed versions of every paper: one with annotations and one clean print. As she stated, the clean print is for a future reading, when she may want to get the idea without being influenced by her earlier comments. Despite the vast number of annotations by highlighting on the papers, none of the participants reported to use such mechanisms that allow persistent highlighting on digital documents or Web resources.

To summarize the above, we identified two main clusters of annotations: relevance adjustment annotations and contributive annotations. Relevance adjustment annotations involve highlighting and other means for marking text to indicate different levels of importance in the text; this type of annotations covered the far majority of annotations. Contributive annotations, explicit readable remarks that are added attached to the text, are used to a lesser extent, but still considered very important. Apart from this categorization, the major result is the striking variety in annotation techniques, which vary from the amount of annotations via the way annotations are done (for example highlighting versus squares and circles) to the material on which the annotations have been placed – within the margin, on a post-it note, in the digital version of the document. These differences in strategies are not merely the result of poor support, but reflect naturally developed behaviour that reflects the needs of the individuals in their working situation.

As a last part of our interviews we asked the subjects to describe how they arrange their papers that lay on their desktops. The relevant categories described were topic, quality, importance, date of reading and task. This simple observation may guide us to design better metaphors of the possible dimensions when trailing online resources.

### 5.2 Paper-Based Versus Digital Annotations

In our study (I) on paper-based annotation we found a total number of 1778 annotations in 591 pages of text – an average of 3.1 annotations per page. The vast majority of the annotations (73%) were accounted for by highlighted or marked text. General notes in the margins covered 18% of the annotations and highlighted section headings 9%. The popularity of text highlighting signifies that annotating is an activity that occurs while reading
light-weight methods that do not force the user to switch tasks and to stay concentrated on the main activity are preferred.

Further evidence for the light-weight character of annotations is given by the statistics on the regular use of the SpreadCrumbs annotation tool (not from the laboratory studies presented). The average number of words in private annotations is 4.56, which is shorter than the average length of short sentences [2]. Further, digital annotations often contained keywords rather than additional information, an indication that they are primarily or secondarily intended for re-finding purposes.

Paper-based annotations are typically located near the relevant text passage, as long as the margins provide sufficient space (Fig.2). In Study II we were interested whether this would be the case for digital annotations as well. The far majority of the participants who used SpreadCrumbs (16 out of 18) did so by carefully placing the notes near the text, table or paragraph where they found the answers.

Whereas most paper-based annotations are private by nature, 9% of the documents discussed with the participants turned out to be part of collaborative work in which two or more people were involved. All except two participants reported that they shared their comments via email or some online communication tool; only two participants shared the same sheet of paper. This would imply that digital annotations would be more supportive for sharing annotations. This is supported by our statistics from the regular use of SpreadCrumbs, of which 32% of the annotations were shared with at least one other person.

Another valuable observation is that all of the participants who share paper-based annotations said that they write them in a different, more careful way when they plan to share them with another reader. We will explore the intricacies of shared annotations in more detail in the next subsection.

Fig. 2 Example of paper based annotations.

5.3 Creating and Using Shared Annotations

In contrast to private annotations (with an average length of 4.56 words), the average length of shared annotations, as collected from the regular use of SpreadCrumbs, is 10.35 words, which fits the average length of short and medium sentences in plain text documents [2]. Shared annotations did not only contain more words, they also tended to be clearer and better structured (Table 1 illustrates this difference). This is further support for the observation that even though shared annotations (or messages, if you like) have many elements in common with private annotations, they serve a different purpose and therefore invoke different user behaviour.

Table 1. Examples of private annotations (left) and shared annotations (right), excerpts from the SpreadCrumbs log file.

<table>
<thead>
<tr>
<th>Private Annotations</th>
<th>Shared Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Conference Deadline: October 29&quot;</td>
<td>&quot;All artists are from Sweden, I think, and do Jazz music (quite soft) but nice...&quot;</td>
</tr>
<tr>
<td>&quot;Flat 64m 2 rooms windthorststr. 8&quot;</td>
<td>&quot;Let me know if there's anything else to be done.&quot;</td>
</tr>
</tbody>
</table>
If users put more effort in annotations that they share, it is likely that they expect the receiver to benefit from these annotations. It turns out that this heavily depends on how the annotation is conveyed to the receiver. In Study III, 50% of the users who received the suggestions from their ‘partner’ via Delicious did not read or even did not notice the additional comments on each bookmark, which were displayed just below the page title and the URL. One participant explicitly told us that she noticed them only in the middle of the task. Another participant said that she noticed the comments, but did not read all of them because she thought they were irrelevant.

By contrast, all the participants who received the suggestions via SpreadCrumbs did notice and read the comments, which were displayed as post-it notes. They all accessed the bookmarked pages and read the shared comments in the context. During the interview after the task, some of them confirmed that their choices were influenced by those comments.

As an illustrative example, consider the following remark sent by the receiver’s ‘partner’: “The British Library is not just a library. It holds a little piece of history like Gutenberg’s Bibles, Magna Carta, da Vinci’s notes, etc...”. Finding such information about the British Library requires a bit more effort, as the home page of the Library does not mention these interesting details. One of the participants made the following remark when making her decision: “I will choose this because has the most interesting comment, I didn’t know that the British Library was also a museum.”

This example clearly shows that if annotations are meant to provide additional information and to influence the receiver’s opinion or choices, they should be presented as such. A text snippet below the title, as provided by many social bookmarking sites, is clearly not sufficient to catch the receiver’s attention.

5.4 Re-finding Through Annotations

As discussed before, digital annotations have the advantage over paper-based annotations that they can support re-finding, given appropriate indexing and search facilities. In Study IV we investigated in what respects annotations may perform better than repeated searches and social bookmarking. As described in the previous section, our participants were randomly allocated to one of the three conditions: search engine, social bookmarking or in-context annotations.

A rather surprising observation was the extensive usage of the browser’s find functionality, which rapidly locates and highlights the given words in the page in view. Typically, participants would enter one or more keywords from the question, or – for the annotation and bookmarking participants – keywords from the answer, often by copying and pasting them from the bookmark comment or post-it note.

Fig. 3a shows that the usage of CTRL+F is far more prevalent in the search engine and bookmarking groups than in the annotation group. This is due to the fact that most participants had positioned the post-it notes near the relevant text passage and that these notes stand out from the rest of the page.

As CTRL+F is primarily meant for speeding up the finding of the relevant text passage, it would be a reasonable assumption that the average browsing time for tasks for which the functionality is used would be lower than for tasks that were solved without using CTRL+F. A
comparison, as displayed in Fig. 3b, suggests the opposite: the completion of tasks in which CTRL+F was used took significantly more time. A likely explanation would be that CTRL+F is mainly used when users encounter pages that do not allow for quick visual scanning due to their length. From Fig. 4 it can be observed that – except for the pages that were re-found using the search engine – the differences in size are not that large. An alternative explanation – which was suggested by the participants’ remarks, but which we only been able to verify by informal manual observation – is that CTRL+F is a coping strategy that allows for scanning the text for relevant items on pages that do not provide an organized structure, clear headings and anchors [21]. The smaller differences for the participants who used bookmarks and annotations suggest that this additional information leverages the issue of scannability in these situations.

![Fig. 4 Average Web page sizes (number of words) distinguishing tasks where the browser’s “find” functionality was used.](image)

### 6 Discussion

On the Web, tools and services for social bookmarking have gained much in popularity. From our discussion on related work it has become clear that these bookmarking services mainly serve the purposes of re-finding and sharing. Digital annotation tools, such as Annozilla and Diigo, provide a mixture of paper-based annotation, social bookmarking and support for collaborative work. The focus of these systems mainly lies on functionality for managing, organizing, searching, re-finding and sharing annotations.

By contrast, as discussed in section two, paper-based annotations are often also used for supporting comprehension, interpretation and thinking. Further, paper annotations may not be very helpful in re-finding a particular piece of paper, but once the paper has been found, the annotations provide important cues on what was deemed relevant.

Consequently, annotations tools need to cover several different user contexts. First, annotation tools should provide intuitive tools for marking text and writing notes in the margin – even if these annotations may not be that useful in searches for re-finding as tags would be. Results from our study indicate that private annotations – both paper-based and digital – are considerably shorter and more cryptic than shared annotations. These annotations are typically created on-the-fly to serve as a marker, or to write down some quick comments before they are forgotten. These annotations are a different breed than shared annotations, which bear more similarity to SMS or Skype messages. Therefore, private annotations should be the default option.

Arguably, keywords and tags written in social bookmarking and annotation systems provide useful information to take into account when searching through the personal history – even though it is reported that such searches are infrequent [26]. Results from Study IV indicate that enriching pages with in-context annotations has a positive impact on the scannability of the page – in particular for pages that are unstructured or contain long text passages. As a result, finding the relevant information within the page took less time and participants resorted to coping strategies, such as CTRL+F, to a lesser extent.

This implies that annotation systems should provide the means not just to attach annotations to a page, but to visually relate them to a particular fragment – similar to the post-it notes in SpreadCrumbs. The annotations needed for drawing the user’s attention to the relevant text fragments are actually the ‘annotations for thinking’ as discussed in the previous paragraph. This emphasizes the need for attaching annotations to the content and
means for highlighting text passages in such a way that it becomes a natural activity while reading. Finally, sharing links and messages are commonly supported by social bookmarking and digital annotation systems. As we have seen, writing for the benefit of someone else is different from writing private texts and therefore these two should not be mixed. Further, we observed that the way in which shared messages are conveyed by the receiver, makes a difference. In our study, comments in Delicious were often overseen, just because of the way they were formatted. Comments and messages that are meant to influence the receiver's opinion should visually stand out. A simple means for achieving this is to present the comment in a box, or to attach the sender's profile picture or thumbnail to the comment.

7 Conclusions

In this paper we discussed the results of four studies on how users create and make use of annotations, on paper or in digital media. We have seen that current digital annotation systems mainly address the goals of future re-finding and sharing – which makes them very similar to social bookmarking systems. What current annotation systems do not sufficiently support is annotating for thinking, which includes mark-while-you-read, highlighting text and attaching comments to specific text fragments. Apart from the cognitive support for understanding and interpretation while reading, these annotations enhance scannability upon later reading. From the results several design implications for annotation tools can be drawn. Annotation systems must be flexible to support a wide range user needs and user contexts that may or may not be similar to paper-based annotations. However, similar to the paper-based world, users will most likely only annotate when there is the need to do so or when it provides certain benefits. These situations do not only include professional use – learning and knowledge work – but also activities as holiday planning or personal finances. Annotating might not become a mainstream activity on the Web, but users should have the means to do so, when they want or need to. The main challenge for annotation systems is on the user interface level, balancing the classic tension between full-fledged features and ease of use. Particular attention should be paid to the question to when and to what extent annotation systems should provide and emphasize social bookmarking features. What has become clear is that private annotations have their own merits and purposes; they invoke different behaviour and serve different user needs than social bookmarking.

Acknowledgements. The authors’ efforts were partly funded by the European Commission in the EU FP7 Network of Excellence Stellar.

References

5. Bush, V. As We May Think. The Atlantic, July 1945.


Annex 3: Methodologies for the Design of Structured CSCL

APPROACHES TO STRUCTURE CSCL: THE MACRO AND THE MICRO DESIGN LEVEL

F. Pozzi, D. Persico - ITD-CNR
F. Fischer, L. Hofmann - LMU

Abstract

This study is rooted in the CSCL research field, where the debate around the use of structured activities to support interactions and collaboration among peers is quite lively. The study discusses a German and an Italian research experience, addressing similar research objectives, i.e. investigating whether and to what extent scripts can be used to make collaboration as effective as possible, with the ultimate end of improving learning through CSCL processes. The CSCL processes analyzed in the two experiences used scripts in different ways or, more precisely, at different levels, that we will call macro (for the Italian study) and micro-level (for the German one). This study discusses and compares the strengths and weaknesses of the two approaches. What can we learn from the two experiences? Is there any possibility – and with what advantages – to integrate the two approaches, so to gain from both?

1. Introduction

In CSCL contexts, the debate around how to support students’ collaboration is quite lively. In particular the focus is on whether and to what extent structuring the interactions among students can enhance the effectiveness of the collaborative process. The debate is generically referred to as “scripted vs. free collaboration”. At the last CSCL conference (held in Rhodes summer 2009), one workshop was organized around the topic “Scripted vs. free collaboration: alternatives and paths for adaptable and flexible CS scripted collaboration”. As Prof. Demetriadis (Demetriadis et al., 2009) enlightedly stressed on that occasion, the need for an international arena addressing the issue of whether and to what extent it is worthwhile structuring and guiding collaboration in CSCL is clearly emerging from the literature. Indeed, on the one side there are studies claiming that free collaboration may fail to engage all team members in productive interactions (Hewitt, 2005; Bell 2004, Liu & Tsai, 2008; all cited in Demetriadis et al., 2009); on the other side, there is a hypothesized danger of over-scripting the activities (Dillenbourg, 2002), in that too much guidance (or an excessively structured activity) may hinder learners creativity and freedom, therefore causing a loss of flexibility (Dillenbourg & Jermann, 2007). As a matter of fact, the level of structuredness of a collaborative activity may vary considerably, ranging from unstructured or moderately structured activities to highly structured ones. Some instructional design approaches suggest that the level of structuredness should be determined by choosing among different collaborative techniques at macro design level. Each technique usually specifies the nature of the task, the time schedule and the group composition. An example of moderately structured technique is the Discussion, where groups are required to perform a common task (for example collaboratively writing a document) within a given time schedule. Apart from time constraints – this is a technique where interactions among students take place freely, without pre-determined procedures. On the other hand, highly structured techniques, such as for example the Jigsaw, are those where the social structure of the groups, or the kind of task, and/or even the time schedule of the activity, are more strictly defined, so that interactions are “guided”. Examples of such techniques are the Case Study, the Role Play, the Brainstorming, the Peer Review, etc, each one presenting different degrees of structuredness. The techniques are usually chosen by the designer and adapted to the educational setting, prior to the educational event, during an instructional design process taking into consideration various variables, such as course objectives and content,
characteristics of target group and context constraints. They also meet precise teaching needs (and often also methods/practices), in just the same way as in face-to-face settings (Persico et al., 2008).

Other approaches tend to determine the level of structuredness at micro design level, where an attempt is made to affect, or even determine, the collaborative process within a given activity by using so-called collaboration scripts to guide step-by-step the interactions among students. Scripts may determine even at the finest grain the social interactions of students (social scripts), the contents to be discussed (semantic/content scripts), or the structure of the argumentation (argumentative scripts) (Stegmann et al., 2007). In this view, the theoretical idea is that learners can have different prior collaboration competence and experience. If this competence is sufficient (e.g., they know how to attack an argument in a counter-argument in a given domain), then no micro-scripting is required. If it is not sufficient, learners will need external guidance on the micro level to optimally benefit from working on the task (Kollar, Fischer & Slotta, 2007) – a problem that can be called under-scripting.

In this study two CSCL approaches are compared. The former is oriented to provide a structure to the activities at a macro-design level, the latter guides interactions through scripts at micro-design level.

The aim of the comparison is twofold: on the one hand, to reflect on the pros and cons of the two approaches in such a way to derive instructional design principles; on the other hand, to study the feasibility of a “contamination” between the two approaches and explore the possible consequences of blending them.

1.1 Structuring activities at macro design level – the Italian experience

In recent years the Istituto Tecnologie Didattiche (ITD) – CNR has been asked by the “SSIS - Scuola di Specializzazione dell’Insegnamento Secondario” (which is the Italian pre-service teacher training institution) to design and carry out a course on Educational Technology, in Italian “Tecnologie Didattiche (TD)” for the SSIS trainees. Aim of the course was to promote the development of instructional design competence and the implementation of educational technology in the school context (Delfino & Persico, 2007).

Although each year the course (hereinafter referred to as “TD-SSIS”) had its own specificities, in terms of learning objectives, contents, activities, schedule, etc., all of its editions had an online component based on a CSCL approach, where trainees were required to carry out collaborative activities under the guidance of tutors. Given the wide variety of trainees, in terms of expectations, interests, background and technological competence, the instructional design choices invariably led to a modular and flexible learning process, capable of meeting the diversified needs of a large and heterogeneous target population. For this reason, the various editions of the course envisaged an alternation between face-to-face lectures and online activities carried out using a CMC system (Centrinity FirstClass™ or Moodle, depending on the course edition). In particular, face-to-face sessions were devoted to lay the bases for both a better understanding of the subject and an effective participation in online activity, whilst online work was mainly collaborative and asynchronous. The size of the audience was usually notable (more than one hundred students per year), the trainees cohort was divided into virtual workgroups, each supported and coordinated by a tutor.

In our courses different kinds of online collaborative activities are proposed to students. Their structures also vary, but the most frequently adopted are: Jigsaw, Discussion, Role Play, Case Study. In the following, we briefly describe these techniques and the method adopted by the Italian research team to investigate their effectiveness.

In particular, the task of a Jigsaw usually requires that the tutor identifies a number of sub-topics or points of view for the subject to be learnt. Then the task is organized into two phases: during the first phase each team studies in detail one subtopic (or the whole subject form one point of view). During the second phase, new groups are formed consisting of one person for each of the previous teams. These new groups have to discuss and produce a shared document or another kind of artifact, requiring the competence developed by each member in the first phase. The organization of teams is usually quite complex and requires the tutor to be quite directive about it. It implies dividing students into a certain number of groups in the first phase of the activity (the number of groups depending on the topic and its organization in sub-topics) and a re-aggregation in new teams in the second stage of the activity, in such a way that each group is composed of at least an expert for each sub-topic.
The organization of the jigsaw groups requires respect for a rigid schedule: the groups of the second phase can only be formed if those of the first phase have finished their task and are ready to act as experts in the second phase. For this reason the time dimension is highly structured in the jigsaw activities.

The way an activity based on the Discussion is organized is rather different. In this kind of activity students are required to critically analyze a problem, a topic or a subject. The key idea in this case is to make sure learners consider it from different points of view, by taking into consideration the pros and cons of different solution methods as well as the opportunities and threats of any given situation. As a consequence, the role of the tutor is to facilitate the emerging of different ideas and points of view, fostering reflection and supporting in depth investigation of the subject. To this end, teams can be handled in a flexible way. They can be of different sizes, even of size that vary in time, and they can be formed by the learners themselves. Discussion usually needs to be ignited and should not go on forever, but there is no need to stick to the schedule too rigidly, unless there is an external constraint imposing the end of the activity by a given deadline.

In activities based on Role Play, trainees usually have to take on a role (generally they can choose a role from a pre-defined set, trying to cover all the possible roles). The role usually entails a “point of view”, in that playing a part means trying to reason as if one was that character. Some claim that learners should choose roles they are far from, so to understand others points of view. The size of the teams mostly depend on the number of roles to be played, and the roles determine their internal structure. There are no particular constraints on time, safe for external constraints. Usually students play their role throughout the whole activity, except perhaps the conclusive session.

Finally, the task is highly structured by the Case Study method. The tutor must specify quite clearly what the students are supposed to do to carry out the Case Study. For example, the learners might be required to work out a solution for a complex problem and then compare it to an experts solution, trying to carry out a sort of self-assessment. Sometimes, Case Studies can effectively be carried out through a Role Play. Teams for Case Study activities can be flexible. Trainees can choose who they would like to work with. Of course, letting the students form their teams will entail allowing them some time to do so and relying on their self-regulation in forming the partnerships. Like in Role Play, there are no particular constraints on time, safe for external constraints. To allow for a reasonable in depth analysis, time should not be too short.

The above described activities based on the different techniques have been analyzed and evaluated, aiming to identify the pros and cons of their different degrees of structuredness with respect to their ability to enhance rich interactions among learners, as well as a fruitful collaborative process.

In order to investigate the nature of the interactions occurred while performing the proposed online activities, an evaluation framework has been used (Pozzi et al. 2007; Persico et al., 2009). The model considers four dimensions as those characterizing a learning process in CSCL contexts, namely the participative, the cognitive, the social and the teaching dimensions. In the model, each dimension is defined by a set of relevant indicators that can be used to evaluate it; in particular:

- the participative dimension is defined by indicators of: Active Participation (P1), Reactive Participation (P2) and Continuity (P3);
- the social dimension is defined by indicators of: Affection (S1) and Cohesion (S2);
- the cognitive dimension is defined by indicators of: Individual Knowledge Building (C1), Group Knowledge Building (C2) and Meta-Reflection (C3);
- the teaching dimension is defined by indicators of: Organizational matters (T1), Facilitating Discourse (T2) and Direct Instruction (T3) (Persico et al., 2009).

As far as the methods and means that allow to gauge these indicators, content analysis of all the messages exchanged by the students during the activities (or a sample thereof) is usually carried out. In particular, the indicators concerning the participative dimension are gathered directly from the data tracked by the CMC system, whereas the analysis of the cognitive, the social and the teaching dimensions is based on a “manual” content analysis of the messages exchanged among students. The unit of analysis chosen is usually the “unit of meaning”, i.e. each message is split into semantic units and each unit is assigned one indicator. The coding procedure is typically carried out by two independent coders, who work separately after a period of training. In order to calculate the inter-rater reliability between the two coders (i.e. the agreement between the two), a sample of messages is then selected.
and coded. The inter-rater reliability is usually calculated using the Holsti coefficient and considering the agreement on each unit of meaning. Disagreement is solved through discussion. Once the two coders have reached an acceptable degree of accord, they split the sample of messages in two parts and work independently on the two parts. Diagrams and simple statistics are then used to compare the effect of the different activities on the learning process.

The main results obtained up to now show that the techniques with a low structure seem to foster more the social dimension, whereas those with a higher degree of structuredness seem to have more positive effects on the cognitive dimension. Data on the participative and the teaching dimensions do not show any particular relation between the level of structuredness of the activity and the development of these two dimensions. Besides, there are dynamics that tend to be invariant across techniques and seem to be intrinsic to the process itself, as to say that the kind of activity selected by the instructional designer at macro-design level may “enhance the probability that productive interactions occur” (Dillenbourg, 2004), but cannot – per se – determine a process which seems – at least in the studied conditions - to be dominated by certain constants (Persico & Pozzi, in press).

1.2 Structuring activities at micro design level – the German experience

The setting of the German is university teaching. Teachers in an educational science curriculum complained that students had major difficulties in applying the theories of their domain in an acceptable way, especially from the argumentative point of view. Instead, students rather resort to their own prior experiences in analyzing educational problems. In discussions, they hardly risk any conflict and do not formulate counterarguments that would potentially facilitate knowledge building. Over one decade, a series of over 20 experiments with more than 1000 participants have been carried out on the general question, what effects collaboration scripts could have in online discussions. In the following, we describe one study to exemplify the approach. In this study, the main question was: “to what extent can university students be supported by different kinds of collaboration scripts to apply theories of their domain argumentatively in discussing real problems and cases of their domain online?”. The study participants were students of Educational Science from the LMU, Munich, Germany, attending a compulsory introductory lecture in Educational Science. One of the 12 weekly sessions of this lecture was replaced by one online discussion. The topic of this online discussion was Weiner’s attribution theory (1985). The underlying aim of the activity was the facilitation of argumentative knowledge construction with scripts, especially by enhancing the quality of argumentation. 96 first semester educational science students were split in 32 groups of three. All students’ triads were then assigned to different experimental conditions.

The groups worked online on an 80-minute collaborative learning phase. The three group members communicated via a web-based discussion board. After having read a three-page description of Weiner’s attribution theory (1985), students had to apply it to analyze three problem cases describing motivational phenomenon and reach agreement on a final analysis. (For an example of a problem case, see Table 1). The students exchanged text messages that resembled emails.

Table 1: The “math case”, one of three problem cases learners need to analyse and discuss

| As a student teacher in high school, you participate in a school counselling session with Michael Peters, a pupil in the 10th grade. |
| Recently I’ve started to realize that math is just not my thing. Last year I almost failed math. Ms. Weber, my math teacher, told me that I would really have to make an effort if I wanted to pass 10th grade. Actually, my parents stayed pretty calm when I told them this. First mom said that nobody in our family is a math wiz. My father just kept smiling and told the story about how he cheated on his final math exams by copying from other students and using cheat sheets. “The Peters family”, he said, “has always been a math teacher’s nightmare”. Once when I was slightly tipsy at a school party, I told this story to Ms. Weber. She said that it was not a bad excuse, but not a good one either. She said it was just a number of excuses you could come up with to justify being lazy. Last year I barely made it through mathematics, so I am really nervous about the upcoming school year!” |
The computer-supported learning environment contained a discussion board with a main page with an overview of all message headers. The learners could read the full text of all messages, reply to the messages, compose and post new messages.

In a first experiment, we implemented a 2X2 factorial design to compare the effects of an epistemic script, a social script, and their combination. In the second of the selected experiments, we implemented a 2X2 factorial design to compare the effects of (b) an argument construction script, an argument sequencing script, and their combination.

Dependent variables were measured in the process and as outcomes: In the process, we focused on three central aspects in discourse (see Weinberger & Fischer, 2006), namely the **epistemic dimension** points to how learners work on their task; the **argument dimension** refers to the way learners construct arguments and sequence them in their discussion, whereas the **dimension of social modes** of co-construction indicates how learners built on the ideas of the learning partners. With respect to outcomes, we measured knowledge on argumentation as well as knowledge on the domain (motivational theory) in pre- and post-tests. Different tests on factual knowledge (multiple choice tests) and application-oriented knowledge (individual case analyses) were administered.

The epistemic script structured the knowledge building task, i.e. it asked questions with respect to the cases that guided the groups to apply Weiner’s theory step by step to the case information.

The argumentation construction script supported the learners to formulate sound arguments according to a simplified Toulmin model (1958). Main focus was to facilitate the distinction of grounds, claim, and qualifications. Text windows with prompts were provided as scaffolding.

An argument sequencing script facilitated the classical dialectic sequence of argument, counter-argument, and integration, the effect of which on knowledge building had been recently shown (Leitão, 2000).

A social script realised a peer review procedure, where each learner analysed one case using Weiner’s theory, then received two critical feedbacks, revised the initial analysis, followed by a second round of feedback and then a final analyses submitted by the learner responsible for this case.

![Figure 1: The ideal final structure of the discussion board when structured by argument sequencing script](image-url)
Main results were that of the collaboration script types improved processes in a dramatic though highly specific way when compared to unscripted control conditions (e.g., the epistemic script strongly improved the systematic application of theory concepts in the case analyses, the argument construction script strongly improved the formal quality of the arguments). With respect to outcomes, the two argumentation-related scripts improved knowledge on argumentation in comparison to unscripted control condition, but not domain knowledge. Only the social script had positive effects on the domain-specific knowledge as well. The epistemic script had even negative effects on domain knowledge – students did solve the problems well together, when supported by the script, but they did not internalize the concepts and the strategy of analysis.

2. Discussion

In this section we set the two approaches one against the other and try to identify their main strengths and weaknesses, aiming to reflect on whether and to what extent it would be useful to blend them, so to take the best from each of them. As we have seen in the previous sections, scripts and techniques reify different design approaches to structure collaboration, yet they are both describable in terms of Task, Team and Time, i.e. both the described experiences point at a Task to be performed by students, the necessity to build up Teams to carry out the Task and Time constraints according to which the Task has to be accomplished. What differentiates the two approaches is the level of structuredness of each of the three above mentioned dimensions. At macro level the Task, Team and Time structuredness can be variedly associated (for example the Case Study is highly structured as far as Task is concerned, but Teams and Time are low structured while in the Discussion all of the three dimensions are low structured,) and they can be, to some extent, changed (or tuned) on the fly. On the other hand, at micro level the Task is always highly structured, while Team and Time are strictly pre-determined a priori. The first consideration that can be drown is that structuring activities at a macro level lets a wider margin of freedom to both students, whose interactions are not completely guided even when the activities are highly structured (as for example in a Jigsaw), and tutors, who may intervene and tune the micro-design of the activities at run time, by providing stimuli to enhance those dimensions of the process that result to be weak. This – on the other hand –
requires a certain effort by the tutor, whose workload is usually rather heavy in these contexts.

The use of scripts at micro design level, on the other side, allows a more structured process, where there is little chance for the students to be “disoriented” by the task (as this is usually detailed step by step) and where the tutor’s contribution (and effort) is limited (if any). It is no coincidence that the degree of task structuredness is nearly always high in these activities. The downside is that scripts may prevent students from expressing themselves freely, as their degree of freedom is determined a priori.

Choosing between macro and micro level designs may depend on the target population: the former approach is more useful when students are adult and have already developed a good degree of self-regulation in their learning, as well as the ability to argue in favour of their positions and to think critically. In the scripting terminology, if learners have access to mentally represented collaboration scripts (e.g., knowledge on how to differentiate observed data from their interpretation in formulating an argument, or how to attack an argument, or how to give critical feedback) then the macro-structure is enough to coordinate the activities in classrooms and small groups. An additional micro-level script could determine an overscripting problem. But if this knowledge is not available to students, simply telling them that they should discuss or role-play or play the expert in a jigsaw will most probably not yield good learning outcomes, even if the process is apparently flawless. This could be underscripting. Adopting a micro level approach, on the contrary, is more advisable when students are younger and need support to improve their ability to discuss and argument.

Another argument for the decision whether structuring on the micro or macro level comes from another difference between the two approaches. Micro-scripts are typically meant to be internalised as cognitive strategies in a Vygotskian sense (higher cognitive functions first appear on a social plane, then on the individual plane). From this perspective, a dialectic process of argument, counter argument and integration is not only there to improve the elaboration of the content argued about; in addition, students should internalise this strategy as a cognitive tool of critical thinking. Macro scripts are didactic aids to learn content and are often not meant to be internalised as cognitive strategies that are supposed to be already at the disposal of the individual.

Nonetheless, looking at the two approaches described in this study, a question arises, on the opportunity and the possible benefits that could derive from blending them in some contexts. Yet, it seems that no experiences are reported exploring such a direction. This may depend mainly on the fact that the theoretical background and often even the context where one research team operates, tends to impose a view, in such a way that the research problem, which still is common to the two approaches (e.g. “How can I effectively support the collaborative learning process?”), is usually tackled either at the macro or at the micro level. In this respect, we should acknowledge that there is a certain fragmentation in this field, and the existence and continuous development of new approaches, models, tools and even projects, that share the same aims but fail to really integrate and exchange their results, is a proof of this.

Starting from this panorama, this study, originated within the STELLAR Network of Excellence, is a step forward towards the real contamination – or even integration – of approaches defined by two different research teams. It is in fact a starting point and an excellent chance to speculate on blending a macro and a micro level approaches (as far as theories, designs and research methods are concerned), in order to improve both the structuredness and the flexibility of the process.

3. References


Annex 4: Identifying knowledge development - a context-aware approach for analyzing knowledge processes

Authors

Know-Center

Stefanie N. Lindstaedt  Ulrike Cress Andreas S. Rath Johannes Moskaliuk Nicolas Weber Joachim Kimmerle Didier Devaurs

Abstract

When people use wikis to work jointly on shared digital artefacts, this may lead to the collaborative creation of new knowledge. The consideration of the digital products and the insights into the construction process itself may lead to a better understanding of knowledge-building processes. In turn, this understanding of knowledge building may help to design environments that support the improvement of knowledge building. Therefore, the central research question for our paper is twofold: on the one hand it identifies indicators for knowledge maturity; on the other hand it provides some considerations on how knowledge maturing may be supported. In order to provide answers to these questions, we conducted a study in which participants had to work with a wikitext with the instruction to improve the text and add further arguments. On the basis of the users' behaviour, we developed an ontology-based task detection approach that identified knowledge maturing processes with a rate of 79.12% (the findings of the tool were gauged with the ratings of two experts who evaluated people's actions with regard to knowledge maturing). The findings are discussed against the background of the question on how it can be ensured that knowledge workers contribute to the development of knowledge.

1 Introduction

Using wikis or other social software tools for knowledge work may lead to new forms of collaboration and learning (Raitman, Augar, & Zhou, 2005; Reinhold, 2006): Individuals can use wikis to work jointly on shared digital artifacts that were provided through the Internet or a local area network, say, of an organization. This will not only lead to accumulation of knowledge, by which the knowledge of many individuals is brought together and made available to others, but also to knowledge emergence, i.e. the creation of new knowledge (Johnson, 2001; Moskaliuk & Kimmerle, 2009). Thus, writing wiki text in a collaborative way is not only a method to share information but also to construct new knowledge (Cress & Kimmerle, 2008). During the process of writing as a joint activity, users generate new ideas and innovation and discuss own arguments with others. They construct shared meaning and build a mutual understanding (Erkens, Jaspers, Prangsma, Kanselaar, 2005). This may lead to an evolution of knowledge. Emergent knowledge that was not part of the individual knowledge of a single user before may arise during this collaborative activity.

Thus, learning is always embedded in a social and cultural context (cf. Lave & Wenger, 1991; Vygotsky, et al., 1987). Shared digital artefacts like wikis or other social software tools are external representations of this social and cultural context and contain the knowledge of the corresponding community. External representations of knowledge reflect and stimulate
individual learning processes within a community or an organization and lead to a development of knowledge over time. The concept of knowledge maturing (Schmidt, 2005) describes this phenomenon from a macroscopic perspective. It considers how external knowledge representations (such as digital artefacts) correspond, are derived from or inform internal knowledge representations. The concept of knowledge maturing describes how knowledge matures from the level of individuals to the level of communities, and, finally, to the level of organizations. It focuses on how knowledge matures from expressing individual ideas to formalizing knowledge on an organizational level.

The process of knowledge maturing is condensed in digital artefacts (e.g. in a wiki text). With the help of these digital artefacts it becomes possible to evaluate the amount of knowledge that is available in the community of participating users (e.g. a class or an organization). Additionally, by considering the writing process closely, the knowledge maturing process itself may be analyzed and new insights may be gained into how people construct new knowledge. This consideration of the digital products and the insights into the construction process itself may lead to a better understanding of knowledge maturing. In turn, this understanding of knowledge maturing may help to design environments that support the improvement of knowledge within organizations, e.g. by giving feedback on the current status of a text.

These considerations lead to the idea that the quality of a text and its maturing over time is not only an important goal of each knowledge organization, but may also be used as an indicator for evaluating successful knowledge building. Therefore, the central research question for our paper is twofold:

(1) Maturity indicators: How can the quality of a text be measured and used as an indicator for knowledge maturing processes within communities?

(2) Knowledge Maturing support: How can it be ensured that knowledge workers write good texts and contribute to the development of knowledge?

In order to provide empirical answers to these questions, we conducted a study to collect data as basis for an ontology-based task detection approach. In this study participants had to work with a wiki, improve the text and add further arguments. The participants obtained additional information that contradicts the content in the wiki. This should provoke knowledge maturing and lead to prototypical editing tasks we could use to train our task detection tool and cross-validate its results. The goal is to identify features for knowledge maturing, which will make it possible to measure behaviour automatically. The results of this study are supposed to lead to a better understanding of knowledge maturing processes in a wiki and of the factors that influence them.

In the following we will first of all explain the theoretical background and introduce a framework model that explains the development of knowledge over time as co-evolution of cognitive and social systems (section 2). We will distinguish between accommodation and assimilation as two essential processes of knowledge building. Then we describe the research setting we designed and how we use it to develop and validate our task detection approach (section 3). In section 4 we describe our task detection approach in detail. We present the results in section 5. As a conclusion we discuss the results of the task-detection approach against the background of the question on how it can be ensured that knowledge workers contribute to the development of knowledge.

2 Theoretical Background

The co-evolution model of cognitive and social systems (Cress & Kimmerle, 2008) may be understood as a framework to describe knowledge processes that take place in a wiki. This model describes individual learning and collaborative knowledge building as a co-evolution between cognitive and social systems. The co-evolution model (Cress and Kimmerle, 2008) considers two relevant systems: The social system wiki and the cognitive system of a user.
Both systems are independent from each other and build their border to the environment with help of a specific mode of operation. Here, the authors refer on Luhmann's systems theory (Luhmann, 1986, 1995, 2006) and transfer his ideas to the context of knowledge building with wikis. The social system means the communication within a (virtual) community that becomes manifest as written text in a digital artefact (e.g. the wiki). The cognitive system consists of psychological processes like learning, reasoning, problem-solving or perception.

According to this model, both the social and the cognitive system develop over time and become more and more complex. The information in a wiki (that represents the knowledge of a community) evolves in the course of time. At the same time, the knowledge in an individual’s cognitive system increases. This mutual development as a co-evolution of social and cognitive systems can be understood as knowledge building or knowledge maturing. Cognitive conflicts (Piaget, 1977a, 1977b) that an individual perceives are considered as the key incitement factor of this co-evolution. In the sense of Piaget a cognitive conflict occurs if new information from the environment does not fit existing knowledge. These cognitive conflicts motivate individuals to contribute to the wiki or to change their own knowledge structure in order to establish equilibrium between own knowledge and new information.

The model by Cress and Kimmerle (2008) specifies the co-evolution process and describes two different processes on the basis of the ideas of Piaget: assimilation and accommodation. Assimilation means active shaping of the environment by interpreting and explaining current experiences, giving them a place in existing schemata. Accommodation means adaptation to the environment in the form of qualitatively changing one’s own cognitive schemata. In Piaget’s understanding assimilation and accommodation are processes in the cognitive system; the co-evolution model, however, expands Piaget’s point of view by describing accommodation and assimilation not only from the perspective of an individual’s cognitive system, but also from that of a social system. Users assimilate information from the artefact into their own cognitive schemata, and they accommodate by modifying their schemata induced by information from the wiki. Analogous processes of learning and knowledge building may take place in the social system: in the case of assimilation users add pieces of information from their own knowledge. This, however, will not change the basic message and structure of the wiki, but only add supplementary aspects. Accommodation is also possible in a wiki if users contribute their knowledge in such a way that the message is changed and, sometimes, new structures are being created. Accommodation tends to result in some qualitative modification of the artefact, whereas assimilation has primarily to do with quantity, introducing additional arguments or new examples but no fundamental innovation.

To sum up, the co-evolution model is considered as useful framework to describe how individuals use digital artefacts like wikis to jointly construct knowledge. The constructed knowledge manifests in digital artefacts and is therefore accessible for deeper analysis. In order to identify features for assimilation and accommodation actions we have to consider two relevant dimensions: The artefact dimension and the usage dimension (Mentzas, 2007).

The first dimension, content artefacts, provides a static picture of the world and is probably the best managed type of knowledge entity. It can take the form of notes, contributions, threads, protocols, lessons learned, learning objects, courses, pictures, videos, podcasts, etc. In many organizations textual contents are the most prevalent contents. For content maturing support, we rely on knowledge discovery algorithms mainly using statistical methods and some shallow natural language processing for text mining and text analysis. These methods extract features in form of feature vectors from a textual information object, from which textual similarity measures can be derived. These in turn are then used for text classification, clustering and for other kinds of operations.

The analysis of external knowledge representations would not be sufficient considering the analysis and support of organizational knowledge maturing processes. Hence, a further issue that has to be taken into account is the way in which these external knowledge representations are being used. In the usage dimension we seek to learn from the behaviour
of people interacting with digital artefacts. The interaction traces can tell something about the person (e.g. the role, interests, knowledge or skills the person is likely to have), about the knowledge asset the person is dealing with (e.g. the context or task a knowledge asset was created, used, changed or shared), and about the activities performed within an organization (e.g. process and task executions).

Looking at the artefact dimension and the usage dimension we are able to gain new insights how individual use texts to build knowledge and how the knowledge matures over time. Integrating the theory of Cress and Kimmerle (2008) with the knowledge maturing idea we suggest that the process of accommodation is a central maturity indicator. This is true for both, the cognitive system and the social system. Assimilation means a quantitative development of knowledge where new information only completes existing knowledge as accumulation of knowledge. Accommodation leads to better understanding of complex information, produces integrative mental models, new ideas and innovation, or emergent knowledge. This can be described as qualitative development of knowledge. The quality of text depends on accommodation processes. We propose that emergent effects usually occur through processes of accommodation in artefacts. This will lead to a higher complexity of the wiki and, accordingly, to knowledge processes in other people’s cognitive systems.

In addition to psychological approaches for measuring individual knowledge in the cognitive system (e.g. the experiment by Moskalik, Kimmerle, & Cress, 2009) it is interesting to take into account the processes in a digital artefact as indicators for processes in the social system. If we measure the quality of text and its development over time we can conclude underlying knowledge-building processes within communities (maturity indicators) and can use this results as feedback for users to support knowledge building (knowledge maturing support).

The main goal of the current study is to examine how accommodation (as opposed to assimilation) happens in detail within a wiki system and how knowledge maturing may be analyzed in an automatic way. We started with a qualitative assessment of user interactions on knowledge artefacts based on expert observations. In order to identify features for accommodation actions we designed a study where participants had to work with a wiki and add new information. Then, we developed an approach for the selection of (quantitative) interaction-based features in order to detect accommodating activities as maturity indicator.

### 3 Research Setting

We designed a research setting in which participants had to work with a wiki and add new information. The wiki and the additional information were presented at people’s computers. The additional information was presented in the form of short arguments that participants could browse. In order to induce cognitive conflicts, the existing information in the wiki and the new information contradicted each other. This was supposed to provoke accommodation processes and is therefore considered an ideal condition to train the task detection tool.

All runs took place in a silent work environment in a secluded room during daytime. Participants had a total of 50 minutes to edit the wiki, the complete study lasts about 90 minutes. Participants were instructed to use the provided laptop computers and that it should solely be operated via the attached mouse and the keyboard and not by using the touchpad. Participants then started by viewing a welcoming page after which they had to click through several information and instruction pages regarding further details about the study and the difference between qualitative (accommodation-like) and quantitative (assimilation-like) changes they would make. Participants were instructed to save their changes after each set of alterations belonging together (editing task). Participants’ task was to complete a given text of a wiki (initial state text), so that a scientifically balanced article about violent computer games and their possible danger for users and the society would result. Two windows were available for the participants during the study: a wiki page and a page with the additional
information. Those two windows were accessible through two tabulators at the top of the page. The wiki page (wiki-tab) presented a text that, initially, was biased toward a contra violent computer games position. Here, only the risks and dangers of violent computer games were presented. The page with the additional information (info-tab) contained ten different arguments that were biased towards a pro violent computer games position. These arguments invalidated the arguments in the wiki text or explained why the contra points were one-sided or wrong. Participants could copy, paste, delete and edit the text freely or type in new text.

Participants are instructed to click on save each time they finished alterations belonging together. We expect that this would lead to single editing tasks. Each editing task could be rated on the dimension of assimilation and accommodation by the participants themselves (user-ratings), by experts on the same dimensions (expert-ratings), and could be measured using readings-scores.

**User-Ratings:** Whenever a participant clicked on “save”, a box popped up. In this box, one had to make two rating of the changes between the last and the present saving point (one single transformation), on a five-point Likert scale. The first rating is about “qualitative text-changes”, scale ranging from “no qualitative changes” to “many qualitative changes” (accommodation-like transformation), the second is about “quantitative text-changes”, scale ranging from “no qualitative changes” to “many qualitative changes” (assimilation-like transformation).

**Expert-Ratings:** Each transformation was rated by two independent experts on the same dimension like the user-ratings (qualitative text-changes; quantitative textchanges) on a five-point Likert scale. The overall correlation between the ratings of the two experts was significant. The correlation for the two accommodation ratings was $r_{acc}(156)=.797$, $p<.01$; for the assimilation rating, the correlation was $r_{ass}(156) = .741$, $p<.01$. For further analysis we build an average of the two rating for each transformation.

**Reading-Scores:** The objective of computing reading scores is to analyze content to facilitate the assessment of the maturity of a document. Reading scores are calculated from quantitative metrics like sentence length, number of syllables or number of words. We applied the *Flesch Reading Ease test* (Si & Callan, 2001), the *Gunning fog index* (Gunning, 2004), and the *Flesch–Kincaid readability test* (Flesch, 1948) to each transformation. More specifically, we computed the reading scores on the state of the document before the transformation and after the transformation. For each reading score we subtracted the respective resulting values, which we used as one of the variables for the correlation. The second variable for the correlation computation was the average over the accommodation ratings of the two experts. Since the used variables showed no normal distributions we used the rank order correlations. The results showed no significant correlations.

The goal of this research setting was to collect data on the basis of which we could develop an approach for the selection of interaction-based features in order to detect accommodating activities as maturity indicator. We used a context-aware system for observing users in performing changes to the wiki page. This observation is on a fine granular level that takes people’s interactions with the wiki into account. For discriminating different levels of accommodation for the observation wiki transformations, we apply our ontology-based task detection approach. We describe this approach in the following section.

### 4 Task Detection Approach

Context-aware (or sentient) systems are systems that can adapt their operations or behaviour to their current context of use, without explicit user intervention. Context awareness thus enables to increase the usability and effectiveness of a system by taking into account environmental elements (such as time or location), individual and organizational elements (such as user’s identity or position), as well as elements relative to the interactions...
of the user with the system (such as pressed button or entered character). The first context-aware systems, developed in the 1990s, were mainly focused on providing functionalities specific to the user’s location (Baldauf et al., 2007). Today’s context-aware systems are much more sophisticated, and integrate complex mechanisms for the acquisition and storage of context, the abstraction and understanding of context, and the adaptation of the system behaviour based on the recognized context.

Context information may be gathered in a variety of ways, such as applying (physical or virtual) sensors, recording network information and device status, or browsing user profiles and organizational databases. Then, a context model is needed for storing the recorded user context data in a machine processable form. Various context model approaches have been proposed, such as key-value models, markup scheme models, graphical models, object oriented models, logic-based models, or ontology-based models (Strang & Linnhoff-Popien, 2004). However, the ontology based approach has been advocated as being the most promising one (Strang & Linnhoff-Popien, 2004; Baldauf et al., 2007) mainly because of its dynamicity, expressiveness and extensibility.

An important aspect of the user’s context is the current task she is performing. Detecting the user’s task enables to provide her with personalized and relevant support (Dey et al., 2001; Coutaz et al., 2005). By task detection we mean task class detection also referred to as task classification, as opposed to task switch detection. Task switch detection involves predicting when the user switches from one task to another (Shen et al., 2009). Task classification deals with the challenge of classifying usage data from user task execution into task classes or task types. Automatic task detection is classically modelled as a machine learning problem, and more precisely a classification problem. This method is used to recognize Web based tasks (Gutschmidt et al., 2008), tasks within emails (Shen et al, 2006) or tasks from the complete user’s computer desktop (Shen et al., 2006, 2009; Lokaiczyk et al., 2007; Rath et al., 2009). This relates to the present study in the sense that we consider accommodation and assimilation processes as being refined editing tasks that we would like to detect.

4.1 Automatically detecting Accommodation

Solving our editing task classification problem is done based on the following steps (illustrated by Figure 1): (i) The user context data is captured by system and application sensors. (ii) Features, i.e. parts of this data, are chosen to build classification training instances, which is done at the task level. (iii) To obtain valid inputs for machine learning algorithms, these features are first transformed into attributes. This transformation may include data pre-processing operations, such as removing stop-words and application specific terms, or constructing word vectors. (iv) Attribute selection (optional step) is performed to select the best discriminative attributes. (v) Finally, classification/learning algorithms are trained and tested.

![Figure 1: This figure visualizes the complete user interaction context ontology task detection pipeline (UICO pipeline) starting from (1) the automatic unobtrusive user interaction observation mechanisms to](image)
detecting the user's task. The automatic population of the user interaction context model (2) is displayed in two ways (i) the instantiation of entities in the conceptual model (conceptual view) and (ii) in the ontology model (ontology view). In (3) the feature engineering process for transforming a task instance into a training instance is shown. The training instance is further fed to attribute selection and learning algorithms for (4) detecting the task.

4.2 User Interaction Context

The first step of the detection process consists in capturing the "user context". Our view of the "user context" goes along with Dey's definition that context is "any information that can be used to characterize the situation of entities that are considered relevant to the interaction between a user and an application, including the user and the application themselves" (Dey et al., 2001). We refine Dey's perspective by focusing on the user interaction context that we define as "all interactions of the user with resources, applications and the operating system on the computer desktop. Resources are digital artefacts on the computer desktop, e.g., documents, web pages, e-mails, persons, appointments and notes." (Rath et al., 2009). In the case of the wiki environment applied here both tabs (wiki-tab and info-tab) were perceived as an unique resource. Context observation mechanisms are used to capture the behaviour of the user while working on her/his computer desktop. Low-level operating system and application events initiated by the user while interacting with the desktop are recorded by context observers, also referred to as context sensors. We distinguish between system and application sensors, based on the origin of the data they deliver. For capturing the user interaction context in the wiki environment of the present study we employed context sensors that have already been utilized in our previous user interaction context observation (Rath et al. 2008) and task detection experiments (Rath et al., 2009). We developed sensors for Macromedia Flash which was the base technology for the wiki editor and the page with the arguments (info-tab). Additional fine-granular user interaction context information included (i) switch to info-tab, (ii) switch to wiki-tab, (iii) text formatting (BOLD, UNBOLD, ITALICS, UNITALICS, UNDERLINE, UNUNDERLINE, COLORCHANGE, ALIGN_LEFT, ALIGN_RIGHT, ALIGN_CENTER, ALIGN_BLOCK, FONT_INCREMENT, FONT_DECREASE, FONT_NAME_SWITCH), (iv) text editing, (v) text selection and (vi) selection of a specific argument on the info-tab. The content of the wiki page as well as the text around the cursor were also recorded for each single user interaction.

The conceptual representation that we propose for the user interaction context is a semantic pyramid. At the bottom of the pyramid are events that result from the user's interactions with the computer desktop. Above events are event blocks, which are sequences of events that belong logically together, each event block connecting the user's actions associated with a specific resource acted upon. At the top are tasks that are grouping of event-blocks representing well defined steps of a process that cannot be divided into sub-tasks, and in which only one person is involved. The layers of the semantic pyramid represent the different aggregation levels of the user's actions. The semantic pyramid is illuminated in Figure 2.
4.3 User Interaction Context Ontology (UICO)

A context model is needed for storing the user context data in a machine processable form. We have defined a user interaction context ontology (UICO) (Rath et al., 2009a) which is illustrated in Figure 3. The UICO contains 88 concepts and 272 properties. It is modelled in OWL-DL, by using the Protégé ontology modelling tool. The ontology web language (OWL) is a W3C standard for modelling ontologies widely accepted in the Semantic Web community. From the 272 properties there are 215 datatype properties and 57 objecttype properties. From a high-level perspective, the concepts of our ontology can be grouped into five different dimensions: the action dimension, the resource dimension, the user dimension, the information need dimension and the application dimension. Figure X2 illuminates the user interaction context ontology.

The action dimension consists of concepts representing user actions, task states and models. The Action concept is refined by the sub-concepts Event, EventBlock and Task. Examples of user Event sub-concepts are Print, Close, Save, Copy, Paste, Cut, WebSearch, Post, Reply and Forward. An Event concept stores information about the application in which it happens, through the datatype properties hasWindowTitle, hasApplicationName and hasProcessId. The different types of task states are borrowed from the Nepomuk Task Management Model (Grebner et al., 2007). In this model, a task can be New, Running, Suspended, Completed, Terminated, Finalized or Archived. The only model available at the moment is the TaskModel which is used to categorize a task.

The resource dimension contains concepts for representing resources on the computer desktop. Examples of possible Resource are File, TextDocument, Presentation, Spreadsheet, E-Mail, Folder, Person, Location, Organization and OnlineResource. Relations can be defined between concepts of the resource dimension and of the action dimension for modelling on which resources what kind of user actions are executed, via the objecttype
property isActionOn.

The **user dimension** contains only the *User* and *Session* concepts. It is related to the action dimension in the sense that each *Action* is associated with a *User* via the objecttype relation *hasUser*. The *Session* concept represents session and user login information.

The **information need dimension** represents the context-aware pro-active information delivery aspect of the UICO. An information need is detected by a set of fixed rules based on the available context information. The *InformationNeed* concept has properties to define the accuracy of the detection and the importance to fulfil the information need in a certain time-frame. For details about information need detection it is referred to (Rath et al., 2007). This dimension is not used during our wiki study.
The **application dimension** is a "hidden" dimension because it is not modelled as concepts in the UIOC. This dimension is present is such a way that each user interaction happens within the focus of a certain application, e.g., the user's desktop, Microsoft Word or the Microsoft Windows Explorer. The Event concept holds the information about the user interaction with the application by the datatype properties `hasApplicationName` and `hasProcessId`. Standard applications that run on the Microsoft Windows desktop normally consist of graphical user interface (GUI) elements. Also console applications have GUI elements such as the window itself, scroll bar(s) and buttons for minimizing, maximizing and closing the application. Most of the GUI elements have an associated accessibility object, which can be accessed by context sensors. Datatype properties of the `Event` concept hold the data about the interactions with GUI elements. A resource is normally accessed and manipulated by the user within an application hence there is a relation between the resource dimension and the application dimension. This relation is indirectly captured by the relation between the resource dimension and the action dimension, i.e., by the datatype property `hasApplicationName` of the `Event` concept. For a user it is not convenient to manually enter the data about her context on such a fine-granular level. Hence semi-automatic and automatic mechanisms are required to ease the process of 'populating' the user interaction context ontology.

### 4.4 User Interaction Context Ontology Population

The contextual information sent by the context sensors is used as a basis for populating the context ontology, i.e. instantiating its concepts. The `Event` concept can be directly instantiated by the sensor data. In order to instantiate the `EventBlock` concept, events have first to be aggregated, using application-specific as well as generic static rules and heuristics. As an event-block represents a sequence of events associated with the same resource, this aggregation process heavily relies on the resource discovery process.

We use three different techniques for discovering resources. (i) The *regular expression* approach identifies resources in the sensor data based on specific character sequences predefined as regular expressions. This is used to identify files, folders, web links and email addresses for example. (ii) The *information extraction* approach extracts person, location and organization entities in text based elements of the sensor data, using the *KnowMiner framework* (Granitzer 2008). (iii) The *direct resource identification* approach finds data about a potential resource directly in the sensor data, and build the resource by directly mapping certain fields of the sensor data to properties of the Resource concept. These three techniques produce what we call *used resources*, in the sense that the user has interacted with them. We are also interested in unveiling relations among these *used resources*, or between these resources and other resources. We say that a resource is an included resource if its content is part of the content of another resource. A resource is a referenced resource if it is mentioned and identified in the content of another resource (e.g. names of persons, locations and organizations, paths of folders and files, URLs of web pages and email addresses). For the wiki environment the third type of resource discovery, the direct resource identification, plays a key role. The special wiki context sensor as described in the previous section provided the information on which elements of the wiki edit page or on which argument the user was working on.

The information about these elements sent by the context sensor was used to directly construct resources in the ontology with a unique URI.

A rule-based aggregation of user actions into tasks might be a reasonable approach for well-structured tasks, such as administrative or routine tasks. But is obviously not appropriate for tasks that involve a certain freedom and creativity in their execution which is the case for text editing and text manipulation. To handle such tasks the idea is to automatically extract tasks from the information available in the using context ontology by means of machine learning.

---

techniques. Once detected, these tasks will also populate the ontology.

4.5 Feature Engineering

50 features were engineered based on the concepts and relations of the user interaction context ontology (UICO). We have defined 50 features that can be grouped in six categories: (i) ontology structure, (ii) content, (iii) application (iv) resource, (v) action and (vi) switching sequences. The ontology structure category contains features representing the number of instances of concepts and the number of datatype and objecttype relations used per task. The content category consists of the content of task-related resources, the content in focus and the text input of the user. The application category contains the classical window title feature (Oliver et al. 2006; Shen et al., 2006; Lokaiczyk et al., 2007; Granitzer et al., 2008) and graphical user interface elements (accessibility objects15) features. The resource category includes the complete contents and URIs (URLs) (Shen et al., 2006) of the used, referenced and included resources, as well as a feature that combines all the metadata about the used resources in a 'bag of words'. The action category represents the user interactions and contains features about the interactions with applications (Granitzer et al., 2008), resources types, resources, key input types (navigational keys, letters, numbers), the number of events and event blocks, the duration of the event blocks, and the time intervals between event blocks. The switching sequences category comprises features about switches between applications, resources as well as event and resource types.

We use the machine learning toolkit Weka (Witten & Frank, 2005) for parts of the feature engineering and classification processes. The following steps are performed to pre-process the content of text-based features (in this sequence):

(i) remove end of line characters, (ii) remove markups, e.g. \&lg and \[CDATA,
(iii) remove all characters but letters, (iv) remove German and English stopwords, (v) remove words shorter than three characters. We transform text based features into vectors of words with the StringToWordVector function of Weka. For numeric features, we apply the Weka PKIDiscretize filter to replace discrete values by intervals. A complete listing of all features with a short description is given in Table 1.

<table>
<thead>
<tr>
<th>Id</th>
<th>Feature</th>
<th>Id</th>
<th>Feature</th>
<th>Id</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Action Category</td>
<td>20</td>
<td>semantic type</td>
<td>37</td>
<td>UICO objecttype relation</td>
</tr>
<tr>
<td></td>
<td>EB duration</td>
<td>21</td>
<td>action type of E</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>resource types interaction</td>
<td>22</td>
<td>accessibility object name</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>control input keys</td>
<td>23</td>
<td>accessibility object description</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>number of EBs</td>
<td>24</td>
<td>accessibility object role</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>included resource interaction</td>
<td>25</td>
<td>accessibility object role description</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>referenced resource interaction</td>
<td>26</td>
<td>accessibility object value</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>resource interaction</td>
<td>27</td>
<td>accessibility object help</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>letter input keys</td>
<td>28</td>
<td>accessibility object help topic</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>task duration</td>
<td>29</td>
<td>application name</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>applications interaction</td>
<td>30</td>
<td>window title</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>navigation input keys</td>
<td>31</td>
<td>raw event source</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>EB resource interaction</td>
<td>32</td>
<td>content of EB</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mean EB duration</td>
<td>33</td>
<td>content in focus</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mean time between EBs</td>
<td>34</td>
<td>user input</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>used resource interaction</td>
<td></td>
<td>Ontology Structure Category</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>action element of E</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>median time between EBs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>number input keys</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>median EB duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: This table shows all the 50 features classified into 5 feature categories. All the features were extracted based on the user interaction context ontology (UICO). E and EB stand for event and event block respectively.

4.6 Task Detection

The fourth step of the ontology-based task detection pipeline is to detect the task based on the features engineered from the populated UICO. In case of this study, for each transformation of a study participant a new training/class instance is built. The training instances are given to machine learning algorithms to train/build a classification model and then to decide to which class another training instance belongs to.

5 Results

The participants were 10 graduate students from Germany their mean age was 25.30 (SD=0.51). 4 of these were women, 6 were men. During the wiki study we recorded a dataset of 158 editing tasks. Each participant saved on average 15.8 (SD=7.29, min=7, max=26) editing tasks.

5.1 Expert Ratings & User Ratings

As reported above the correlations between the two independent experts for each transformation were on a high level. We understand this as evidence for the validity of the expert ratings and as a hint that it is possible to estimate one single transformation on the basis of the concept of accommodation and assimilation. However, there was also a high correlation between assimilation and accommodation for the expert ratings: \( r(156)=.826, p<.01 \). This is basically in line with the results from Moskaliuk et. al (2009) who reported correlations between assimilation and accommodation on a medium level. On basis of this result we assume that assimilation and accommodation processes are tightly connected. One can only integrate new information in the existing wiki on the wiki-tab if as a first step new information is added from the info-tab to the wiki-tab. In our setting assimilation is a precondition for accommodation and its therefore not possible to measure this two processes independently.

We expected a high correlation between expert and user ratings. Contrary to our expectations the correlations between user ratings and the (average) expert-ratings for assimilation are non-significant: \( r(156)=.112, p<.01 \); However, looking at the correlation between expert and user rating for single user we found a high variance from \( r_{user3}(12)=-.611, p<0.05 \) to \( r_{user2}(17)=.783, p<0.01 \).

The correlations for accommodation between user ratings and the (average) expert ratings are significant but only on a medium level: \( r(156)=.436, p<.01 \). The results regarding the correlation for single user show also a high variance from \( r_{user9}(6)=.250, n.s \) to \( r_{user10}(5)=.810, p>.05 \) but are mostly on a medium positive level.

The low correlations between user and expert ratings for assimilation and accommodation leads to the assumption that the participants are, against our assumption, not able to rate their own transformation as assimilation or accommodation. One possible explanation is the different granularity of the participants' transformations (the variance of transformation per participants reaches from 7 to 26), which made it hard to rate it in an appropriate way. This argumentation is supported by the overall correlation between assimilation and accommodation for the user ratings: \( r(156)=-.194; p<0.05 \). The range of the correlations for single participants shows also a high and non-systematic variance.

As a consequence of these results we decided to build the classes for the machine learning pipeline based on the average (accommodation) ratings of the two experts. This is described in detail in the following sections.

5.2 Automatic Detection of Accommodation Levels

For each editing tasks of a study participant a new training/class instance for the machine learning algorithm is built based on the recorded usage data. During the study we recorded a dataset of 158 transformations. For 19 transformations the log files show now difference
between the two versions and were excluded form further analysis. Based on the average accommodation ratings of the two experts we divided the 139 transformations in equal thirds using the terzils. The leads to three classes; we computed the boundaries of the classes with low\(\leq 1.5\); medium\(\leq 2.5\); high\(>2.5\) (based on the average accommodation ratings of the two experts). This leads to the following 3 classes representing different levels of accommodation: high accommodation (54 class instances), medium accommodation (55 class instances), and low accommodation (30 class instances).

5.2.1 Evaluation of the Machine Learning Pipeline

In order to evaluate influencing factors to this classification problem we varied the following parameters: (i) the learning algorithm, (ii) the set of used features and (iii) the number of attributes generated from the features. Furthermore, the set of used features is varied by including (i) each feature individually, (ii) each feature category individually, (iii) all feature categories or (iv) the top \(k\) best performing single features, with \(k\in\{2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20\}\).

We studied the Naive Bayes (NB), Linear Support Vector Machine (SVM) with cost parameter \(c\in\{2^{-5}, 2^{-3}, 2^{-1}, 2^{0}, 2^{1}, 2^{3}, 2^{5}, 2^{8}, 2^{10}\}\)\(^{16}\), J48 decision tree (J48) and k-Nearest Neighbor (KNN-k) with \(k\in\{1, 5, 10, 35\}\) algorithms. For each classifier/learning algorithm \(l\in L\), for each feature category and each feature \(f\in F\) we selected the \(g\) attributes having the highest Information Gain (IG) value to obtain our dataset. As values for \(g\) we used 50 different measure points. Half of them were equally distributed over the available number of attributes with an upper bound of 5000 attributes. The other half was defined by \(G = \{3, 5, 10, 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 7500, 10000\}\). Which attributes are finally used by the classifiers depends on the attribute selection algorithm (IG in our case). We measured the accuracy \((a)\) of the used algorithms \((l)\), the number of attributes \((g)\), the micro precision \((p)\) and micro recall \((r)\).

5.2.3 Results of the Classification

In Table 2 an overview of the best results about the about the performance of detecting transformations with low, medium and high accommodation by stratified 10-fold cross-validation for each feature category, for all feature categories combined, each single feature as well as the k top performing single features. The evaluation results show that a combination of four UICO features achieved an accuracy of 79.12% with the Naive Bayes algorithm for detecting low, medium and high accommodation tasks. In comparison, the probability of randomly guessing whether a transformation belongs to the low, medium and high accommodation class is 39.57% on our dataset. This means that by applying the ontology-based task detection approach significantly improves the accuracy. A detailed discussion of the features is given in the following:

**Feature Categories:** The best performing feature category was the content category with an accuracy of 77.03% and the NB algorithm \((g=175, p=0.86, r=0.77)\). The combination of all 50 features was closely behind with an accuracy of 74.12% with the same algorithm but required 1500 attributes \((p=0.84, r=0.72)\). The action category achieved the third rank in the category ranking with an accuracy of 69.07% with KNN-10 algorithm and with 25 attributes \((p=0.80, r=0.66)\). The difference between the best and the worst accuracy values was 23.68.

**Single Features:** The best performing feature was the content in focus, with an accuracy of 74.07% with the NB algorithm and with 75 attributes \((p=0.84, r=0.72)\). The content in focus feature belongs to the content category which was the best performing category with 77.03%. The second best performing feature with an accuracy of 70.44% was the control input keys feature with only one attribute and the SVM algorithm \((p=0.81, r=0.66)\).

\(^{16}\) The interval was chosen according to the libSVM guide at: [http://www.csie.ntu.edu.tw/~cjlin/libsvm/](http://www.csie.ntu.edu.tw/~cjlin/libsvm/)
attribute represents the number of times a control key, e.g. SHIFT, RETURN or INSERT, has been pressed by the user during the editing tasks. The distribution of values for the attribute showed that control keys were less used for low accommodation tasks than for high accommodation tasks. The values for the medium accommodation tasks are distributed in a balanced way. The user input feature ranked at third place among the best performing single features with an accuracy of 69.89% with the KNN-5 algorithm and with 10 attributes (p=0.80, r=0.66).

**Top K Features:** The best combination of the best performing single features was the top k=4 feature combinations with an accuracy of 79.12% with the NB algorithm and with 100 attributes (p=0.88, r=0.79). The third best combination was the top k=3 combination which only utilized the best three performing single features but also achieved a high accuracy value of 78.53% (l=NB, g=100, p=86, r=76). The top k=4 and the top k=3 performed almost as well in terms of accuracy with only 0.59% difference. All top k feature combination except the worst performing one outperformed all single features; all feature categories as well as the combination of all 50 features. This shows that not all features are helpful in the classification decision.
Table 2: Overview of the best results about the performance of detecting tasks with low, medium and high accommodation by stratified 10-fold cross-validation for each feature category, for all feature categories combined each single feature as well as the k top performing single features. The learning algorithm (l), the number of attributes (g), the micro precision (p), the micro recall (r), the ranking in the corresponding section (RS) and across sections (RG) is also given.

<table>
<thead>
<tr>
<th>Set</th>
<th>$R_S$</th>
<th>$f$</th>
<th>$l$</th>
<th>$g$</th>
<th>$a$</th>
<th>$p$</th>
<th>$r$</th>
<th>$R_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Content Cat.</td>
<td>NB</td>
<td>75</td>
<td>74.07</td>
<td>0.84</td>
<td>0.72</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>All Categories</td>
<td>NB</td>
<td>1500</td>
<td>74.07</td>
<td>0.84</td>
<td>0.72</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Action Cat.</td>
<td>KNN-10</td>
<td>25</td>
<td>70.07</td>
<td>0.80</td>
<td>0.65</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Resource Cat.</td>
<td>SVM-C = 2^5</td>
<td>3000</td>
<td>68.41</td>
<td>0.77</td>
<td>0.60</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Ontology Str. Cat.</td>
<td>SVM-C = 2^5</td>
<td>5</td>
<td>74.07</td>
<td>0.74</td>
<td>0.57</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Switching Seq. Cat.</td>
<td>KNN-5</td>
<td>10</td>
<td>74.07</td>
<td>0.77</td>
<td>0.60</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Application Cat.</td>
<td>NB</td>
<td>10</td>
<td>73.35</td>
<td>0.65</td>
<td>0.50</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>Single</th>
<th>Features</th>
<th>Top $k = 4$</th>
<th>Top $k = 6$</th>
<th>Top $k = 8$</th>
<th>Top $k = 10$</th>
<th>Top $k = 15$</th>
<th>Top $k = 20$</th>
<th>Top $k = 30$</th>
<th>Top $k = 50$</th>
<th>Top $k = 70$</th>
<th>Top $k = 90$</th>
<th>Top $k = 150$</th>
</tr>
</thead>
<tbody>
<tr>
<td>content in focus</td>
<td>NB</td>
<td>75</td>
<td>74.07</td>
<td>0.84</td>
<td>0.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control input keys</td>
<td>SVM-C = 2</td>
<td>1</td>
<td>73.44</td>
<td>0.81</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>user input</td>
<td>KNN-5</td>
<td>10</td>
<td>69.90</td>
<td>0.80</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>res. interact.</td>
<td>J48</td>
<td>175</td>
<td>69.78</td>
<td>0.80</td>
<td>0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>used res. interact.</td>
<td>J48</td>
<td>5</td>
<td>69.01</td>
<td>0.78</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>resource content</td>
<td>SVM-C = 2^{-2}</td>
<td>75</td>
<td>69.01</td>
<td>0.78</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UICO concept</td>
<td>J48</td>
<td>104</td>
<td>68.35</td>
<td>0.79</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>referenced res. interact.</td>
<td>SVM-C = 2^4</td>
<td>298</td>
<td>67.69</td>
<td>0.78</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E type switch seq. 2</td>
<td>J48</td>
<td>10</td>
<td>67.00</td>
<td>0.78</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>res. types interact.</td>
<td>J48</td>
<td>14</td>
<td>67.58</td>
<td>0.78</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>used res. metadata</td>
<td>SVM-C = 2^{10}</td>
<td>1000</td>
<td>66.98</td>
<td>0.76</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EB duration</td>
<td>J48</td>
<td>9</td>
<td>66.21</td>
<td>0.77</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>applications interact.</td>
<td>SVM-C = 2^{-1}</td>
<td>3</td>
<td>66.15</td>
<td>0.75</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>included res. interact.</td>
<td>SVM-C = 2^{-1}</td>
<td>75</td>
<td>66.04</td>
<td>0.75</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UICO objecttype rel.</td>
<td>J48</td>
<td>57</td>
<td>65.60</td>
<td>0.77</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nr. of E/EB</td>
<td>J48</td>
<td>10</td>
<td>65.40</td>
<td>0.76</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>navigation input keys</td>
<td>NB</td>
<td>1</td>
<td>64.80</td>
<td>0.74</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>included res. content</td>
<td>SVM-C = 2^5</td>
<td>50</td>
<td>64.78</td>
<td>0.77</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EB res. interact.</td>
<td>SVM-C = 2^{-1}</td>
<td>125</td>
<td>64.78</td>
<td>0.74</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>referenced res. content</td>
<td>SVM-C = 2^{-5}</td>
<td>50</td>
<td>64.73</td>
<td>0.77</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2.4 Comparison with other Task Detection Approaches

The evaluation results show that we achieved an accuracy of 79.12% and a precision of 0.86 for classifying transformations to low, medium and high accommodation levels. In comparison, by randomly guessing to which class a transformation belongs to someone would have a probability of 39.57% on our dataset. This is a good result in respect to the granularity of the task we try to detect: a transformation of a single wiki page. With granularity we mean here (i) the short duration of a transformation, (ii) user interactions on only one webpage and in one application, i.e., the browser.
Dataset: For us a "task" is a transformation of a single wiki page that can be categorized into a low, medium and high accommodation. The boundaries of the tasks were freely chosen by the participants of the study. Tasks used in other task detection experiments were higher level. Examples are "buying a book" task (Oliver et al., 2006), business tasks (Lokaiczyk et al., 2007) or tasks as in (Granitzer et al., 2008) comprised of email handling, paper writing, research, documentation or information collection.

Detection Performance: Existing task detection approaches focus on more higher level tasks and report similar performances: an accuracy of 76% with a precision of 0.49 (Oliver et al., 2006), 85% with a precision of (Lokaiczyk et al., 2007), an accuracy of 74.51% with a precision of 0.91 (5 classes) and an accuracy of 76.42% with a precision of 0.90 (4 classes) (Granitzer et al., 2008) and a precision of 0.8 (96 and 81 classes) (Shen et al., 2006).

Features: The most popular features identified for having a high discriminative power among tasks are the window title feature (Oliver et al. 2006; Shen et al., 2006; Lokaiczyk et al., 2007; Granitzer et al., 2008), the file path/web page url (Shen et al., 2006;), and the content in focus feature (Granitzer et al., 2008). The task detection results on our dataset show that the window title and the file path/web page url features do not work well. This is because of the fine granularity of the tasks we were trying to detect. A whole transformation of a study participant only involved the Wiki simulation environment web page, i.e., a single flash application, and the browser.

Attributes: In terms of attributes used for training the machine learning algorithms an interval of 200-300 attributes is suggested to be sufficient by (Shen et al., 2006; Granitzer et al., 2008). Our results agree that only a small ratio of attributes is required to successfully detect a specific accommodation task.

Classifiers: In the task detection experiments reported in (Lokaiczyk et al., 2007) the SVM learning algorithm was mentioned as the one with the highest achieved accuracy. In (Granitzer et al., 2008) the good performance of the SVM learning algorithm was confirmed and the high accuracy achieved by the KNN learner highlighted. On our datasets the SVM and the KNN learner showed also good results but was beaten by the Naive Bayes algorithm.

6. Conclusion

In the research presented here we have developed a tool that is able to analyze how people interact with different artefacts in order to create a text. This context-analysis tool provided a method to measure processes of accommodation and assimilation in an automatic way. Our research provided also a more detailed differentiation of knowledge maturing processes. This is an important aspect in the larger context of social media and knowledge management, since the study provided some insights into the processes that lead to accommodation of knowledge. Consequently, our results may help to improve the understanding on informal learning and knowledge maturing.

The first goal of our study was to measure the quality of text as an indicator for knowledge building within communities and to identify 'accommodation patterns' (as opposed to 'assimilation patterns'). Based on the ratings of two experts we could distinguish among three classes of low, medium and high accommodation. The ontology-based task detection approach – that was applied in order to answer this identification question – finally yielded an identification rate of 79.12%. This was the first step for developing methods for the automatic detection of accommodation processes as indicators for knowledge building and knowledge maturing. The results of our study led to the following conclusions:

- It is possible to detect the accommodation level of a wiki text transformation with only a small ratio of features. This allows an efficient detection of text maturing and knowledge building on the basis of user behaviour. The features achieving the highest
accuracy values, however, were related to the content of the wiki article and the users’ interactions with this content. Since content-related features are often domain dependent, it might be the case that they are not well generalizable with regard to other domains.

- The results led to a better understanding of accommodation processes. Expert ratings according assimilation and accommodation correlate on a high level. This means that the two constructs are not independent from each other. From our perspective the process of accommodation is central for text maturing and knowledge building. It seems to be a proper strategy for further studies to focus on accommodation as main concept.

The second research question aimed at answering the question how it can be ensured that knowledge workers write good texts and contribute to the development of knowledge. Here, the results of the user ratings are important. One could assume that users are not able to monitor their own work very well and can hardly decide, whether their edits will lead to a maturing of the text. Thus, it could be a good strategy to use the classification results (and the reading scores) as a feedback for users that guide them to become successful writers.

In a next step we would like to use the automatic way of analyzing accommodation processes for further studies that may identify possible support mechanisms to improve and/or encourage accommodation (towards supporting knowledge maturing). We also plan to implement our tool as a MediaWiki extension and to apply the task-detection method to real-world settings with real communities working on shared digital artefacts. In addition to the online detection of knowledge-building processes a post-hoc analysis of revision histories of collaboratively written text would be of great interest. This is would allow to analyze existing text corpora in the Internet, e.g. of the Wikipedia.

Altogether, we hope that our previous findings and our considerations concerning potential future studies will stimulate other researchers to initiate corresponding research on their own that might shed some light on the understanding of processes of knowledge building and knowledge maturing.

References


Deliverable D3.1: Task management model.


Proceedings of the 1st Workshop on Context, Information and Ontologies (CIAO 2009), workshop as part of the ESWC 2009, Heraklion, Greece. New York, NY: ACM.


Schmidt, A. (2005). Knowledge maturing and the continuity of context as a unifying concept for knowledge management and e-learning. Proceedings of the Fifth International Conference on Knowledge Management (I-KNOW 05), Graz, Austria.


Annex 5: Phone, Email and Video Interactions with Characters in an Epidemiology Game: towards Authenticity

Muriel Ney\(^1\), Celso Gonçalves\(^1\), Nicolas Blacheff\(^1\), Claudine Schwartz\(^2\), Jean-Luc Bosson\(^2\)

\(^1\)Laboratory of Informatics of Grenoble (LIG), CNRS and University Joseph Fourier 961 rue de la Houille Blanche, 38402 Grenoble Cedex, France
\(^2\)Techniques for biomedical engineering and complexity management (TIMC), University Joseph Fourier, Domaine de la Merci, 38710 La Tronche, France

Abstract. A key concern in game-based learning is the level of authenticity that the game requires in order to have an accurate match of what the learners can expect in the real world with what they need to learn. In this paper, we show how four challenges to the designer of authentic games have been addressed in a game for an undergraduate course in a medical school. We focus in particular on the system of interaction with different characters of the game, namely, the patients and a number of professionals. Students use their personal phone and email application, as well as various web sites. First, we analyze the authenticity of the game through four attributes, authenticity of the character, of the content of the feedback, of the mode and channel of communication and of the constraints. Second, the perceived authenticity (by students) is analyzed. The later is threefold and defined by an external authenticity (perceived likeness with a real life reference), an internal authenticity (perceived internal coherence of the proposed situations) and a didactical authenticity (perceived relevance with respect to learning goals).

Keywords: authenticity, immersion, simulation, role-play, epidemiology, communication

1 Authentic and Immersive Games

Computer simulations afford the possibility of showing a phenomenon when it is impossible or impractical to confront learners with such a phenomenon in the physical world. However, a simulation alone may not be sufficient. An important component of an educational approach including a simulation is to allow an authentic learning experience, the virtual equivalent of a real-world experience. This is precisely what a learning game is about, simulating real-world possibilities to act, and consequently helping players to become more confident of his or her ability to recognize and handle similar situations, should it happen afterward in real life. However, authenticity does not mean a perfect reproduction of reality. What we mean by authentic are the main characteristics of the type of situations at stake, that is, those characteristics that require learners to mobilize the knowledge targeted (the learning goals) in order to be successful in the game.

One mean to create authenticity in learning games is immersion, making learners feel like a certain situation is real although they know it is not. According to Brown [1] the main indicator of immersion is the degree of involvement of the players. Game designers know that players may be engaged in different ways, mobilizing them differently, whether they are challenged on their ability to act rapidly, to find the best strategy, to communicate with different people, etc. There have been numerous definitions of immersion in games in previous studies (e.g. tactical, strategy or sensory immersion) [2], [3]. In this paper, we focus on immersion in an authentic situation, and more specifically, a situation that involves many moments of interactions with people. We call "interactional immersion" an immersion that relies mostly on interactions with other players or with characters of the game. Examples of such games that focus on the social aspects of the situation are Second Life or the Sims.
In this paper, we address the following four challenges posed to designers of authentic and immersive learning games:

- Find a compromise between the requirement of the real life reference and the learning context.
- Allow the learner to appropriate the meaning of the situation (what is the point in learning terms).
- Blur the line between the virtual and the real experiences.
- Create engaging interactions for players with realistic characters.

We show how these challenges have been taken up in a game for an undergraduate course in a medical school.

2. Design Methodology

The Laboratorium of Epidemiology (LoE) is collaboratively designed and then used by three parties: researchers, teachers and students [4], [5]. On one hand, medical students enrolled in a course of biostatistics and learn to develop critical thinking on statistics and epidemiology studies. On the other hand, researchers assess the conditions of learning with such a simulation. LoE is thus both an educational project integrated in a medical school, and a research project, a laboratorium, allowing repeated data collection campaigns that are not singular events in students and teachers lives. This integration is an attempt to reduce data collection bias and produce well-documented corpora.

LoE uses learning with live experiences of social interactions, learning by doing, and problem-based learning and as such follows a socio-constructivist approach. In LoE, students conduct their own epidemiology survey and present evidences in an article. By doing so, they acquire skills in critical reading of medical articles. Indeed, with LoE students learn about issues in applying methodologies in epidemiology surveys, in particular, when defining a main objective, designing a protocol, collecting data (with issues like sample size, data quality, ethical consideration, etc), using statistical tools and analyzing data, making a decision based on statistical results, and presenting results in a scientific form (article and talk) [5]. LoE also addresses patient-centred skills (history taking), communication skills (with different professionals), and teamwork skills [6].

LoE is based on a computer simulation and a role-play. Students play the role of public health physicians and experience an otherwise inaccessible professional situation: an epidemiological survey at the hospital. For the sake of authenticity, we used real data to simulate patients at different hospitals. This allowed us to tell students that everything was real. We used a national database containing about 10,000 patients suspected of being affected by one of the tromboembolic diseases. Then the mission given to students was first to evaluate the risk of these diseases at a hospital (look for signs and symptoms and give evidence based on statistics) and second to propose a decision tool (which patients are at great risk and should be tested first). The mission was sufficiently broad that students could further define their own objectives for their epidemiology survey.

To design an immersive simulation that offers students such a professional experience, we first performed a task tree analysis based on a series of workshops with an epidemiologist [4]. The epidemiologist was asked to focus on facts and actions rather than on a conceptual view of his practices. Combined with the learning goals of the course, this work resulted in a task tree presented in figure 1. It starts from the problem to be solved while tasks on the lower leaves correspond to the actions to be performed to solve that problem. Between the two, there are steps and sub-steps. On figure 1, seven main steps of the task tree are indicated. Only one of them is further detailed into sub-steps and actions.
The task tree approach allows visualizing the line of tasks that a learner with a given level of expertise may follow [7]. It was a basis to define a compromise between the requirement of the professional reference and the learning context. Finally, this work also helped unravel the important interactions. We will now see how the different professionals and the patients were represented as characters of the game. The system of interactions between students and characters of the game is summarized in Table 1. We indicated how the character is represented in the game (e.g. by a text giving it a name, gender and position) and whether the level of representation is low (few information) or high (many information including a picture, a voice, etc). Then the mode and channel of communication with characters are given. Some of them will be detailed below. Finally, the learning goals of the interactions are indicated.

**Table 2. System of interactions between students and characters of the game.**

<table>
<thead>
<tr>
<th>Character of the game</th>
<th>Representation level of the character</th>
<th>Mode of representation</th>
<th>Mode of communication</th>
<th>Channel of communication</th>
<th>Learning Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head of hospital department</td>
<td>Low</td>
<td>Text</td>
<td>Oral and written</td>
<td>Telephone</td>
<td>Synthesize a protocol verbally</td>
</tr>
<tr>
<td>Expert of the ethical committee</td>
<td>Low</td>
<td>Web site</td>
<td>Written</td>
<td>e-Mail</td>
<td>Write and communicate a protocol</td>
</tr>
<tr>
<td>Patient</td>
<td>High</td>
<td>Video and text</td>
<td>Visual, oral and written</td>
<td>Online structured interview</td>
<td>Translate patient saying into medical data</td>
</tr>
<tr>
<td>Technician of the hospital information department</td>
<td>High</td>
<td>Video and Web site</td>
<td>Written</td>
<td>Web and e-mail</td>
<td>Make a data entry mask and fill it with data</td>
</tr>
<tr>
<td>Congress reviewer</td>
<td>Low</td>
<td>Web site</td>
<td>Written and oral</td>
<td>Web</td>
<td>Write and present statistical results to a scientific community</td>
</tr>
<tr>
<td>Public health</td>
<td>Low</td>
<td>Web site</td>
<td>Written</td>
<td>e-Mail</td>
<td>Propose a recommendation to</td>
</tr>
</tbody>
</table>
The game computer environment is used by teams of three or four students. Teams may gather in groups at various times, and at least for the final simulated congress. Several groups may play the game in parallel. Therefore, the team is the main unit in the game. Although each student has a personal identifier for the web site, the environment will adapt to the team (never to a student individually). For instance, the patient “recognizes” the team when it comes back to a room, or the report on the article submitted to the congress is visible by the whole team. The only thing that teams of the same group share are regular meetings in class with a tutor. In our implementation, there are eight four-hour meetings and there are about 180 students distributed into six groups (with about 6 teams each).

As we will see in the following sections, patients are played by actors while professionals are played by tutors. In class, the tutor’s role is to help students on the epidemiology methodologies. Tutors also play the different characters of table 1, but not necessarily for their own group of students. This way, they discover the characters’ feedback content with their students and do not have to pretend they do not know it in advance. Another reason is that it allows attributing the characters to tutors according to their competences.

The environment is a pure Web platform that can be reached from any Internet-enabled computer. Server-side code is object-oriented PHP built upon open-source Symfony framework, and client-side code is rich HTML including Adobe© Flash and Ajax components. A second environment is used by researchers who can visualize computer trails of the users (not shown). A third environment is devoted to tutors, and mainly allow them to follow students’ productions and to interact with them (table 1). This environment should not only facilitate the task they have to do, but also help them to “be” the character of the game that is supposed to interact with the students. In the following, we specify the immersive and authentic components of most interactions of table 1.

3 Description of the system of interaction

We describe successively some of the email, phone, video and web interactions of table 1.

3.1 Interaction with Experts by Email

Students interact with two groups of experts (table 1). One is the Ethical Research Committee that, in real life, controls all the protocols before they can be applied on patients. Another is the Public Health Commission that commissions epidemiology surveys. We use similar communication mean for both, for the same reasons, and we describe in details only the first one.

Students have to send their protocol to the Ethical Research Committee for approval, before they can apply it at the hospital with patients. The learning goals of this interaction are to learn how to address to professionals in a formal way and to learn how to present a protocol in a written form (table 1). In real life, people address protocols to this committee by regular mail or, more and more often, by email. We choose the latter (figure 2). The benefit of this system is to give life to a character (member of this committee) that will receive students’ productions. This way, we hope that students will address problems in terms of the underlying concepts and skills, rather than in terms of what they think that the teacher will expect them to do [8], which happens when the teacher is the receiver.
Students use their personal e-mail application. They send their protocol as a pdf file that should be introduced by a short message. This is compulsory since only approved protocols can be implemented at the hospital.

The immersive interface that we use for tutors is a standard Webmail interface. It is immersive in the sense that it is separated from their usual mailbox (they might probably answer in their name by mistake, otherwise) and it is organized with a dedicated signature and e-mail address. Furthermore, it is accessible to tutors only, from the game web platform.

### 3.2 Interaction with Hospital by Phone

In order to be able to interview patients, students have to ask for the authorization of the head of each medical department of the hospital(s) they wish to visit (table 1). Students learn to be convincing, quick and thus prepared. It also put them in a rather intimidating situation since they are talking to a physician with a heading position. In real life, students do not have to meet him/her in person (this person is very busy and it is difficult to make an appointment) but rather they have to give a call.

We designed the following interaction (figure 3) where students have to use their personal phone. They make a call and, having listened to a brief message, they have to formulate their demand in the way they think useful to describe it to the head of the department. Later, a tutor listens to their message and writes an SMS. Students thus receive an answer on their personal phone by SMS that gives them an agreement or an argued refusal. An agreement allows the team to access to the patients' rooms. When a tutor sends a SMS with an agreement, it opens automatically the rooms to the given team of students. This phone call is thus a compulsory stage to continue the epidemiological survey. A team can make several calls, either to have access to several departments, or to repeat their demand after a refusal.
“specify your objectives and constraints for patients and staff”. We use IP convergence to integrate VOIP technology into the web application: Phone2Web via IMAP for tutors to check phone messages and Web2SMS for tutors to send SMS to students.

3.3 Interaction with Patients Based on Video

The interaction with patients is one that has been extensively simulated, obviously due to its central place in medical practices. Simulated patients are used to overcome the disadvantages of involving real patients that have limited availability and that provides low variability of cases [9]. Reviewing the literature, we built the following list of main strategies that have been used (table 2).

Table 3. Bibliographical search on patient interaction systems. We specify the representation model of the character (here the patient), the mean of communication with the patient, the type of activity the student can perform, and the learning goals.

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Character representation model</th>
<th>Mean of communication</th>
<th>Type of activity</th>
<th>Learning Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>[10]</td>
<td>Text</td>
<td>None</td>
<td>Information research</td>
<td>Diagnostic</td>
</tr>
<tr>
<td>[11]</td>
<td>Hypertext Hypermedia</td>
<td>Website</td>
<td>Information research</td>
<td>Diagnostic</td>
</tr>
<tr>
<td>PULSE - U.S. Office of Naval Research</td>
<td>Avatar</td>
<td>Clicking</td>
<td>Information research Communication</td>
<td>Diagnostic Interview</td>
</tr>
<tr>
<td>[12], [13]</td>
<td>Video</td>
<td>Website</td>
<td>Information research Communication</td>
<td>Diagnostic Interview</td>
</tr>
<tr>
<td>iStan - Medical Education Technologie s, Inc.</td>
<td>Model</td>
<td>Oral</td>
<td>Gestures Communication</td>
<td>Diagnostic Interview</td>
</tr>
<tr>
<td>[14]</td>
<td>Student</td>
<td>Oral</td>
<td>Gestures Communication</td>
<td>Diagnostic Interview</td>
</tr>
</tbody>
</table>

The patient can be physically present in the room (last three rows in table 2). First, the patient can be an actor. Simulated or standardized patients are individuals (actors) who are selected and trained to portray a patient, including accurate and consistent responses to questions [6]. Second, the patient can be a student. A student plays the role of a patient in front of peers, which is a cheaper alternative to simulated patients and sometimes leads to similar outcomes [14]. Third, the patient can be a model that talks and bleeds (like the commercial model iStan). The talking is done through the mouth of the model by a teacher. The teacher is behind the scene watching the student dealing with the model, thanks to a camera (the so-called Wizard of Oz method).

In all the other cases, the patient is represented on screen or paper. First, the patient can be described textually (first row in table 2). Many medical schools across the world including the one where our experiments are conducted have adopted Problem-Based Learning [6]. Patient cases are most often presented on paper. It has the drawback that the patient is only a source of information, and not a person with words, feelings, in a social and cultural context [10]. Second, the patient can be described through a multimedia online document. An authoring system that is currently in use in several healthcare disciplines worldwide is Web-SP [11]. The patient is represented by a number of textual information and images (his/her photo, X-rays, etc) accessible by checking boxes. Third, the patient can be a video-game character. The most accomplished example today is PULSE, which includes a 3D
virtual hospital where students guide their avatar and interact with lifelike patients (doing more than interviews since physical examinations and interventions can be performed). Finally, the patient can be a person video recorded. Clips of either true patients or actors are used, but in most cases only for demonstration (physical examinations skills, doctor-patient relationships, etc), rather than for problem solving [6]. In a recent example, interviews in psychiatry are performed using a branching system of videos allowing students to personalize their interview [12], [13].

LoE includes video clips of actresses and actors portraying patients (up to 20 patients in each of the six hospitals). In our case, a benefit of creating virtual interviews was to script the responses according to learning goals. The scripted responses includes signs and symptoms, communicated verbally or not (e.g. breathing heavily can be a sign), related to the disease at stake or not (some information are irrelevant). The system allows the students to interact with the simulated patient by selecting one of five preformed questions (figure 4): present illness, present lifestyle, past history, family/social history, and medications. For each question, there is an appropriate video clip associated with it that provides the interviewer with the patient’s response. The patient is speaking directly to the camera and thus to the students. This response lasts generally around one minute, and can only be seen once by the team of students. On-demand video has been implemented using Adobe® Flash technology.

![Fig. 8. Screenshot of one of the rooms with patient and questionnaire.](image)

When entering the room, a video is immediately starting: either the patient is absent (video of a nurse), not available (video of a physician doing a physical examination) or the patient is there (video of the patient either saying hello or telling that he/she saw the interviewer already). A probabilistic algorithm regulates the presence/absence of patients.

Students must visit at least fifteen patients before they can ask the Information Department to complete their database with hundreds of cases, which will allow doing statistics. With this system, students learn to listen to patients with great attention, to identify accurately symptoms and signs within the patient narrative, to be prepared (know what to look for in advance), to control data quality (e.g. by confronting notes from two interviewers of the same patient), and to manage time (patients are not always available). Moreover, since students have little opportunity to conduct interviews with real patients that early in their studies, this system could help them feel more comfortable when conducting their first patient interviews.
3.4 Interaction with Institutions through their Web Site

In order to be able to participate to the simulated medical congress at the end of the game, students need to submit their article. Students learn to write a scientific article, how it is structured, and how to choose and present their evidences (statistics may come in different forms). In real life, applicants to a congress interact with the scientific committee either by email or via the web site of the congress. We chose the latter.

We designed the following interaction (figure 5) where students have to upload their article on the congress web site. Later they will get a report on their article, visible on the web site. In most cases, they will be asked to revise their article and send it again. They also have access to the program of the congress. Articles are divided into two groups by the organizers of the congress and best articles are given a longer time of presentation.

Fig. 9. Diagram of interaction between students and the scientific committee of the medical congress.

The tutors who play the role of the referees connect to the congress website and fill a pre-structured form that will make a report on the article. They have a list of points they should pay attention to, to help them in this task. They play the role of referee for students they do not know.

4 Analyze Authenticity in a Game

Serious games are playing an important role for training people on real world situations. However, an entirely realistic simulation is neither practical nor desirable. On one hand, game designers add realism by adding more and more elements, and thus adding complexity to the situation. On the other hand, when students spend too much time to get familiar with too many details they can skip the main learning goals. In the context of learning, a key concern is thus the level of authenticity that the game requires in order to have an accurate match of what the learners can expect in the real world with what they need to learn. Failure to achieve the right level of authenticity runs the risk that the learners may adopt a different strategy in the virtual world than would be desired for learning.

Designed with the desire to make the right compromise between the professional requirements and the learning goals, LoE includes several characteristics to facilitate a move to the right degree of authenticity. Another very good example from which LoE was inspired is the WallCology project [15], which simulates a virtual ecosystem of insects within the unseen space of classroom walls. Interestingly, Moher and co-authors [15] pointed limits in this compromise, not only technical limits, but what we may call epistemic limits. For example, they decided to use imaginary creatures instead of authentic ones in order not to frighten young children and to avoid stereotyping the living conditions of some learners. The limit to authenticity was therefore defined by the learning context.

Authenticity is both a function of the game as well as of the perceiver, and we shall now look at each one separately in the context of LoE.
4.1 Authenticity of the Game

Considering the authenticity of our system of interaction, we identified four attributes (see table 3). First, the characters students are interacting with may be more or less authentic. This is decomposed into two issues. First, the authenticity of the character refers to its personification: what information is available and in what form (text, photos etc). Second, the content of the feedback the character sends to students may be more or less authentic, depending on how adapted to student’s actions this feedback is. Another issue is the channel of communication (by email, phone, etc) and the mode (textual, verbal, visual). Finally, the constraints of the interaction may be considered. For instance, in the interaction with patients, it is important that a patient does not repeat at will his/her answers, which is a constraint that was reproduced in our system. This particular constraint is important to the learning context. Indeed, when a patient does not repeat oneself, students need to prepare their interview and also they need to look for ways to control their data, two learning goals of the course.

<table>
<thead>
<tr>
<th></th>
<th>Authenticity of character</th>
<th>Authenticity of feedback content</th>
<th>Authenticity of communication mode and channel</th>
<th>Authenticity of constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head of hospital department</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Expert of the ethical committee</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Patient</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
</tbody>
</table>

In Table 3, the authenticity of the character is said low when this character is hardly represented. It has only a name (head of hospital department) or not even a name but only a competence (e.g. for the group of experts of the Ethical Research Committee). Regarding the patient, it is only the mode and channel of communication that is low. Indeed, students do not really hold a dialogue with a patient as they would in real life. This table shows that several attributes may be tested when considering the impact of authenticity on learning outcomes.

4.2 Perceived Authenticity

Authenticity is also a function of the perceiver, in our case the student. We call it perceived authenticity. One aspect is students’ judge likelihood that an event happening during the game could happen in the real world. This is also referred to as perceived realism [16]. This relates to the perceived fidelity of the game, that is, whether students feel that the simulated interactions mimic a reality.

To study perceived authenticity, we analyzed (1) the trails (verbal, written) of the interactions of seven four-student teams with different characters, and (2) the interviews of two students at three moments during the game. The seven teams (28 students) are second year medicine students at the Grenoble Medical School in a compulsory course of biostatistics paying the game LoE over a period of 12 weeks. We successively analyzed the phone, email and video interactions studying for each whether students behaved as they would with real people [5].

The first data set consisted of phone messages left by students on the hospital answering machine asking the head of the hospital department authorization to interview patients. We collected and analyzed 15 messages. A majority of the messages (10/15) addressed the person in charge of the hospital department in a formal way. This indicates a relation of
hierarchy that is adequate in the context of this mission. This is to compare to the information available to students about the receiver of their message. In this case, all they know about the characters are name, gender and position of the receiver of their call. This proved to be enough. On the contrary, for most teams (9/15) patients are only sources of information rather than people. These characters (the patients) have no representation in the environment (students had no information about these characters at this stage of the game although they will later see them on videos).

The second data set consisted of e-mail messages. The students sent 12 messages to the experts of the Ethics Research Committee, introducing an attached research protocol. We analyzed whether students played the game, that is, addressed these messages to experts as they should, using the following six criteria: putting a message subject, giving a personalized name to the added file, greeting formally, presence of a message addressed to the receiver, closing formally, presence of a signature of the writer. Most messages (11/13) showed at least four of these criteria. Therefore, students used appropriate communication features to address to this character by email.

The third data set consisted of audio-taped intra-team interactions. Students’ verbal interactions were recorded as they were using the web application to perform their patient interviews (i.e. view videos). We analyzed a four-hour session for one team. The verbal interactions were synchronized with the computer trails. Perceived authenticity is thus analyzed through students’ verbal reactions to videos. Students had immediate reactions to what patients say, usually an interjection (e.g. “a tumble!”, “he’s deaf!”). After each video, they sometimes made inappropriate comments (e.g. “I hate these old people unable to answer a single question”). However, they might do the same after interviewing a real patient. We observed several indications of personification of the patients: Students showed feelings for the patient condition (“honestly, I am worried about him”, “Did you see his blood pressure? We’ll have to take care of him.”), students talked about a person rather than about a system (“the patient is not there, the nurse said that he is gone for some test”), students said hello to the patient after the hello-video. Moreover, students reacted to probabilistic events (presence/absence of patients): “look if the patient is back from the test” or “be careful, maybe the patient in room 4 will go away soon”. Therefore, we collected several indications that students perceived the situation as authentic in the sense that they played the game and behaved as in a real situation. The authenticity of the character, of the content of its feedback and of the constraints (see table 3) all contributed to their perceived authenticity.

Next, we looked for indications on perceived authenticity in three interviews with two students. After the phone interaction, these students mentioned that what makes the interaction more “interesting” and more “realistic” is: using their personal mobile phone, talking to an answering machine that limits the duration of the call (need to be prepared), leaving a phone message to a human being and not a machine. They said that they did not know exactly who was going to listen to their message and they had an image of a “person with a white blouse” (a physician). In a later interview, one of the students talked about the video interaction with patients. She pointed out that it was not necessary to interview patients, although it would be necessary in real world. Indeed, in her group it was not clear why students had to interview patients before they can ask for new data and this looked like an optional step. This pointed out a type of authenticity that we did not envision before, namely, an internal coherence of the mission proposed. Indeed inconsistencies within the mission can also disrupt students’ engagement. Furthermore, the other student underlined that interviewing patients was useful for his training. In other words, he perceived this activity as relevant from the learning point of view. To summarize, a situation may be perceived as authentic from three points of view: it can be perceived as realistic, coherent and/or relevant. In our case, one student perceived the interviews with patients as realistic, relevant but not coherent. Another one perceived it as realistic, coherent but did not mention any thing about its relevance.
5 Conclusion

At the beginning of this paper, four challenges posed to the designer of authentic and immersive games were proposed. We are now going to show how these were tackled in the design of the game LoE and more specifically of the system of interaction with characters of the game.

First, to find a compromise between the requirement of the real life reference and learning context we combined a task tree analysis with learning goals. The task tree allows showing all the necessary actions to be taken towards the goal. Combined with the learning goal analysis, one can choose those actions that will be more important for learning and those that could be done in place of the students or skipped.

Second, to allow the learner’s appropriation of the meaning of the situation, we introduced a form of monitoring on the learning process as part of the interaction system. This allowed them to revise their work and reflect on what has been learnt or missed. Furthermore, we included two moments of institutionalization. The latter is a term used in the Theory of Didactical Situations [17] to denote the process of giving a status to the activities done and the knowledge acquired by the students. First, after students had been working on their data analysis and when they started writing their paper, a teacher gave a two-hour interactive session on statistics. Second, at the very end and right after the simulated medical congress, the teacher debriefed the congress with the students and then de-contextualizes the knowledge that was embedded in a specific problem during the game.

Third, to blur the line between the virtual and the real experiences, we allowed students to use their personal phone and email application. This design aimed at immersing students into the game. The ubiquity of mobile phones is providing a fading of the line between the game space and real world experience [18]. An example is the game Majestic, one of the first so-called alternate reality games, in which players were receiving phone calls on their personal phones. In our case, phone calls and emails are both received and sent by learners, like no other game does, as far as we know.

Fourth, to create engaging players’ interactions with realistic characters we addressed the authenticity of different attributes of the interaction: personification of the characters, content of the feedback from these characters, mode and channel of communication and constraints. Different degree of authenticity may be attained for these different attributes. Then we studied authenticity on the learner side, which is partly referred to as perceived realism [16]. In our case, we speak of the perceived authenticity of a mission proposed to students, which includes several situations of interaction. We found indications of three types of perceived authenticity, that is, the extent to which each situation seems (1) similar to a real one (external authenticity of the mission), (2) coherent within the game (internal authenticity of the mission) and (3) relevant from a learning perspective (didactical authenticity of the mission). This is partly in line with research on film narratives [19] where internal realism and external realism are pointed out and correspond to (1) and (2). We suggest that the three types of authenticity should be taken into consideration when designing authentic learning games. However, we only caught a glimpse on how these three are related to authenticity attributes of the game with the game LoE, and further work is needed.

References

5. Gonçalves, C., Ney, M., Balacheff, N., Bosson J-L.: Student’s Problem Appropriation in an Epidemiology Game, paper presented at ECGBL 2009, European Conference on Game-Based Learning, Graz, Austria (2009)


Annex 6: Technology for classroom orchestration

Pierre Dillenbourg and Patrick Jermann
Swiss Federal Institute of Technology Lausanne (EPFL)


Abstract. We use different criteria to judge teaching methods and learning environments as researchers and teachers. As researchers, we tend to rely on learning gains measured in controlled conditions. As teacher, the skilled management of classroom constraints results in the impression that a specific design "works well". We describe fourteen design factors related to the metaphors of classroom orchestration and education ecosystems and illustrate their embodiment in three learning environments. These design factors provide a teacher-centric, integrated view of educational technologies in the classroom. We expand this list of factors to include the main constraints that designers should consider to address the difficult methodological issue of generalizing research results about the effectiveness of methods and designs.

Keywords: educational technologies, classroom orchestration, ecosystem, design based research

1. Are metaphors useful for educational design?

On a regular basis, new metaphors emerge in the field of learning technologies. Are they 'old wine in a new bottle' or do they convey a novel idea? Is 'inquiry-based' learning more than 'learning from simulations'? Is 'educational data mining' different from 'student modelling'? Is 'orchestration' just a new buzzword? It probably is, but nonetheless this chapter argues that the idea of 'orchestration' conveys a new flavour to classroom technologies. However, we cannot evaluate the effectiveness of a design metaphor in terms of learning outcomes. There is no way to establish that "all learning environments whose designers have been inspired by a metaphor X are more effective than those who have not". Design metaphors trigger analogies rather than quantifiable specifications. Hence, a first step to assess their usefulness is to disentangle the elements that can be found in these metaphors and relate them to concrete features of learning environments. This is the goal of this analytical contribution.

In the first section, we explain why the relationship between design and learning should be analyzed in terms of classroom life and not only in terms of learning outcomes. Then, we decompose this relationship by extracting a set of 'design factors', first from the metaphor of 'orchestration' (Section 3) and then from the metaphor of 'educational ecosystem' (Section 4). These factors relate pedagogical and technological design choices with classroom life. In the next 3 sections, we illustrate these factors with three learning environments we have developed and tested in real contexts. Our conclusions revisit orchestration as the management of constraints systems.

This contribution is restricted to the domain of formal education in co-present settings. It does not deny the interest of research on informal learning as well as on distance education but stresses the social responsibility of our research community to contribute to schooling.
2. "It works well (in my class)"

As most scholars in learning sciences, we are not only researchers but also teachers in our own university. Surprisingly, these two roles sometimes co-exist almost independently from each other.

As teachers, we sometimes have the pleasant feeling that the method we are using is "working well". What does it mean? It probably corresponds to situations during which students are engaged in the proposed activities: they attend the course, they pay attention to our talk, they argue with each other at the semantic level, they ask relevant questions, they make smart suggestions,..., in short, they are 'with us'. This "it works well" is just a personal feeling, a subjective opinion, not an objective measure even if it may rely on cognitive cues: the students' answers indicate that they understand, they come with original solutions or new ideas, they deliver satisfactory assignments, they do well at the exam,..... Moreover, "it works well" means that the method is compliant with our constraints: the students’ workload is reasonable, our teaching workload is acceptable, there is no main bug in the technology, the teaching material can be reused next year, the students are satisfied (as shown by the course evaluation), as well as the director and the students' parents. As teachers, we have to care about many of these constraints that we do not really conceptualize as researchers: homework, parents, discipline, room size, friendship, security, etc. In summary, there are many conditions for a method to "work well" and we, as teachers, are delighted when these conditions are met

As researchers, we did not pay that much attention to things that "work well". A first obvious critique is that the teacher's opinion that it "works well" may be unfounded. It may be that students are actually bored while the teacher perceives them as engaged. It may be that students are highly engaged but do not learn much through the activity. For instance, the LOGO environment "worked well" for many years but did not reach its promises in terms of learning outcomes (Pea, 1983; Tetenbaum & Mulkeen, 1984). Similarly, many colleagues stick to lecturing because, from an economical viewpoint, it "works well", that is, it satisfies several of the constraints we will analyze in this chapter. Educational research reached its maturity when it replaced stories about methods that supposedly "work well" by empirical evidence of learning gains (pre-post). The fact that teachers feel their method "works well" is probably a necessary condition for success, at least for sustainability of the method, but certainly not a sufficient condition. The second critique against this idea is that "it works well" refers to a broad beam of interweaved factors that can hardly be disentangled in a formal experiment. Rigorous experiments may not address more than a few independent variables while "it works well" refers to many parameters. While design-based research leads to pedagogical methods that "work well", they still face founded criticisms with regard to the generalisability of their findings.

Can we reconcile the viewpoint of teachers who need methods that "work well" and the viewpoint of researchers for whom "it works well" can hardly be viewed as a scientific statement? Wouldn't we be extremely happy as researchers to know that the teachers involved in our empirical studies continue to use -on a voluntary basis- the methods we tested in their class? Wouldn't we be upset as teachers to be asked to test methods that don't work well? The gap between the teachers' and researchers' viewpoint is related to the difference of scale in their preoccupation. On the one hand, the researcher has to narrow down his focus to a specific activity or process (e.g. peer tutoring), hopefully isolated from any other source of influence which would be considered as a bias. It is interesting–yet dramatic- to note that the 'teacher effect' is often viewed as an experimental bias (F. Fischer, personal communication)? On the other hand, the teacher has to consider the much broader scope of what's happening inside the class, and to some extent also outside the class. So, there is a huge difference of scope in the concerns of each side. The models we analyze in this chapter are compatible with these two levels since they articulate in a single picture individual cognitive processes, social interactions and the class life. We hence refer to them as 'integrative models' versus other models that have a more local scope (e.g. the zone of proximal development, the repair theory, ...),
Integrative models are not learning theories. They do not predict how learning may result from specific activities. Instead, their strength is to be rather 'agnostic' with regard to learning theories. Theories provide us with a particular lens to grasp the complexity of learning in classrooms, but they are somehow exclusive. Isn't it legitimate to combine them, for instance to reason on the metacognitive skills can be internalized through social interactions (Blaye, 1988)? Isn't it legitimate to design activities that combine self-explanation and socio-cognitive conflict? Teachers don't have to choose between Bloom, Vygostky or Piaget. Researchers don't have to choose either. Students have a single brain; they don't switch between brains when moving from individual to social activities. When two peers perform a classification task, how could we dissociate individual induction from social argumentation?

The interest of the models we discuss hereafter is to pay attention to the multiple levels that enter into play when explaining that a method "works well". It is also their weakness: by addressing many aspects, they remain quite abstract. We hence attempt to extract more specific factors that relate models with design choices. We analyze to general models, orchestration and ecosystems, and then instantiate them with homemade specific models ("SWISH" and "Der Erfahrung")

3. The "Orchestration" model

Many scholars have used "orchestration" to refer to the design and real-time management of multiple classroom activities, various learning processes and numerous teaching actions (Brophy & Good, 1986; Tomlinson, 1999; DiGiano, & Patton, 2002; Fischer et al., 2005; Gravier et al., 2006; Dillenbourg & Fischer, 2007; von Inqvald, 2009). Intuitively, this metaphor is appealing. Music writers and teachers both have to harmonize multiple "voices". They both need a fine-grained control of time. They both translate a global message (emotions, knowledge) into a sequence of atomic actions (notes, interactions). Rothstein-Fisch and Trumbull (2008) explain that they use the word 'orchestration' instead of 'management' because it "denotes bringing about harmony" (p. 101). Moon (2001) uses 'orchestration' to refer to "the process of managing a whole learning group in such a way as to maintain progress towards the learning outcomes and improvement of practice for all" (p. 120)

Unfortunately, the classroom "orchestration" seems to have a different meaning than its musical counterpart. In music, orchestration refers to writing the score that an orchestra will play. It does not refer to the activity of the conductor when the orchestra is playing. The metaphor is applied in a more orthodox way to computer science where "service orchestration" (Peltz, 2003) refers to the definition of some workflow or architecture that connects different systems (namely web services). Applied to the educational context, the proper meaning of "orchestrating" would correspond to instructional design and not to the real time management of classroom activities.

Nevertheless, metaphors don't have to be perfect to be inspiring. The key difference between music orchestration and classroom orchestrations is that, when orchestrating a classroom, the score has often to be modified on the fly. Could we imagine that orchestra players read their score on a computer display (rather than on paper) and that this score changes dynamically, depending for instance on the audience's emotional state? Good teachers perform "reflection in action" (Schön, 1983), i.e. they are able to change the score as they are playing it. A good learning environment should let good teachers modify the score as often as necessary. The words 'dynamic orchestration' could convey the combination of design/planning and real time adaptation. Let us stress that 'dynamic' does not mean free improvisation, it is neither plan-free or goal-free, but refers a certain degree of freedom around the instructional plan.

Is that stretching the model of orchestration too far? Metaphors are not correct or incorrect; the question is whether they are useful or not. In HCI, designers have used metaphors such as the desktop, the cockpit, the window, .... Instructional designers also developed metaphors such as 'frames' (Merrill, Li & Jones, 1992), 'anchored' instruction (Bransford et al, 1990), ... A metaphor is 'useful' if it helps the designer/teacher to take decisions, namely because it provides a global structure to articulate multiple local decisions. A metaphor creates an educational flavour that local design principles do not convey. We analyze now
why 'orchestration' does convey flavours that a concept such as 'classroom management' would not convey. We try to operationalize these flavours by extracting a certain number of design factors.

3.1. Teacher-centrism

Orchestra players religiously follow the movements of the conductor. This quasi-mimesis is an exaggeration of the message this model should convey to educationalists but it stresses the importance of the teacher's role. In the last decades, many colleagues have redefined the role of teachers as facilitators (Carey, 1994) along the common place slogan "from a sage on the stage to a guide on the side". Orchestration has a different flavour: teachers are not on the side, they are the conductors, they are driving the whole activity. They are managing in real time the activities that occur at multiple planes. They share their passion for the content. Their body language conveys their interpretation (speed, intensity, …) to the musicians.

Factor 1. Leadership. Teachers act as the drivers of the scenario and lead the collective (i.e. class-wide) activities.

Our community has somehow confused constructivism with "teacherless" (Dillenbourg, 2008). Let us state clearly that promoting the role of teachers as orchestrators does not imply that they have to lecture intensively or that they have to make a show. Orchestration would be compatible with "teacher-centric constructivism": it is the students who have to learn through their activity but teachers have the leadership of the whole scenario: lectures may be integrated into constructivist perspective for several purposes.

*Debriefing* lectures "work very well": these are not simple feedback sessions but lectures that address new course contents through the data produced by the students themselves (experiment results, projects, assignments,…). The drawback is that these lectures have to be prepared after receiving the students' contribution which require last minute preparation or even some improvisation. There is a "time for telling" (Schwartz & Bransford, 1998): lectures "work well" in terms of learning when students have previously acquired some experience of what the lecturer is going to say; when they have the meanings but not the words.

Lectures enable teachers to "qualify" the knowledge for instance to explain to students why an equation is beautiful, why a theory is somehow obsolete, why these results are surprising,… This 'personal touch' or meta-knowledge can hardly be made explicit in instructional material without over-simplification. It's the beauty of human presence to be able to convey these subtle cues.

Providing teachers with a strong leadership implies that they have the power to drive the system and not to be prisoners of an instructional plan. Acknowledging the role of teachers does not simply mean to let them tune some options or parameters but to truly empower them with respect to technology. This implies for instance that teachers are allowed to bypass decisions taken by the system and to flexibly rearrange their own scenario. Of course, we will see that flexibility has some limits described in section 8.1

Factor 2. Flexibility. Teachers have the possibility to change the learning scenario on the fly, as far as it makes sense.

As leaders of a class, teachers have the responsibility of what students do in this class. They don't have the same degree of control that the orchestra conductors have over their musicians, but they still have to be 'in control' of their class. A method 'does not work' if students become distracted, if they talk to each other when the teacher speaks to them, if they read their email while they should interact with a simulation. Some methods may reduce the authority of teachers because some students start to do silly things while waiting for late groups, because teachers lose face when encountering technical bugs, etc.

Factor 3. Control. Teachers maintain in the classroom the level of interest and concentration necessary for the on-going activities.
A method that "works well" facilitates teachers' task of maintaining interest and concentration among their audience. The class behaviour has not been not a main concern in learning sciences research but it is a major issue for teachers (especially new teachers), directors, parents and even students.

3.2. Cross-plane integration

In computer-supported collaborative learning (CSCL), the idea of "script" emerged as a method that shapes the way collaboration will occur within a team (Dillenbourg, 2002). More precisely, scripts aim at triggering specific types of interactions that are known to generate learning gains such as providing explanation, solving conflicts or mutually regulating each other. For instance, to increase conflicts within a team, some scripts, referred to as micro-scripts, will prompt peers to provide counter-evidence against the claims made by the other (Weinberger, Ertl, Fischer & Mandl, 2005). Another approach is to detect people who have opposite opinions and ask them to perform a task together, or to give peers different documents to read so that they end up with different opinions (Jermann & Dillenbourg, 1999). These so-called "macro-scripts" (Dillenbourg & Hong, 2008) also include individual activities (read, summarize, write,...) and class-wide activities (introductory lectures, vote, debriefing,...) in addition to small group activities. We also refer to them as 'integrative' scripts (Dillenbourg & Jermann, 2007) in the same way we talk here about 'integrative' models.

Factor 4. Integration refers to the combination, within a consistent scenario of individual, small group and class-wide activities, as well as activities beyond the class.

Anecdotally, we have used a notation (Figure 1) that is very similar to music scores. Time is represented from left to right. Lines refer to what Vygostky designed as planes: the intrapsychological plane (level 1 in fig 1: individual cognition), the interpsychological plane (level 2: small group interactions) and the social plane (level 3: where culture is located). It is clear that mental activities occur most of the time at the three levels in parallel. The "notes" placed on the score do not represent cognition but concrete activities.

These levels constitute an arbitrary segmentation of a social scale continuum. However, looking at practices, most classroom activities occur along these 3 levels: solo, class or in between. We added three levels that can be encountered: (level 4) activities with other classes in the school; (level 5) activities with local community (parents, correspondents, local fieldtrips,...) and (level 6) activities with the anonymous world via internet (e.g. on-line newspapers, polls,...). This segmentation is very pragmatic: level 5 adds logistics and security concerns to perform out of school the activities that are performed in the school at level 4.
3.3. Sequentiality

A piece of music is more than a set of notes, the sequence makes it all. A method that "works well" is not just a random sequence of activities; there is something in the sequence that turns discrete activities into a consistent whole, there is a scenario.

Orchestration has a flavour of linearity. Some verses or refrain can be moved and repeated but nonetheless, the idea of orchestration refers to some sequentiality. In terms of learning, the order of activities partly determines the cognitive process. Whether a concept definition is introduced before or after the examples will switch between an inductive and a deductive approach. Letting students evaluate their solution before or after providing a feedback changes it all. The order matters. For many years, non-linearity has been perceived as an educational plus, e.g. when moving from linear programmed instruction to branched programmed instruction or when moving from texts to hypertexts. It was a new feature offered by technologies that certainly facilitates adaptation to individual needs. But, despite being the enemy of individualisation, linearity has pedagogical advantages. It may lighten teachers’ cognitive load of orchestration by streamlining activities. It simplifies issues of data consistency at the technical level. It also creates a classroom atmosphere by knowing that all students experience the same thing more or less at the same time. Of course, there is design trade-off with factor 2, since flexibility should enable teachers to escape from linearity, but the flavour of linearity is nonetheless strong.

Factor 5. Linearity. The method is a simple sequence of activities that almost all students will perform at almost the same period. It is easy to explain to the students.

The notion of scenario implies continuity across activities: the groups remain the same, the students keep their roles or rotate in a predictive way, the activities concern the same objects or issues (as in problem-based learning or project-based learning). Technically, this continuity may be implemented as a workflow: the output of an activity is used as input for subsequent activities.

Factor 6. Continuity. The successive learning activities are articulated around shared data structures (objects, groups, assignments,...) that circulate via a workflow.

A method that "works well" is not emotionally flat. It includes phases of tension, where energy is accumulated, and phases of relaxation where accumulated energy is exploited. It is difficult to report this from a scientific stance, but the best methods are salty.

Factor 7. Drama. The emotional state of students varies across activities, with highest moments that trigger engagement for the rest of the scenario.

3.4. Time Management

One of the main constraints that teachers have to cope with is time. Not only is teaching time limited, but moreover it is segmented into time slices. The flexibility of this segmentation depends on the institutional context. A method will only be described as "working well" if most of the following timing issues can be orchestrated.

A method that "works well" must be reasonable in terms of total time necessary. Time for learning has always been a critique against constructivist approaches and an easy argument in favour of lecturing since the rate of content delivered per minute is high. The time necessary must be proportional to the importance of the domain to be taught within the year curriculum. This importance is measured in terms of credits in higher education or class-contact hours in lower school levels. This means that the time available depends upon the curriculum relevance of the topic (see section 4.2).

Factor 8. Relevance. The total time that a method requires to teach X should be proportional to the importance of X as specified in the curriculum

A method that "works well" must be flexible in terms of time segmentation. Should the next
activity require a full period or can it be segmented into two periods? Should the teacher stop 15 minutes before the break because it is not worth starting a new activity for 15 minutes? Can students save work-in-progress and reuse it later on?

Factor 2 (again). **Flexibility.** Teachers have the possibility to adapt the next pedagogical activity to fit the time slice available.

The dynamics of a method that "works well", the feeling of drama, require the right intervention at the right time. This is a trivial statement but isn’t true that the best joke can be destroyed by a bad timing? This means that some activities must be set up when the students are hot, when the energy is present. This new instance of factor 7 (drama) will be illustrated in section 5.

3.5. **Physicality**

Finally, conducting an orchestra is rather physical. It is not a virtual action in a virtual space. The conductor is physically engaged. The spatial layout of the room and the location of each musician are very important. The orchestration of classroom activities encompasses the spatial organization of tables, chairs and tools. This layout must facilitate the transition between the different forms of grouping in the scenario and depicted in figure 1. It must enable students to move when they have to move, for instance because they are switching roles within a group or switching between groups as in some Jigsaw scripts (Aronson, 1978). The teacher must be able to pass between groups and to be present across the space. The students' and teachers' location must enable them to see what they are supposed to see.

Factor 9. **Physicality.** Compared to previous models that stressed virtual learning spaces, orchestration refers to the concrete layout of the physical space in the classroom and to the physical movements of the different actors.

The physical layout of an orchestra enables the teacher to perceive the activities of all musicians at any time. On this point, the model does not match very well classroom practices: when the teachers moves between tables (what a conductor does not do), he does only perceive a subset of the students. What the conductor and the teacher have in common here is the need to keep permanently an awareness of the activity performed by each actor. Even a simple lecture requires maintaining an approximate model of the level of attention across the different sectors of the lecture room. A good teacher monitors the state of the class and reacts as soon as something goes wrong. New attention technologies may help some teachers to maintain a global representation of the state of his class, not by doing deep student modelling but by analyzing social signals in real time: what student are looking at, with whom they interact, how they sit on their chair, etc.

Factor 10. **Awareness.** An orchestration technology helps teachers to be aware of the activity state of his students, at a behavioural level.

One interesting problem with the orchestration metaphor is that the conductor has an awareness of all musicians but does not see the audience in his back. Good teachers are known to have ‘an eye in their back’, but the model only works well if we consider the students as the orchestra players and not as the concert audience.

**4. The "Ecosystem" model**

It is very tempting to describe education as an ecosystem, as any system with multiple circular causal components (Copeland, 1979, Stoll & Fink, 1996; Dillenbourg, 2008). This metaphor has been applied outside the field of environmental sciences for instance to describe a city or a society as an ecosystem. Authors such as Resnick (2002) or Brown and Adler (2008) use the word ecosystem to refer to the global learning space outside the school, namely the digital sphere. As for 'orchestration', these metaphors are used in a shallow way. Again, is it just a buzzword or is it interesting for a learning technology designer to consider a classroom as an ecosystem? Does it help to understand why some method
"works well"? Actually, the notion of ecosystem is quite different from the 'orchestration' since an ecosystem does not include a central orchestrator (unless some divinity is considered as orchestrator). Nonetheless, some authors in natural sciences talk about "orchestrating an ecosystem" such as a Sequoia Forest (Brigg, 2001). The word 'orchestration' is then used in a broad sense as a synonym of 'management'. The ecosystem model has also been applied to business cases. For instance, the study of various communities using the SAP environment is described as the orchestration of an ecosystem (Iansiti & Lakhani, 2009).

4.1. Species

In an ecosystem, one cannot start feeding the rabbits without considering the impact on foxes. As any ecosystem, a classroom shelters the interactions between several animal species: students, teachers, teaching assistants, special support staff, directors, inspectors, parents, ... This has trivial implications which are nonetheless worth mentioning.

Methods that 'work well' should not only be designed for learners but also, as mentioned in section 3, for teachers, but also for parents, directors, ... For instance, the learning technologies that our community is developing rarely pay attention to the homework and, through it, to the role of parents. Technologies could improve the integration of homework in the flow of learning and enable parents to follow what their children have done. Isn't it strange that, while many scholars on learning technologies are parents, we behave as researchers as if we were not? Some ethical arguments may explain why researchers neglect these possibilities: homework increases the effects of social disparities (Ferguson, Ludwig & Rich; 2001); not all parents have Internet access; parents might interfere too much with the teachers’ job, ... However, it is a fact that parents and homework are part of the ecosystem.

Factor 4 (again). Integration of homework and parents’ interaction in the educational workflow.

Each species includes a large variety of animals whose features (size, fitness...) follow more or less a Gaussian distribution of animals. One may exceptionally find a rabbit that is 1 meter long but not many. A method that "works well" cannot be generalized if it is designed for exceptional teachers. The researcher maybe lucky to find one exceptional teacher, for instance one who cope easily with chaos in the classroom, but this will restrict generalisability of the results. The very fact that teachers accept to participate into our experiments is an unavoidable bias in the representativity of these volunteers. The trivial implication is that pedagogical methods must be adaptable to differences between teachers. Our point goes further: designing for the ecosystem means that most teachers should be able to apply our method, not only a few of them.

Factor 11. Design for all. The method can be conducted by most skilled teachers, not only by exceptional teachers

4.2. Selection

A method that 'works well' in an education ecosystem has to pay respect to the rules that govern life and death within the ecosystem.

Teachers are not free to teach what they want; they have some degree of freedom in primary school, almost no freedom in secondary schools and a bit more at university level. Our community has been quite creative in designing activities that address skills that are not in the curriculum or only in the 'meta' section with transversal skills. An extra-curriculum investment from teachers and students is acceptable for a short duration (the time of an experiment), but such an environment will only be used over long term if is justified by the importance of its learning objectives within the curriculum.

Factor 12. Curriculum relevance. How important are these learning objectives within the curriculum of these students?
Whether we like or not, the educational ecosystem is shaped by large scale evaluations such as PISA or TIMSS. Our research community tends to neglect or even disregard these assessments. Instead, we should contribute to the improvement of the quality of these measures. In the meanwhile, these large assessments and more generally all exams are part of the educational ecosystem. A method that "works well" is a method that does not put teachers, directors or ministers in an undefendable position when the results of evaluations are published. At the university level, we don't have large scale assessments, but the same argument applies: a method will not "work well", if the students don't perceive this method as really useful for passing the course exam.

Factor 13. **Assessment relevance.** Is the method compatible with the different assessments that students will have to pass?

In theory, curriculum relevance should guarantee the assessment relevance, but in practice there are gaps: the specific way of measuring skills and knowledge may differ from the way they are taught, the national or international assessments may differ from local curricula, etc.

**4.3. Legacy**

Given the complexity of causal links within an ecosystem, external interventions have to be minimalist. The same applies to the classroom ecosystem. The learning environments we design do not land in an empty world. Every class has its legacy of methods, tools and resources that have been used so far. A new method cannot wash out the past. For instance, a method that "works well" should be compatible with the main book used in the course otherwise teachers would be in the very difficult position of explaining why they do not use the book that has been bought. Another example is that every student or teacher has a legacy of habits with respect to computational tools. If the learning environment proposes a specific email tool or a specific chat tool, it will suffer from a strong competition with the email and chats that students are using, i.e. the ones that already include all their contacts, their histories, their favourite goodies, ... The obvious implication is to design a minimalist intervention that respects this legacy, namely to design a learning environment that is limited to the functionalities strictly required by the pedagogical method. This means design should resist to the temptation to provide a fully integrated learning environment (with yet another agenda, yet another forum, etc.).

Factor 14. **Minimalism.** The functionalities offered by the learning environment are only those specific to the learning scenario and that are not provided by the tools (books, software,..) already in use by the students

However, minimalism has drawbacks: an integrated environment enables interconnection between tools, for instance a link between chat utterances and graphical objects on the display (as in ConcertChat, Mühlpfordt and Wessner, 2005) or a link between a forum and a spreadsheet (as in http://www.sense.us). Despite emerging standards and multiple APIs, it remains a technical challenge to provide this kind of interactions between the environment we design and multiple existing tools (e.g. with the 5 most popular chat systems). In other words, there is a trade-off between minimalist intervention and integration, but the ecosystem model is rather on the first end

**4.4. Sustainability**

A classroom has limited energy resources as any ecosystem. The energetic efficiency of a pedagogical method is a key factor to claim that it "works well". The energy that students will invest is of course limited; it is actually measured in terms of time (i.e. credits). For instance, it often happens that a very exciting project concentrates all the students' energy on one specific course, leading them to fail other courses. But, most importantly, we refer here to the teachers' energy. A method based on teachers' heroic investment over a few months is not a method that will "work well" for several years. Teachers and their team will do it well the first year, even better the second year and then the investment will inevitably decrease over the following years.
Factor 15. **Sustainability.** The energy required to run the method can be maintained over several years.

We should not forget that teaching is a long term repetitive activity: researchers conduct an experiment once or twice but teachers have to repeat it for many years. A method that "works well" must have realistic energy expectations. It may for instance be that groups of 2 are better than groups of 4 for the targeted interaction but they require twice more energy for grading the assignments. Design is always a trade-off.

## 5. The "SWISH" model and the ManyScripts environment

This section and the two next ones illustrate how the factors mentioned actually shape the design of technologies used in the classroom. We report the local design models used for building these environments, their implementation and how they relate to the 15 factors.

### 5.1. The 'SWISH' model

A macro-script is a pedagogical method that aims to trigger specific interactions during teamwork (Dillenbourg & Jermann, 2007). Here are two scripts we want to relate to our 15 factors. Their score-like representation is reported in figure 2.

- **ArgueGraph** (Dillenbourg, 2002) aims to trigger argumentation by forming pairs of students with conflicting opinions. It is run in 5 stages: (1) individuals answer a questionnaire; (2) the system computes their opinion by compiling their answers and locates them on a map that is projected in the classroom and discussed collectively; (3) groups are formed with individuals to maximize their distance on the map; they have to answer the same questionnaire as in the first phase; (4) the teacher gives a debriefing lecture based on the answers provided by individuals and pairs; (5) students have to write a summary of the arguments. Typical ArgueGraph sessions last 3-4 hours depending on the number of questions. Experiments showed that the ArgueGraph conflict mechanism leads students to express more elaborated arguments (Jermann & Dillenbourg, 1999).

- **The ConceptGrid script** (Dillenbourg & Jermann, 2007) aims to trigger explanations. (1) each team has a set of papers to read and a set of definitions to produce; students distribute their work within teams (2) students individually read the papers they have been assigned to (3) students individually enter the definition of the concepts that they have been assigned (4) teams have to build a grid of concepts in such a way that they are able to explain the relationship between two neighbours on the grid (by entering a short text in the system); (5) the teacher gives a debriefing lecture based on the grids build by students. A typical ConceptGrid sessions lasts 2-3 weeks, including the time for their readings. Experiments showed that ConceptGrid forces students to elaborate upon each other's explanations (Dillenbourg & Hong, 2008).

These examples of macro-script rely on the same design principle that makes the team interactions more difficult than if the teacher let them freely collaborate. Collaborative learning is often described as the side effect of the effort engaged to build a shared understanding of the task at hand. This effort is represented as $\Delta$ on figure 3: there is a natural divergence in teams at the beginning ($\Delta_1$) and the script actually increases this divergence ($\Delta_2$) in order to increase the intensity of interactions required to finally minimize...
this divergence ($\Delta_3$). In ArgueGraph, the effort necessary to reach agreement is higher since we pair students who disagree. In ConceptGrid, the effort necessary to build a consistent grid is increased by the fact that no student has enough knowledge to do it; they have to explain it to each other. Both scripts increase the collaborative effort by 'splitting' the team in terms of respectively opinions and knowledge. The nature of this split determines the type of interactions that student will need to engage in to complete the task. Hence, the design principle was termed 'split where interactions should happen' (SWISH). Other examples are provided in Dillenbourg & Hong (2008).

Figure 3. Macro-scripts reduce team convergence to increase the effort to be engaged.

5.2. The ManyScripts environment

The two examples of scripts are available in web-based platform. Teachers can select a script (ArgueGraph, ConceptGrid,...), edit the contents, set up the groups and tune some parameters. Figure 4 presents a snapshot from the ArgueGraph script in the ManyScripts platform: the system synthesizes all the justifications that individuals and pairs have associated to their answers. Teachers use this display during his debriefing lecture (Phase 4).

5.3. Design factors

In both scripts, the teacher has a salient role (Factor 1: Leadership) especially during the debriefing phase. These are not genuine lectures but 'organic' lectures based on what student have produced. The drawback is that these lectures cannot be prepared in advance. We usually had one coffee break between the duo answers and the debriefing phase in ArgueGraph. This semi-improvisation is against factor 10 (DesignforAll) since many teachers don't like to improvise, even if in this case the lecture is facilitated by the environment (figure 4). In ConceptGrid, we usually asked students to complete their grid one or two days before the debriefing lecture, which requires less improvisation but nonetheless heavy work to analyse all grids the night before the course. This stressful preparation work reduces the sustainability (factor 15).

Both scripts illustrate factor 4 (integration) since they integrate individual, small group and class wide activities in a meaningful way. They are linear (factor 3) and based on a workflow that provides continuity (factor 6): for instance, data from individual answers (ArgueGraph, phase 1) are automatically processed to form conflicting pairs (phase 2), individual and pair answers (phase 3) are collected for debriefing (Phase 4).
Drama (factor 7) is embedded in the SWISH model: this model is about increasing and relaxing tensions. However, smaller design elements make the drama higher in ArgueGraph than in ConceptGrid. The conflict phase triggers a degree of engagement that goes beyond the 'didactic contract' (Brousseau, 1998): in many cases, we had to tell students "Stop arguing now with your friend even if you have not convinced him; this is just a didactic game; we have to continue". In ArgueGraph, we also noticed the following phenomenon. When being asked to choose between answers A and B, many students expressed their frustration not to be able to answer some mix of A and B. This frustration pushed them to stand up literally during the debriefing session (phase 4) to defend themselves. They needed to explain that they answered B but actually wanted to answer something slightly different. Their frustration raised their level of participation. Once, we modified the interface of the environment and let the student express subtle choices (e.g. to answer 'in some cases' instead of 'yes' or 'no') and this killed the drama. Finally, we found that the energy generated by this design was very fragile. It depends upon the timing (factor 7) of the script. Once, we did the debriefing activity one week after the argumentation phase instead of right after and all energy had disappeared from this debriefing session.

The time flexibility (factor 2) was crucial in ArgueGraph. We had bad experiences with an early version of the environment in which all students were expected to provide their individual answers before to move to the group formation. The whole class was stuck if one student left the room or lost the Internet connexion. Now, ManyScripts offers the possibility to move on to phase 2 even if all answers from phase 1 have not been provided. In terms of timing, teachers actually make sure that the delay between the first student who completes the questionnaire and the last one is not too high. If many students have to wait for other, it quickly generates all sorts of undesirable behaviour (factor 3: Control).

Since ConceptGrid has a duration of several weeks, we had to implement flexibility (factor 2) in terms of group formation. Once a group of students is formed, what happens if a student drops out the class? If the teacher decides to make groups of 4, what happens if the number of students in the class is not a multiple of 4? ManyScripts implements two functionalities for coping with these accidents (Dillenbourg & Tchounikine, 2007). If a student is missing in a group, the teacher may turn on the 'SPY' feature which enables the group in which role-X is missing to borrow the definitions produced by the students who play role-X in any other group of the class. This makes sure that all students have the same workload (factors 12 and 14). Conversely, if the group has for instance 5 students for 4 roles, the teacher may activate the 'JOKER' function: a 'joker' student may provide a definition that belongs to any other role. These functionalities are not perfect solutions but ways to carry on the script despite an imperfect situation. A method that "works well" should survive to these common
accidents in everyday classroom life.

Finally, with respect to factor 13 (assessment relevance), ConceptGrid connects the script to the course exam: the environment includes a button to print all the work produced by a team so they could use it during the exam.

6. The "Erfahrraum" model and the TinkerLamp environment

6.1. The "Erfahrraum" model

The initial vocational education is structured as a dual system in Switzerland and other German-speaking countries: students spend about 4 days per week in a company and one day per week at school. These apprentices are between 16 and 20 years old and represent 69% of the teenagers in the educational system. Vocational education constitutes a specific ecosystem, with its culture, its laws, its actors (e.g. the companies, the corporate associations), quite different from general education. Our main research hypothesis was that technologies may play a specific role in a dual system as a bridge between what apprentices do at their workplace and what they do at school. The model of "Erfahrraum" combines two ideas into a German name that reflects the socio-cultural context in which this model emerged. "Erfahrung" means 'experience'. Learning from experience is of course not a new idea: it is rooted in Dewey's work. Experiential learning is defined by Keeton and Tate (1978) as learning "in which the learner is directly in touch with the realities being studied" (p. 2). A dual training system combines this direct experience with some distanciation, more abstract activities in the classroom. Learning technologies are envisioned as ways to capture the apprentices' experience in order to exploit it during these more abstract activities in the classroom. "Raum" means room, as we insist on the physical orchestration of the room (factor 9).

The model has been implemented by different technologies in three different contexts. In the first context (Gavota et al., 2008), we capture workplace experience of dental assistants by asking them to write down their experience in a wiki-like environment. The school activities, peer commenting and text revision, exploit the diversity of experiences across a class of 20 apprentices: some work in small cabinets with a single dental surgeon, some in large cabinets with several surgeons, some with old fashioned technologies, other with high-tech cabinets, etc. In the second project, (J.-L Gurtner, University of Fribourg) captures workplace experience life in the context of car mechanics and pastry making. They call apprentices on their workplace on the phone and/or ask them to take pictures. Later on, this raw material feeds classroom activities. Finally, we instantiated the Erfahrraum approach in the context of logistics (Jermann, Zufferey and Dillenbourg, 2008). Our observations of logistics apprentices in their warehouse revealed a gap between what apprentices are asked to do in their warehouse and what they are supposed to learn. While the official curriculum specifies that they should acquire logistics skills, the apprentices mainly follow their boss’ instructions (move this over there). They are rarely involved in decision making such as flow optimization, warehouse layout or storage management. We refer to this as the 'abstraction gap', i.e. the difference in the degree of abstraction between the tasks they experience and the tasks they should master. In school, the apprentices encounter more abstract logistics problems but the drawback is that they do not connect these tasks to their warehouse experience.

6.2. The TinkerLamp environment

The TinkerLamp is an augmented reality system designed to run tabletop tangible simulations. The simulation which we developed in close collaboration with teachers from a professional school allows logistics apprentices to build a warehouse model by placing small-scale shelves on the table. Besides shelves, users can place tangibles which represent architectural constraints (e.g. pillars to sustain the roof of the warehouse, loading docks, offices and technical service rooms). The building elements for the model are scaled to allow the construction of a realistic warehouse (32 by 24 meters in reality).
The physical small-case model is augmented through a video projector placed above the table (figure 5). All objects (shelves, pillars, cardboard) are tagged with fiducial markers (similar to a 2 dimensional bar code) which enable a camera to track their position on the table (Fiala, 2005) and enable the system to project graphical representations (augmentations) on top and around the objects. The physical layout of the warehouse is used as input to configure a simulation. Projected forklifts represent the movement of goods in and out of the warehouse.

Figure 5: TinkerLamp: a tabletop tangible simulation in the domain of logistics. Left: four TinkerLamps used concurrently in a classroom. Right: a closeup of a small-scale warehouse designed by apprentices. Information is projected on the floor and on top of the shelves.

The simulation is controlled by a paper-based interface called TinkerSheets (figure 6). Small tokens can be placed on a paper form which is recognized by the system and allows users to set parameters like the type of warehouse management, the number and type of forklifts, or the type of augmentation which is displayed. Master sheets allow setting the main parameters relevant for a particular logistic concept and companion sheets either allow setting supplementary parameters or to visualize general simulation output (summarized numbers, graphs, etc).

Figure 6: TinkerSheets. Left: a master sheet from the official curriculum is used to parameterize the warehouse simulation. Right: a companion sheet is used by an apprentice to visualize different types of surfaces in the warehouse.

The TinkerLamp is used on a regular basis in 4 different classes in two Swiss schools. Teachers are enthusiastic and describe it as environment that "works well". More formal evaluation have shown the positive effects of the tangibles on learning (Schneider, Jermann, Zufferey, and Dillenbourg, submitted) but the usability of TinkerSheets has not been empirically proven.
6.3. Design Factors

The TinkerSheets offer interesting opportunities with respect to our 15 factors. First, they make curriculum relevance (factor 12) very tangible. The official curriculum – jointly designed by public authorities and the relevant corporate association has the form of a large binder. Teachers may simply take one sheet from this binder and place it under the TinkerLamp in order to set up the activities described in the sheet. In addition, teachers may annotate these sheets with personal comments (e.g. which simulation parameters work better for this activity) that will be very useful for the next year reuse of this activity (Factor 15: sustainability).

These sheets can also be used for designing homework (Factor 4: Integration), for instance by printing a warehouse performance sheet at the end of a warehouse design activity. Teachers may then ask apprentices to perform some analysis at home and to come back for the next week with new simulation parameters. Their homework could typically be assessed by putting the homework sheet under the TinkerLamp and see how the warehouse performs.

The continuity across activities (factor 5) becomes tangible: what connects successive activities is not an invisible workflow, but a concrete sheet of paper that is passed from hand to hand between the different phases of the activity. This “tangible workflow” has advantages: it is simple and concrete, publicly visible, documents can be annotated, shown to echo other…Actually, the virtual workflow still exists since any sheet is associated to a fiducial marker that connects it to a data structure, but this invisible data set has a concrete clone in the physical world.

The physical orchestration (Factor 9) of the activities was initially not trivial. Controlling the simulation required teachers to use finger-driven menus which were not easy to use. This difficulty did not empower teachers as drivers (factor 1). We replaced these menus by TinkerSheets that have a much higher usability. Once they are laid out around the display area, they constitute some kind of cockpit in which is relatively easy to see all available options. Initially, we used a very large version of the TinkerLamp (the TinkerTable, 1.5m X 2m?) which occupied a large space for only 5 students. What would then happen with the students who are not working (factor 3: control)? Working with a subset of the class is acceptable for one experiment but not sustainable on the long term (factor 15). One teacher came with an innovative idea: while one team was working on the table, he projected their work on a screen placed elsewhere and on which he could discuss with the rest of the class what one group was saying. The new tool was then integrated in the ‘legacy’ environment of a whiteboard and a beamer (factor 14). Later on, we produced a smaller simulation environment (40 X 50 cm) in such a way that 4 groups can work simultaneously in the same class under four different TinkerLamps.

7. The Shelve and Lantern environments

7.1. The Recitation Section Model

Recitation sections are managed in our university in a way similar to many other places: the students receive a set of exercises to be carried out with the help of teaching assistants (TAs), mostly PhD students. The sessions gather between 20 and 60 students in a room with 2 or 3 TAs’. The students work in groups of 2 to 4; some prefer to stay alone. They raise their hand to attract the TA’s attention who joins the group and provides help for a few minutes. Occasionally, if several teams face the same difficulty, the TA gives a collective explanation on the blackboard.

7.2. The Lantern Environment

We observed several recitation sessions and noted several regulation problems: students spend a lot of time chasing the TA instead of working on their assignments. TAs do not necessarily come in a 'fair' way (first asker is not first helped), some students ask help from the very beginning of the exercises without even trying seriously to solve it, most students do
not complete all exercises while exam items have a level of difficulty similar to the last exercises of the session (factor 12), etc (Alavi, Dillenbourg & Kaplan, 2009). To help the TA to orchestrate the session, we developed two awareness tools. The ‘Shelve’ is a central display where teams indicate with a remote control device which exercise they do and when they need help. The ‘Lantern’ is a small device that teams put on their table and which they use to provide the same information either by turning it (to indicate their move through the exercises) or by pushing on its top (to ask TA’s help). Both the Lantern and the Shelve display the same information: which team is working on which exercise (each exercise has a different colour), how long they have been working on that exercise (the height of the colour bar), do they need help (it blinks) and for how long (it blinks faster).

Figure 7. The Shelve environment.

Figure 8. The Lantern (left) and its use during orchestration.

7.3. Design Factors

The Lantern helps tackling the problems of time management (factor 7) and control (factor 2). The main difficulty for the orchestrator is to manage multiple teams of different sizes working at different speeds and requiring different types of help at different times. The experiments (Alavi, Dillenbourg & Kaplan, 2009) showed that students using the Lantern or the Shelve spend less time chasing the TA: once they push their device to ask for help they can concentrate on their exercises. Control (factor 3) is improved because teams do not simply wait doing nothing.

Overall, this environment provides TAs with awareness (factor 10). As we stressed, it does not carry any knowledge diagnosis or any smart computation. If simply provides an overview of who is doing what and who is looking for help. It also illustrates the physicality of orchestration. The only difference between the Lantern and the Shelve is the spatial layout. Besides the fact that Shelve is centralized and Lantern is geographically distributed, both devices display the same information, in the same way. The Lanterns connects spatially the information displayed (exercise being done, time in exercise, waiting time) with the team being concerned by this information. Yet, this single difference led to different interaction patterns in the class. The central display induced more comparison between groups, independently of their location, while the distributed version led to the emergence of clusters, i.e. sets of 2-3 small groups, located close to each other, and which interact with each other (e.g. a group A helps group B because A sees that B has been waiting for help at an exercise that A has completed (Alavi, Dillenbourg & Kaplan, 2009).
8. Conclusions
8.1. Implications for learning technologies

We presented 3 examples of technologies developed in our lab but many other colleagues have developed environments that are close to the orchestration model and inspired us as the NIMIS classroom (Hoppe et al., 2000), personal response systems, classroom-scenarios based on handhelds (DiGiano, & Patton, 2002; Zurita & Nussbaum, 2004) or even on multiple mice (Infante et al., 2009). Our goal in presenting our examples was to show that the orchestration model, although being somewhat abstract and metaphorical, can be turned into concrete implementation choices. We hereafter list several technological choices that are more or less directly connected to the orchestration model.

1. Orchestration technologies are different from distance education since they are designed for classroom life. They may include on-line activities but the most salient part of the scenario occurs in face-to-face.

2. A part of e-learning technologies are document-centric, while orchestration technologies mostly support activities (e.g. the simulation for TinkerLamp) and overall diverse forms of interactions (as in ArgueGraph or Lantern).

3. Orchestration technologies should have a high usability for the teacher. Of course, the usability also concerns the students but orchestration implies that teachers may easily interact with the technology despite being overloaded by other tasks (managing groups, lecturing, etc.)

4. These technologies make orchestration quite physical, i.e. they have to cope with the spatial organisation of the classroom and other spaces as well as with the location and movements of students and teachers. The development of HCI towards location-aware services is hence pushing learning technologies towards orchestration.

5. These technologies make orchestration very concrete, i.e. something that can be manipulated by the actors (the tangible aspects) and also something that can be perceived by all actors (the ambient dimension). The evolution of HCI towards physical interaction (tangibles, roomware, …) is supporting the pedagogical evolution towards orchestration.

6. Since orchestration technologies have to integrate different activities into a scenario, they need to include some workflow functionality, i.e. the storage and reuse of data between activities. These activities can be performed with heterogeneous software which makes the workflow more complex to implement.

7. Combining the two previous points, when these digital data are represented as physical objects, the workflow management becomes a public task and hence becomes easier. This applies to the TinkerSheets but also to explain why walking to someone with a PDA to share is richer than sending him the data by WiFi (as in Roschelle & Pea, 2002).

8. Implementing flexibility is a complex design issue. A learning environment that would be completely open, where teachers could change everything, would not convey anymore a pedagogical idea. For instance, in ArgueGraph, changing the group formation criteria from "pair students with opposite opinions" to "pair students with similar opinions" is technically easy but pedagogically meaningless. We refer to this as intrinsic constraints, i.e. design choices that correspond to the core idea of the script (Dillenbourg & Tchounikine, 2007). Flexibility can be increased by reducing extrinsic constraints, i.e. design choices implemented for technical reasons (e.g. data consistency problems if one students drops out) or sometimes simply because the designer did not consider leaving this choice open. Designing environments with a clear scenario but which still allow for flexibility remains a challenging trade-off for which new software architectures would be interesting.

9. The integration of a new learning technology in the classroom legacy (factor 14) is important. It takes ages to convince a student who is used to chat on-line with system-X to move to system-Y. There is a need to design our learning environments from the very beginning either as very open systems that can interact with any other web service or simply to consider our learning environment as web services. The advent of 'cloud computing' may push learning technologies towards a better technical integration.
8.2. Implications for Design-Based Research

Design-based research (Collins 1992; Sandoval and Bell; 2004) relies on prototyping-testing cycles with the participation of all classroom actors. If this participatory design is carefully conducted, it should produce a method that "works well". However, the generalisability of the method and hence of the associated learning effects is quite low. It "worked well" in that classroom with these actors, because it matches their constraints, but will it work in a different classroom? Despite our conviction that DBR is the best method for studying classroom orchestration, we must acknowledge that the generalization remains a problem. Conjecture maps help the interpretation of results and their potential generalization. Conjectures embed the relationship between interventions and expected learning outcomes or expected process changes and are hence not very different from experimental hypotheses. One solution we have chosen for the logistics project was to complement the DBR approach by a more formal experiment with a control group. When a method "works well", it means that it is compliant with the many constraints of the context where it has been tested. It's never "it works well universally" but "it worked well in my class this year". Understanding which constraints were satisfied and why the method satisfied them is the condition to generalize the DBR results. This generalisability is not only for the sake of science but also because it is our social duty to come up with methods that "work well" beyond a single context (even if no method will "work well" universally). We elaborate this in the next section.

8.3. Orchestration as constraints management

1. Teachers who have to orchestrate the classroom and its technologies actually face a multi-constraints management problem.
2. Curriculum constraints: how relevant is the topic with respect to the learning objectives listed in the curriculum, will these students be motivated by this topic?
3. Assessment constraints: are my learning activities compatible with summative evaluation exams, large scale assessments studies,...?
4. Time constraints: how much time is necessary, how much time is available (see constraint 1) and how much flexibility do we have around these two factors?
5. Energy constraints: how much time and energy must teachers engage to prepare and run this method, how long can they sustain it, ...?
6. Space constraints: do I have the space necessary in my classroom do set up these activities, is the classroom layout compatible with the interaction I expect to trigger,...?
7. Safety constraints: Can I keep control of my class? Can I be sued in court because some accident may occur during the field trip? ...
8. There are of course other constraints such as financial constraints (obviously), cultural constraints (the Tinker lamp simulation fits well the specific culture of warehouse workers), teachers' personality constraints (e.g. risk-aversive versus pioneers), motivational constraints (a method that "works well" should boost the teacher's self-esteem),.... Overall, we should mention the prerequisites constraints: no method "works well" if students don't have the pre-requisites. We could build a long list of constraints; these factors have nicely been explored by Bielaczyc (2006). Our point here was to have a shorter list of constraints and to underline which ones are particularly relevant to the concept of orchestration. The six constraint classes mentioned above match the 15 design factors we listed.

8.4. Final word

It is true that "it works well" is a subjective feeling and not a rigorous statement. However, it is a social reality. No successful empirical study will lead to any generalization if the teachers do not acquire the conviction that "it works well". It is hence a key item on the agenda of learning sciences to understand what teachers mean by "it works well". It is a key challenge to improve research methods so that they combine this concern with generalisability.
9. Acknowledgments

The environments listed in this paper have been developed by Fabrice Hong, Guillaume Zufferey, Son Do-Lent, Bertrand Schneider, Hamed Alavi, Frédéric Kaplan, Olivier Guédat. We thank the students and teachers that have been involved in the various experiments. The work on the TinkerLamp is supported by Dual-T, a leading house funded by the Swiss Federal Office for Professional Education and Technology. The work on the Lantern is supported by the Swiss National Science Foundation (Grant PDFM1-118708). The work on ManyScripts was supported by Swiss Center for Innovation in Learning (St Gallen) and a group of KALEIDOSCOPE, a former European Network of Excellence. The work on orchestration is the theme of a group within STELLAR, a new European Network of Excellence on learning technologies.

10. References


Maastricht (The Netherlands), September 16-19, 2008.. Lecture Notes in Computer Science, Springer,


von Inqvald, E (2009) Teachers’ implementation and orchestration of Cabri: Initial use of a dynamic geometry software package in mathematics teaching. VDM Verlag, Saarbrücken


Annex 7: Location-based and contextual mobile learning

Elizabeth Brown1, Dirk Börner2, Mike Sharples1, Christian Glahn2, Tim de Jong2 and Marcus Specht2

1 Learning Sciences Research Institute (LSRI), Exchange Building, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham, NG8 1BB, UK
elizabeth.brown, mike.sharples @nottingham.ac.uk

2 Centre for Learning Sciences and Technologies (CELTEC), Open University of the Netherlands, Valkenburgerweg 177, NL-P.O. Box 2960, 6401DL Heerlen, The Netherlands
dirk.boerner, christian.glahn, tim.dejong, marcus.specht @ou.nl

Abstract: This study starts from several inputs that the partners have collected from previous and current running research projects and a workshop organised at the STELLAR Alpine Rendezvous 2010. In the study, several steps have been taken, firstly a literature review and analysis of existing systems; secondly, mobile learning experts have been involved in a concept mapping study to identify the main challenges that can be solved via mobile learning; and thirdly, an identification of educational patterns based on these examples has been done. Out of this study the partners aim to develop an educational framework for contextual learning as a unifying approach in the field. Therefore one of our central research questions is: how can we investigate, theorise, model and support contextual learning?

1. Introduction

This paper forms a comprehensive review and is intended to act as both an introduction to the field for new researchers and also consolidate aspects of previous publications into a source of critical reference concerned with location-based contextual mobile learning, for those already working in this area.

The distinguishing aspect of mobile learning is the assumption that learners are continually mobile. Rather than seeing learners as physically present in a certain place, such as a classroom or a museum, learners are active in different contexts and frequently change their learning contexts. These contexts can be described in part by a set of parameters, including location, social activity, and learning goals. To gain a fuller understanding, it is necessary to examine how contexts for learning are artfully created through continual interactions between people, technology, and settings, and how these ephemeral learning contexts might be supported and maintained through deployment of new context-aware technologies.

This needs an analysis of what parts of context are important for effective and efficient support of learning. It also requires understanding how learning can be transferred and continued across contexts and life transitions, such as between home and school, or from college to workplace.

A key issue for design of contextual learning technology is whether it is necessary to implement explicit models of context that instantiate and interpret the parameters of location, time, activity, goals and resources so as to offer personalised learning resources; or alternatively whether learning is better enabled by contextual tools such as dynamic maps, guides and probes that offer more generic awareness and support for learners to explore
their environment. The best solutions might come from combining model and tool-based
technologies, which makes technology integration a central concern.

As a first step to relate the current practice in mobile learning to contextual learning we have
performed a concept mapping method with the core question on “What are the educational
problems that can be solved with mobile learning?”. The method and basic results will be
described in the following.

In a second step we have identified current research approaches, projects, and best
practices in location-based and contextual mobile learning. Thirdly, we have identified
educational patterns based on the best practices and the related educational problems.

2. Educational problems solved by mobile learning

To identify the main educational problems that can be solved by mobile learning, 20
international high-profile experts have been involved in a Group Concept Mapping
Methodology. The structured participative approach combines both qualitative and
quantitative methods and makes use of ideas and opinions generated by the experts.

The methodology consists of two stages and a subsequent evaluation and interpretation of
the date. In the first stage the experts were asked to generate some ideas on the following
aspect: “The educational problem that mobile learning tries to solve is…”.

The experts were free to generate as many ideas as they wanted to, while each statement
should describe exactly one specific idea. Thereupon 11 out of 20 experts generated 82
ideas (e.g. “Maintaining continuity of learning across settings, such as between classrooms
and museums on school field trips”) elaborating on the given aspect. For a complete list of
the collected statements see in Annex 1.

In the second stage the experts are asked to sort the generated ideas into groups of
similarity, where a group is defined on how similar the contained ideas are to one another.
Every group is then described with a short phrase or title. Additionally they are asked to rate
the generated ideas on a 1-to-5 scale on importance and feasibility where, for importance, ‘1’
means the statement describes a less important educational problem that mobile learning is
trying to solve and ‘5’ means the statement describes a highly important educational
problem. Respectively, for feasibility, ‘1’ means solving the described educational problem
through mobile learning is not feasible and ‘5’ means it is feasible.

The second stage of the methodology is still in progress. The evaluation and interpretation
results will be included in this deliverable later on.

3. Contextual mobile learning

Context can be broadly defined as the formal or informal setting in which a situation occurs;
it can include many aspects or dimensions, such as location, time (year/month/day),
personal and social activity, resources, and goals and task structures of groups and
individuals.

If learning becomes mobile, location becomes an important context, both in terms of the
physical whereabouts of the learner and also the opportunities for learning to become
location-sensitive. The properties and affordances of one’s location vary enormously and
hence other contexts become even more important, such as the task or goal or the user; the
ubiquity of network access (GPS, wifi etc); the time of the year or day or even the weather.
Seasons can change the visual nature of the landscape whilst inclement weather can turn an
enjoyable day out into a disappointing and demotivating trudge along a wet and muddy
footpath.
The field of context-aware computing has developed a variety of context definitions, mostly starting from location or object context. Zimmermann et al [1] give a pragmatic definition of context. Following their approach the context of a person or an object can be defined by five distinct parts:

- **identity context**, this includes information about objects and users in the real world. With respect to users, their profile can include preferences, acquired-desired competences, learning style etc. This facet of context can also refer to information about groups and the attributes or properties the members have in common.

- **time context** ranges from simple points in time to ranges, intervals and a complete history of entities

- **location context** is divided into quantitative and qualitative location models, which allow to work with absolute and relative positions

- **activity context** reflects the entities goals, tasks and actions

- **relations context** captures the relation an entity has established to other entities, and describes social, functional and compositional relationships.

Generally speaking the notion of context-aware systems originated out of ubiquitous computing and the adaptation of a computer system to its changing environment. Computers that become mobile or embedded in different environments should basically be able to sense their environment and react to environmental changes. In ubiquitous computing, context is used in two ways: to encode information in order to aid later retrieval; and as a means of personalising the end user experience depending upon events from the user or usage patterns [2].

Early work in context-aware computing was largely directed towards integration and abstraction of data from environmental sensors (such as absolute and relative time and the user’s physical setting including location, surroundings, and conditions), plus features indicating the user’s current activity, preferences and social surroundings (such as the availability of other users nearby or accessible online), see Want et al [3] and Abowd et al [4] for some examples. This approach considered the environment as a shell encasing the user, which can be described by scalar properties such as current time, location (positioning coordinates) and a list of available objects and services. The work has led to successful demonstrations, providing context-based content and services. For example, researchers from the University of Birmingham developed mobile technology that can provide location-dependent services that adapt to whether the user is walking, standing or sitting [5].

The problem with this approach is that it models the user as receiving data from an environment, rather than interacting with it. It has been noted that this model is inherently limited, and does not capture some important and useful aspects of context (see e.g. Dourish [2]). The ‘environment as shell’ approach does not acknowledge the dynamics of interaction between people and their environment, for example as we move or modify objects around us to create a supportive workspace or form an *ad hoc* social network out of people with shared interests either in the immediate location or available online. It leads to the classic ‘AI Frame Problem’, of determining which aspects of the current situation are relevant, and how these can be operationalised as a set of machine-interpretable features. Furthermore, if we regard context as a negotiated construct between communication partners in the world, then it is likely that context as acquired by sensing mechanisms might not match the continually evolving negotiation.

This study – whilst embracing aspects of ubiquitous computing – has a focus upon ubiquitous, or mobile, learning. We believe these terms to be related but completely distinct from each other, which each having its own emphasis.

Mobile learning is not just about the mobility of the learner or the device, but also mobility across contexts. As we spend more time physically on the move, it is essential to realise that
contexts might change rapidly; this is also true in the more long-term sense of change, which might encompass lifelong learning. A big question for technology-enhanced learning is how contextual mobile learning can be supported by various learning scenarios and the technologies/devices being used.

The MOBIlearn project examined context in detail and developed a Context Awareness Subsystem (CAS), intended to provide a way of recommending content that was context-dependent and also to store these recommendations. Context was seen as “a dynamic process with historical dependencies” – in other words, a changing set of relationships that may themselves be shaped by those relationships. For example, the information presented to someone visiting a museum or art gallery for a second or third time, might not be as appropriate for someone for whom it is their first visit. Context was not seen here as mere data, provided by the triggering of sensors, but more as an interactive negotiation between people and their environment, including local artefacts/resources and activities of the user.

Synchronizing learning activities with the physical environment in that sense can be concluded as a promising approach from various theories of learning and cognition. According to information processing theory and cognitive load theory, human working memory has limited capacity and learning content should be structured such that the information load does not overwhelm the learner. Furthermore, multimedia learning theory states that each sensory channel (visual and auditory) has limited processing capacity and learning is optimal when the information presented on one sensory channel augments that presented on the other. On the one hand, the limited processing capacity means the information delivered to the learner should be limited to the information relevant in the current learning context. On the other hand the complementary distribution of information given on different channels must be given. The implications for contextual learning are that information with and across contexts must be appropriately structured to reduce cognitive load and thus maximise retention.

Wenger and Lave state that knowledge needs to be constructed by the learner in a realistic context that would normally involve the application of that knowledge. An authentic learning environment often provides a variability in stimuli, or multiple perspectives on the theory learnt, and needs context-dependent, highly interconnected knowledge; several aspects that are emphasised by cognitive flexibility theory as important for learning. Especially, the variability in stimuli and learning tasks available in an authentic context may result in a better generalisation of the knowledge constructed by learner. As stated with the encoding specificity theory, a realistic context would moreover result in improved recall because the stimuli presented at the time of recall (the authentic situation) would be most alike to stimuli present at the time of learning (learning in a realistic context).

Conversation Theory is a systems theory of learning that describes the process of coming to know as a continual sequence of conversation within and between individuals, groups, and interactive technologies. Successful learning occurs when the participants are able to engage in a shared conversation about the learning topic itself and about the aims and processes of the learning, continually adjusting their understanding through dialogue and exploration. Learning contexts need to be designed to enable such conversations, by providing learning materials that match the learners’ current understanding, and also by provoking and facilitating dialogue and exploration.

3.1 Dimensions in contextual learning

To explore the dimensions in designing contextual learning, a review of current systems for mobile contextualised learning support, presented in de Jong et al [17], is used. Furthermore the authors introduce a reference model that can be used to classify the current research, to identify limitations of current applications, and to discuss new solutions and challenges for contextualised learning support. Table 1 shows the presented reference model that is

---

17 http://www.mobilearn.org/
comprised of five dimensions: content, context, information flow, purpose, and pedagogical model. For each dimension, the possible values are given in the column below.

Table 5: A reference model for mobile social software

<table>
<thead>
<tr>
<th>Content</th>
<th>Context</th>
<th>Information flow</th>
<th>Pedagogical model</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annotations</td>
<td>Individuality</td>
<td>One-to-one</td>
<td>Behaviourist</td>
<td>Sharing Content and Knowledge</td>
</tr>
<tr>
<td>Documents</td>
<td>Context</td>
<td>One-to-many</td>
<td>Cognitive</td>
<td>Facilitate</td>
</tr>
<tr>
<td>Messages</td>
<td>Time Context</td>
<td>Many-to-one</td>
<td>Constructivist</td>
<td>Discussion and Brainstorming</td>
</tr>
<tr>
<td>Notifications</td>
<td>Locations</td>
<td>Many-to-many</td>
<td>Social</td>
<td>Social Awareness</td>
</tr>
<tr>
<td></td>
<td>Environment</td>
<td></td>
<td>Constructivist</td>
<td>Guide</td>
</tr>
<tr>
<td></td>
<td>or Activity</td>
<td></td>
<td></td>
<td>Communication</td>
</tr>
<tr>
<td></td>
<td>Context</td>
<td></td>
<td></td>
<td>Engagement and Immersion</td>
</tr>
<tr>
<td></td>
<td>Relations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>context</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The five dimensions and corresponding values describe the following aspects of contextualised media for learning:

- The content dimension describes the artefacts exchanged and shared by users. In an analysis of the literature the main types of artefacts found were: annotations, documents, messages, and notifications.
- The context dimension describes the context parameters taken into account for learning support. The five values for the context dimension are based on an operational definition of context [1].
- The information flow classifies applications according to the number of entities in the systems involved in information flows and information distribution.
- The pedagogical paradigms and instructional models describe the main paradigm leading the design of contextualised media and the integration of media in real-world contexts.
- The purpose describes applications according to the goals and methods of the system for enabling learning.

Thus, on the one hand, the reference model describes the manipulated knowledge resources, the context in which they are used, and the different flows of information. On the other hand, the higher-level concepts of pedagogical model and purpose define how the content, context, and information flows are used and combined. Hence, by combining different values for each dimension, various forms of contextualised software can be created for different purposes and with different pedagogical underpinnings.

For example, a system with a main purpose of sharing content and knowledge between its users, can be described by using documents from the content dimension, relations context to describe social relations between the users, and a many-to-many information flow. Another example is a location-based information system like RAFT [2], which combines (1) the creation and delivery of documents with (2) locational context, (3) a one-to-many information flow to provide (4) a social constructivist approach for (5) increased engagement and immersion.

During this exploration also some limitations of mobile contextualised learning solutions have become clear. Summarising the following extensions to current state-of-the-art can brought forward based on these limitations provide more integrated systems with a range of functionality:

- better and wider use of metadata,
- more advanced and wider use of notification techniques,
- an improved adaptation to the user’s personal preferences and learning environment or situation by using more kinds of context information than location and identity alone, and use of techniques to derive more detailed or higher level context information by a combination of different context parameters,
more attention to systems aiming at informal and lifelong learning.

3.2 Case studies

Location-based (but not location-aware)

It is essential to note that some of the earlier projects into location-based learning were not themselves location-aware in terms of the technology, due to the limitations of the hardware available at the time. However, these projects were important pre-cursors for later research and so they are included here.

One of the first documented computer-supported field trips was “Wireless Coyote” in 1993. Twenty 11-year old children, divided into 5 groups, used modified tablet PCs to record and share information relating to the environment on a school field trip. This allowed students to share data in real-time with their peers in the other groups and challenged the traditional role of the teacher as well as the students’ own expectations [18].

Another example of how mobile computing has been used in field work is shown in the “Cornucopia” project [19] in 1997. This project involved undergraduate students in a formal teaching environment using mobile devices to record data relating to different varieties of corn, grown in a university test plot. Data was stored on flash memory cards, which were then handed in to the class instructor for them to collate data from each device back in the laboratory, and thence uploaded to a shared web space to facilitate future class discussions.

A related pilot application, “Plantations Pathfinder”, was designed to provide information to visitors in respect of a garden attraction [19]. This electronic guide (also circa 1997) could be updated for tourists much more quickly than printed matter and could record a visitor’s interests as they toured the garden with the device. It also enabled collaboration via input of data that was then uploaded to an online discussion forum, allowing users to view previously-inputted information from others as well as share their own experiences.

Location-based and location-aware

As technological advances were made throughout the 1990’s and networking infrastructures improved, so did the affordances for creating truly location-aware experiences. These projects managed to utilise not only location-based work, but also used devices that were themselves aware of their geospatial position.

Cyberguide was a project from 1997 to investigate handheld intelligent tour guides. Early demos were produced that enabled a visitor to locate nearby attractions, including local bars in Atlanta (a variation named CyBARguide) [4].

Another example of how these location-aware technologies were used can be found in the work of Pascoe et al in 1998, where a PalmPilot running the “stick-e note” system was used by ecologists in Africa to record contextual data relating to the behaviour of giraffes [20]. Context in this project included data corresponding to location; date/time; number of males, females and juveniles; feeding habits; vegetation type/amount and location of faecal samples.

Another location-aware system is “GUIDE”, an electronic tourist guide for use by visitors to Lancaster [21]. GUIDE provided a means of personalising information in respect of a visitor’s own interests and also various environmental parameters. It consisted of an information retrieval system; a map-based navigation system; and also a guided tour, created by the user, depending on what attractions they wanted to see.

In 2002 came the KLIV project. This explored the effects of self-produced videos, filmed on location in the workplace, on the professional education of nurses. It used barcode-equipped medical devices and PDAs with barcode scanners to provide the videos in context. When
needed to be reminded about a piece of equipment or task, nurses could scan the barcode in order to view the video [22][23].

From 2004-2009 a number of environmental education projects have emerged, some of them funded through the MOBILEARN and Equator frameworks. Ambient Wood [24] was an Equator project, in which a playful learning experience was developed where children explored and reflected upon a physical environment that had been augmented with a combination of digital constructs. Savannah [25] was a strategy-based adventure game where a virtual space was mapped directly onto a real space, with school pupils playing at being lions in a savannah, navigating the augmented environment with a mobile handheld device. Using aspects of game play, it challenged children to explore and survive in an augmented space, by successfully adopting strategies used by lions. CAERUS was a context aware educational resource system for outdoor tourist sites and educational centres, consisting of a handheld client and a desktop admin, to provide a visual interface to add maps, define interest regions, etc. It delivered this information to visitors (such as those at the University of Birmingham’s Botanic Garden at Winterbourne) through Pocket PC devices with GPS capability [26]. MyArtSpace was a service that let children 'collect' items in museums and galleries using mobile phones and then build presentations with them [27][28]. It was later developed into OOKL, which is currently running as a commercial enterprise [29].

In the RAFT project, live video conferences have been used to establish a video link between an expert interview in the field and a classroom from which learners could ask questions. As one important finding not only do the participants in the field trip profit from the excursion, but more importantly students in the classroom develop a more realistic interpretation of application contexts and are more interested in the topic in general. Gender differences in the use of new technology can also reduced by such approaches [30].

Another museum-based project is Ubicicero, a multi-device, location-aware museum guide in which artefacts are tagged with RFID for users to interact with via mobile devices [31]. It also takes into account a user’s position and their behaviour history, as well as the type of device available.

dgeoMole was an application developed to support Geographical Information Science (GIS) fieldwork [32]. It presented digital reconstructions of the landscape via a Windows Mobile PDA, allowing on-screen sketching and the delivery of audio relevant to the landscape visible from that point. geoMole was developed as part of the SPLINT (SPatial Literacy IN Teaching) Centre for Excellence in Teaching and Learning (CETL). Another project with an environmental focus was Environmental Detectives [33][34]. It was an example of a game-based scenario in which students were asked to investigate a spill of toxic waste into the environment.

Language learning has also been a focus of some location-based learning, such as the LOCH project [Language-learning Outside the Classroom with Handhelds]. Paredes et al [35] examines how students participate in ‘field activities’ assigned by the tutor, where everyday tasks take place in real physical locations. Recent work by Fisher et al also explored language learning through different modalities, including handheld learning devices [36].

Besides these case studies, important platforms for developing location-based experiences have also been created. These include mScape (that allows media to be associated to geographical points and regions on the Earth’s surface, creating location-aware

---

18 MOBILEARN was a worldwide European-led research and development project from 2002-2006 exploring context-sensitive approaches to informal, problem-based and workplace learning by using key advances in mobile technologies [http://www.mobilearn.org/].

19 Equator was a six-year Interdisciplinary Research Collaboration (IRC), supported by the EPSRC, which focused on the integration of physical and digital interaction [http://www.equator.ac.uk/].

20 http://www.ookl.org.uk

21 http:// www.splint-cetl.ac.uk
“mediascapes” running on mobile phones or PDAs [37]) and PaSAT (a system to build mobile learning games with physical environments as playing spaces, using GPS-enabled PDAs to create location-based activities). PaSAT has been developed by Peter Lonsdale\textsuperscript{22}, a PhD student at the University of Nottingham, and used for the underlying framework for the “BuildIt!” game, used by school children to learn about the physical environment of their school in a goal-oriented game to help structure the children's planning and reflection.

3.3 Alpine Rendevous workshop contributions

\textit{Learning Sciences Research Institute (LSRI), University of Nottingham}

At the University of Nottingham, there are several projects surrounding contextual learning in location-based scenarios, including the aforementioned work with PaSAT/BuildIt.

In the PI project\textsuperscript{23}, secondary school pupils have been undertaking inquiry-based science investigations inside and outside the classroom, supported by a computer-based toolkit that structures their learning activities. In one study the children explored micro-climates in their school grounds, measuring variables such as wind speed and temperature to determine where to site a bench or fly a kite. The toolkit provided a means to probe the environment and collect data, and also to guide them through the inquiry learning process and enable them to present results in the classroom.

The LSRI has also been collaborating with colleagues elsewhere in the University of Nottingham, most notably the School of Geography and the Centre for Geospatial Science. Research is being carried out into how geography students create augmented visitor experiences for tourists to the Lake District, using a range of mobile devices and techniques. This has led to some interesting results relating to affordances of the technologies and end-user experience [38] and has informed future work planned into the placement of media (particularly audio) in the landscape to deliver location-specific information. The most important contexts being considered in this research include a person’s orientation to their landscape; viewsheds (maps of the visibility of the surrounding area); weather (related to screen visibility of the device and also the behaviour of the visitor); motivation/engagement of the visitor; type of visitor (school children; general public; groups or individuals); type of interaction (informal or formal); the appropriate placement of media related to a visitor’s geographical position; and the suitability of different media types used to deliver information.

Previous related research also includes the “Answer Tree” system, a mobile location-based game that used collaborative game card-collecting as a means to teaching children about characteristics of trees [39]. In this project, placement of media related to the technology was an important contextual constraint: GPS visibility is not very good when stood underneath a large tree canopy. Other aspects investigated included the nature of the interaction (both with the environment and with the other groups of learners) and the type of learning taking place ("jigsaw" learning [40]).

There is also ongoing research into informal environmental education, using crowd sourcing as mechanism for creating community-driven media for sharing with the general public. Aspects of this work under particular consideration include the informal learning scenarios that are created by these conditions and other technical and social research questions connected with ubiquitous access, user-generated content, tagging and annotation.

\textit{Centre for Learning Sciences and Technologies (CELSTECC), OUNL}

\textsuperscript{22} http://www.peterlonsdale.co.uk/

\textsuperscript{23} The “Personal Inquiry: Designing for Evidence-based Inquiry Learning across Formal and Informal Settings” is a project from the University of Nottingham in partnership with the Open University to help school students learn the skills of modern science, including learning in informal outdoor settings.
The Centre for Learning Sciences and Technologies at the Open University of the Netherlands works on location-based and contextual mobile learning in several national and international projects, such as MACE\textsuperscript{24}, GRAPPLE\textsuperscript{25}, or STELLAR\textsuperscript{26}. Thereby the focus is on informal and incidental learning, where learners are mobile by changing contexts. Context as such is seen broadly and includes locations, processes, peers, concepts, and events. The projects aim to support learners’ physical mobility and transitions between contexts through detecting, linking, and constructing contexts.

Within the MACE project the delivery, creation, and metadata enrichment of architectural content on mobile devices has been researched. The MACE content can be delivered to the mobile device using location-based and augmented reality browser, such as Aloqa\textsuperscript{27} and Layar\textsuperscript{28} that utilise the available contextual metadata.

On the basis of search filters, where the created content is enriched with metadata specifying certain values of these filters the “ContextBlogger” client uses a GPS location-based filter that provides the users with MACE content about the real-world objects in their vicinity. The client combines social software, a weblog, with information about the context of a learner. The information in the weblog can be accessed using a mobile device, and the content can be filtered through the application of search filters based on context information. The search filters for the contextualised blogging application retrieve the content either related to a specific real-world object or to a specific user location. Furthermore, the learner can also choose to create his/her own content and relate it to a real-world object or location. Therefore, the use of contextualised blogging provides a basis for an investigation of the usage of physical artefacts in learning. On the one hand, the combination with a physical object could provide the basis for learning, on the other, shared objects could be used to build communities of practice and couple the creation of learning networks to physical objects.

With “Locatory”\textsuperscript{29}, an augmented reality game has been developed. The game builds upon an open source augmented reality framework for mobile devices to render virtual artefacts that are laid over an image of the real world. The goal of this project was to explore learning scenarios especially games the make use of Augmented Reality (AR).

Other institutions

A number of other projects and research into location-based contextual learning are being carried out across Europe and North America. These were showcased at a workshop held as part of the STELLAR Alpine Rendezvous in December 2009\textsuperscript{30}, where participants shared their experiences of the work they were carrying out in location-based contextual mobile learning.

Some researchers are focusing specifically on mathematics education. Brendan Tangney [CRITE, Trinity College Dublin, Ireland] is investigating learning mathematics via mobile technology in a real-world setting (e.g. estimating the height of a building monument using the accelerometer of a phone (angle) and GPS distance, or calculating it with estimates on the basis of photographs). Monica Wijers and Vincent Jonker [Freudenthal Institute, Utrecht University, the Netherlands] discussed their work with MobileMath, a GPS-based game to help students learn maths and geography by creating a virtual shape in a real world playing

\textsuperscript{24} Metadata for Architectural Contents in Europe - http://mace-project.eu
\textsuperscript{25} Generic Responsive Adaptive Personalized Learning Environment - http://www.grapple-project.org/
\textsuperscript{26} The STELLAR Network of Excellence - http://www.stellarnet.eu/
\textsuperscript{27} Aloqa Mobile Service - http://www.aloqa.com
\textsuperscript{28} Layar Mobile Augmented Reality Browser - http://layar.com
\textsuperscript{29} Locatory: http://code.google.com/p/locatory/
\textsuperscript{30} “Education in the Wild: contextual and location-based mobile learning in action”, organised by Elizabeth Brown and Mike Sharples (report available Spring 2010).
field. Jonker is looking to extend this work with an idea called “living points of interest [POI]”. The current idea behind LivingPOI is to design gameplays for a set of 6 mini-games that fit within the boundaries of a playground at school where all children have a RFID-tag (passive or active) and where three RFID readers are placed around the playground in order to log all geo-positions during a 10 to 20 minutes gameplay. Jonker’s ideas for mini-games include:

- development of an epidemic virus
- making geographical patterns like squares, triangles
- measuring density during a game where all children move from one place to another on the playground

Other projects include research into the use of handheld devices for use by student teachers (Jocelyn Wishart, Graduate School of Education, and University of Bristol, UK). Differences in the way these devices were used and accepted by different groups of student teachers (science vs Modern Foreign Languages) has left Wishart considering the current social and ethical practices with respect to mobile devices. She has also carried out work into location-based learning through collaborations with colleagues from WildKnowledge and the ‘Mudlarking in Deptford’ project. Wishart’s continuing focus remains on investigating clear, achievable codes of practice for students, teachers and researchers engaged in location based learning to ensure that these opportunities for engagement and learning are successfully integrated into educational systems across Europe.

Gill Clough from the Open University (UK) has researched the influence of location-aware mobile and social technologies for creating and consuming content in the field/informal learning. In particular, she worked with the Geocaching community (a geographically dispersed group who carry out hiding and finding of hidden packets [Geocaches] in the physical world, guided by GPS enabled mobile devices) to carry out this work. The research focused on the activities of community members rather than on a particular piece of mobile or social technology and uncovered detailed information about innovative informal and collaborative learning embedded into the practices of the community. It revealed the considerable efforts individual community members went to in order to create and engage with a variety of location-based informal learning opportunities.

Jacqui Taylor from Bournemouth University (UK) is looking at the impact of educational technology on the psychology of learners (e.g. the individual differences in emotions and affect involved in learning). The key focus of her emerging work in this field is to develop our understanding of learning styles/strategies in contextual and location-based mobile learning, including the evaluation of the emotional impact on the learner.

John Cook from the Learning Technology Research Institute, London Metropolitan University (UK) has written about extending Vygotsky’s Zone of Proximal Development (ZPD) into a Responsive Context for Development (RCD) and provided two cases that illustrated how this was achieved using location aware mobile devices. They specifically set out to design a mobile RCD that is able to respond supportively to, and that acts as a challenge for, learners in Higher Education. The first case was working with school pupils examining change in their playground. They specifically set out to design a mobile RCD that is able to respond supportively to, and that acts as a challenge for, learners in Higher Education. The first case was working with school pupils examining change in their playground. The second case took place in Yorkshire, where archaeology students explored Cistercian chapels and used 3D visualisations of wire-frame models on handheld devices.

Leilah Lyons (Computer Science & the Learning Sciences, University of Illinois at Chicago, USA) has investigated the use of mobile devices to support co-located, synchronous collaborative learning activities in formal learning environments (classrooms) and informal learning environments (science museums). On both cases, she worked with small groups of users who employed mobile devices as personalized interfaces to a simulation activity hosted on a nearby computer. She is collaborating with a zoo on her next project, where they aim to make real animal behaviour the phenomena of interest by placing GPS collars on free-roaming peafowl. Students will remotely track and study the movements of these birds over time in their classrooms. During field trips, they can examine hypotheses devised from afar (e.g., do the birds prefer areas with vegetation cover or open areas?) using GPS-equipped mobile devices to locate and document locations. Visitors will also be able to
examine this long-term data, thus enabling location to act as a connection point between the two types of user and their learning activities.

Finally, Nicola Beddall-Hill from City University (London, UK) has been looking at mobile learning outside the university classroom in case-based learning activities on field trips. Her research investigates the influence that mobile devices have on learning processes and outcomes, and their influence on the use of the device in fieldwork settings. It compares the use of devices introduced for teaching and personal devices used during the learning activity by the group or individual. This project is not based on delivering an intervention per se but instead is a case study into current practices of TEL used on field trips.

4. Educational patterns in location-based and contextual mobile learning

Mobile learning technologies and support can be used in a variety of educational settings and domains. In the following section we will first identify some basic educational functionality connected to the interaction patterns described above. Furthermore, a key idea of contextual learning is to connect the real world with digital media, which is essential for the learning gain. The characteristics of a situation in the real world and the metadata of digital media in this sense are the key parameters that are used to synchronise the augmented reality experience.

![Diagram of connecting real world and digital media via context](image)

Figure 2: Connecting real world and digital media via context

The different interaction patterns enable different experiences that can give different perspectives, insights and forms of interaction for an augmented real world experience. As an example, Head-Up display interaction mostly enables users to get an enhanced vision of their current environment. The educational tool in this case basically filters information according to the current location and the direction a user is looking. The merge between user vision and tags in the enhanced vision basically visualise geo-located information or
interaction facilities. In that sense, a first cluster of educational patterns can be described as ‘Geo-Location Layers’. All kinds of applications enable the perception and interaction with location related information, this basically happens in two ways: on the one hand with a Head-Up display, the augmentation of the video signal of the mobile device enables the direct linking of an augmentation layer and a view of the real world; on the other hand, a more map-oriented metaphor enables the relation of several layers onto maps which gives a more spatial relationship between the entities of the layers.

A second cluster of educational patterns is the “Tagged Environment”. This subsumes approaches in which either computer readable tags are placed in the environment on specific locations or objects, or where these tags are used to identify interaction artefacts with which users can manipulate complex visualisations. With such a tag-based approach, different kinds of game play approaches can also be identified, that are used for learning or enabling other perspectives on learning content.

The most advanced and probably most technically challenging form of mobile AR is the ‘X-Ray Vision’ pattern, as it requires real time mapping of complex structures carried out on the mobile device.
<table>
<thead>
<tr>
<th>Title</th>
<th>Layers of Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>The 'Layers of Interest' (LOI) pattern describes educational scenarios in which users consume information, while on the move. The information is delivered via the users' mobile device and can be manifold, ranging from location-based historical information to contextual language information. The location artefacts (e.g. buildings, landmarks, artwork) are stored in databases enriched with contextual metadata, such as location. Applications that support the pattern usually categorise the available information in 'layers' (e.g. architecture or history).</td>
</tr>
<tr>
<td>Educational Background</td>
<td>The educational focus of this pattern lies in the exploration of physical spaces. In both cases location related data is presented to a learner. Nevertheless via the selection of channels the learners can filter different views on their environment. The pattern basically can only filter information on the level of GPS and compass information, which typically can be used to either filter for physical objects of bigger size. The exploration of the environment also enables the visualisation of learning opportunities in the physical environment of the learner. Mobile learning makes it possible to support the learners in an authentic learning environment. The importance of an authentic context in learning is emphasised by several theories. For example, Dewey [41] and Wenger and Lave [11] state that knowledge needs to be presented in a realistic context that would normally involve its application. Moreover, the functional context approach described by Sticht [42] emphasised that assessment of learning requires a context/content specific measurement in an authentic work context. An authentic learning environment often has a complex and ill-structured nature that provides multiple perspectives on the theory learnt, and needs context-dependent, highly interconnected knowledge. Several aspects that are emphasised by cognitive flexibility theory [12] are as important for learning. Especially, the variety of stimuli and learning tasks available in an authentic context may result in a better generalisation of the knowledge constructed by learners [12][14]. Cognitive apprenticeship, as described by Collins et al [43], supports the learner in dealing with the complexity of an authentic environment by social interaction between a learner and an expert tutor. An important process in a cognitive apprenticeship is the reflection about action carried out by the expert in an authentic work environment [44][45]. In managing the complexity of an authentic environment, the learner can focus on certain cues and information in that rich environment which helps to reduce the information load and will lead to more effective learning.</td>
</tr>
<tr>
<td>Related Patterns</td>
<td>Head-up Learning Layers, Tricorder Layers</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title</th>
<th>Head-Up Learning Layers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>The pattern is deduced from the LOI pattern. As in most cases the sensors for relating information within mobile devices as GPS location, compass, and accelerometer the objects from a learning layer can be visualised in a certain direction, distance, and height from the viewers perspective holding the mobile device. This allows good results for giving estimations about the current environment but does not enable the viewers to develop a bird’s eye perspective on the relationships between objects in the augmented view. On the other hand this kind of augmentation allows senseful learning support when ego-perspectives are necessary and helpful.</td>
</tr>
<tr>
<td>Educational</td>
<td></td>
</tr>
</tbody>
</table>
### Background

The educational background and potential of this pattern can be seen mostly in an explorative approach enabling the user to explore his/her current environment with different layers activated on a display.

The pattern appears to be mostly appropriate for subjects where the relation between the entities in the layers is differing according to the viewers’ direction or distance. Furthermore the merge of different layers that can be structured task specific or the mapping of digital media to locations could become the actual task for a learner. As an example for the last one the Locatory game was created. In the Locatory game the user has to do a mapping of a picture presented to him/her to the real world and drop the picture virtually on the real world object. This technology would also enable assessment methods and collection tasks related to real world objects. Examples can be found in domains as architecture, biology, geosciences.

### Related Patterns

<table>
<thead>
<tr>
<th>Title</th>
<th>Layers of Interest</th>
</tr>
</thead>
</table>

### Title

**Tricorder Learning Layers**

### Description

This pattern basically is a mash-up of contextualised information sources. In the pattern information layers are merged mostly with map information. The basic view is a map that is rendered according to the viewers’ location, and viewing direction and overlaid with a learning layer. Tricorder can also work with head-up perspectives but focus more on the metadata about a certain entity (example could be the statics for air pollution in different directions of the viewers perspective).

### Educational Background

This allows the learners to relate the layered information streams with their current location. It provides the learners with opportunities to learn relations between implicit of explicit characteristics of the current location and situation with digital data sources. As a difference to the head-up display approach in the tricorder pattern the relation between the real world view and the visualised objects is not the most important learning goal, the tricorder pattern enables more to visualise relations between different layers and the entities of these layers.

### Title

**Tagged Environments**

### Description

This pattern enables the integration of digital information via the integration of codes in the physical environment via real time or asynchronous scanning. In learning situations where the actual objects with which the users learn become of smaller size or are presented in an indoor scenario tags are often used to enable simple computer-based sensoring of the context and to trigger integration of augmented visualisations.

The probably most prominent example of this pattern is the MagicBook, in which elements of a physical book can either be made alive of presented in augmented reality by scanning printed barcodes in the book.

### Educational Background

Tagged environment link an educational function, interaction facility, or metadata about objects and the environment to a concrete location or object. The approach gives the learner the possibility to interact in tag-enriched environments with exploration, interaction, or augmented collaboration playfields.

### Related Patterns

| Tagged Artefacts, Tagged Locations, Tagged Interaction Objects, Tagged Playfields |
### Tagged Artefacts

**Description**

The pattern is based on the MagicBook approach; basically it allows users to look at 3D models of learning content. A good example is learning about the solar system by manipulating the complete view of the solar system by moving the personal perspective. Or a Lego 3D becoming alive based on a tag on a packet.

**Educational Background**

The theory is that the traditional methods of learning spatially-related content by viewing 2D diagrams creates a sort of cognitive filter. This filter exists even when working with 3D objects on a computer screen because the manipulation of the objects in space is made through mouse clicks. I am investigating the possibility that the physical manipulation of the earth-sun virtual models in augmented reality will provide a more direct cognitive path toward understanding of content.

### Tagged Locations

**Description**

This pattern integrates codes in the physical environments as the visualisation of the information is highly related to a very specific location.

Tags are used to enrich the objects within the physical environment. The artefacts are usually selected and enriched in an authoring environment and can then be accessed through mobile applications.

Furthermore two types of tags can be distinguished, as they enable an active or passive access to the available information. To access information passively means to get the information presented in context (e.g. location), while the active access demand an interaction using physical tags (e.g. barcode). Using the mobile device the user can trigger an action by reading the tag.

**Educational Background**

A situated view on an object and on the relation between different tagged locations is supported. In this pattern the exact location of the tag is important as the overlay produces a unique merged view for the learner that combines properties of the real world objects and the digital media for augmentation.

### Tagged Playfields or Interaction Objects

**Description**

In this pattern users basically can use real world objects to manipulate augmented worlds, so by moving tagged objects user scan combine molecular structures or build new environments that are visualised in 3D. Furthermore this kind of technology is used in games where user scan interact with playfields via tags by viewing and interacting with virtual objects embedded in the real world.

**Educational Background**

This gives the learner an isometric perspective on the learning content as also is the basis for collaborative augmented reality approaches, in which learners have personalised views on playing fields augmented via tag recognition.

### 5. Summary

This deliverable for STELLAR Work Package 1 has included a comprehensive review into current research into location-based and contextual mobile learning. It has discussed the dimensions important in contextual learning and has presented detailed case studies of location-based research from the past decade. It has also showcased the work of the authors (both in the UK and in the Netherlands) and those who took part in the STELLAR Alpine Rendezvous ‘Education in the Wild’ workshop in December 2009, thus providing an
excellent overview of research in this area being carried out across Europe and beyond, and the issues and challenges being addressed by academic colleagues.

We have also asked experts from the field of mobile and contextual learning to contribute to a concept-mapping study to identify the main issues and challenges that can be addressed by mobile learning. The results from this study are ongoing and will be included in a later version of this report.

Lastly, we are working on an identification of educational patterns, based on examples from the expert concept-mapping study, to develop an educational framework for contextual learning as a unifying approach for the field. This is already partly presented in this paper and the final version of the report will include this in more detail.

6. References


[29] S. Johnson, "Evaluation of the use of mobile phone technology (OOKL) as a recording tool and indicating children’s reaction to working both indoors and out of doors at Kew Gardens."


Annex 1 Statement List

1. Limited access by some learners in remote locations.
2. Lack of support to young learners, which have the mobile technology.
3. Insufficient real life experience in the learning process.
4. Nomads who move from one location to the next while learning.
5. Lack of community building during learning.
6. Low motivation of learners who are mobile technology literate.
7. Not enough collaboration between learners.
8. Learning from any location.
9. Learning at anytime.
10. Learners not able to interact with experts from around the world.
11. Just in time information for immediate application.
12. Learners cannot learn in context.
13. Teachers not comfortable using mobile technology.
15. Not enough self-directed learning activities while learning.
16. Ability to discover and experiment in own context.
17. Access to learning resources and learning opportunities without the restrictions of location, time and cumbersome equipment or facilities.
18. Provision of opportunities to contribute to the development/production of learning resources and course content without the restrictions of location, time and cumbersome equipment or facilities.
19. Provision of opportunities to collaborate, share and publish learning resources and course content without the restrictions of location, time and cumbersome equipment or facilities.
20. Actively participate in learning activities outside of formal educational settings and facilities.
21. Enhance teaching and learning within formal educational settings and facilities through handheld technologies.
22. Finding new teaching methodologies that are suitable for the challenges of, and embraces the opportunities of, the knowledge and information age.
23. Finding new learning strategies that are suitable for the challenges of, and embraces the opportunities of, the knowledge and information age.
24. Flexibility for the learner.
25. Mobility of the learner.

27. Cost-effectiveness for the providers of teaching and learning.

28. Outside in, inside out problem, where cultural practices involving new digital media can be brought into formal learning institution, get enhanced inside the institution and in turn feed back into the digital world at large.

29. The design of augmented contexts for development problem to enable collaborative problem solving where learners generate their own ‘temporal context for development’.

30. The provision of access to knowledge in the context in which it is applied.

31. Refreshing the image and practice of institutional e-learning.

32. Helping educational institutions understand the increasing & near-universal ownership, acceptance and use of mobile devices across society.

33. Taking education out of classroom settings into meaningful settings.

34. Helping educational institutions to offer learning aligned to the students' ownership, experience & use of technology.

35. Perceptions of technologically impoverished provision.

36. Make use of the affordable technologies that students have access to.

37. Design suitable activities for the mobile learners.

38. Assess learning experiences to be accountable for the stakeholders.

39. Interacting with your environment to achieve new knowledge from it.

40. Anything is a potential learning scenario.

41. Easing access to educational opportunities.

42. The perception that there is a lack of student engagement.

43. Students exhibit passivity, boredom, indifference, low attention spans, and fail to complete their studies.

44. Rigid assessment systems stifle creativity and innovation.

45. Inequality of access to computers, learning resources and teachers.

46. Pressured, busy, fragmented, mobile lives leaving little quality time for conventional, place-and-time-dependent education.

47. Blinkered, old-fashioned views about education stopping when working lives begin.

48. Traditionally ineffective instruction and low learner performance in some subjects.

49. Gaps (time lags) between traditionally scheduled learning sessions, limiting achievement, teamwork and collaboration.

50. Under-utilization of potentially rich learning resources in heritage sites, art collections and all sorts of other interesting places.
51. Enable learners in classroom settings to have equal access to rich resources and computational tools to support curriculum learning.

52. Orchestrate new forms of classroom pedagogy that require coordination of individual, small group and whole class activity.

53. Connect learning across contexts, including between formal and informal settings.

54. Maintaining continuity of learning across settings, such as between classrooms and museums on school field trips.

55. Enable enquiry-based learning in novel locations, through novel locations, and about novel locations.

56. Making use of space and environment as a backdrop for engaged spatial learning.

57. Making use of affordances of locations to support learning.

58. Using technology to probe or to enrich understanding of the natural environment, and annotating the environment for the benefit of visitors.

59. Access to information when and where it is required, through 'just in time' browsing of relevant information, and information push to support learning in context.

60. Enable learning through distributed conversation across contexts.

61. Accessibility of information in relevant everyday life and work situations.

62. Documenting real time experiences of learners.

63. Contextualization of e-learning.

64. Transfer of training.

65. Spontaneous collaboration in situated learning.

66. Harness the fact that every student in every university owns a sophisticated communications device.

67. Revolutionize mobile learning, as the iPhone has revolutionized mobile telephony.

68. Make mobile learning a revenue stream for telecommunication companies.

69. Dealing with small screens and difficult data input.

70. The worthwhileness of location-based and contextual mobile learning.

71. Difficulties to reuse the products.

72. Get students to use their mobile devices constantly also in education.

73. Learning in context.

74. Learning across contexts.

75. Self-directed learning.

76. Learning with narratives.

77. Mass-customized learning.
78. Including learners with disabilities.
79. Including learners from rural areas.
80. Developing third world countries' education.
81. Engagement of the learner.
82. Transformation of traditional education according to the needs of information society.
Annex 8: Supporting students to self-direct intentional learning projects with social media

Terje Väljataga
Tallinn University, Centre for Educational Technology/Tampere University of Technology, Estonia // terje.valjataga@tlu.ee // Tel: +372-640-9355 // Fax: +372-640-9355

Sebastian Fiedler
Centre for Social Innovation, Linke Wienzeile 246, 1150 Wien, Austria // fiedler@zsi.at // Tel: +43-1-4950442 // Fax: +43-1-495044240

ABSTRACT

In order to be able to cope with many authentic challenges in increasingly networked and technologically mediated life we need to construct opportunities for participants in higher educational settings to practice the advancement of self-directing intentional learning projects. In addition to teaching general strategies for carrying out these projects more emphasis should be put on acquiring some expertise regarding the selection and combination of a diverse set of technological means for own purposes. The various practices that are emerging around social media seem to be a promising field of experimentation in this regard. The knowledge and skills needed to select, use and connect different social media in a meaningful way form an important part of the dispositions in self-directing intentional learning projects. This paper argues for a course design in which participants are not simply engaged in developing knowledge, skills and orientations in regard to curricular subject matter and the use of technology but actively involved in self-directing intentional learning projects with the support of social media. The theoretical framework of this research is inspired by conceptual ideas developed within iCamp (http://www.icamp.eu) project. We will illustrate our line of argumentation with some empirical data collected from a pilot course taught at Tallinn University, Estonia.

Keywords
Self-directing intentional learning project, Social media, Personal learning environment, Course design

Understanding self-directing learning projects

An essential aspect of today’s post-modern, technologically rich society is to develop the ability to take control and responsibility for our own education, learning, and change. Charles Hayes has claimed that “when we fail to take control of our education we fail to take control of our lives” (Hayes, 1998). Thus, educational experiences need to be constructed in a way that provides opportunities for participants and facilitators to organize and manage their activities in technologically rich contexts. This is an essential aspect to become increasingly a self-directing person (Knowles, 1975) in today’s world.

An extensive amount of research about self-direction and related concepts (self-organization, autonomous learning etc.) exists and has produced rather heterogeneous theoretical understandings in the field of education. Most often self-direction in education is defined as “a process in which individuals take the initiative with or without the help of others, in diagnosing their learning needs, formulating learning goals, identifying human and material resources for learning, choosing and implementing appropriate learning strategies, and evaluating outcomes” (Knowles, 1975). It is important to note that self-direction does not mean the total isolation and purely individual work and activities, rather it is seen in the social context, where the “self” is influenced by others (Lindeman, 1926).

Candy (1991) offers an overview of the various strands of research that can be found under the label of self-direction in education. Activities and strategies of actors who either want to
support or execute „self-direction“ are the focus of a research perspective that Candy calls **activity-oriented**. Such an activity-oriented perspective can either be applied to actors who operate within or outside formal instructional settings. Candy (1991) suggests speaking of **autodidaxy** in reference to the latter and of **learner control** if actors operate within formal instructional settings.

The second **disposition-oriented** perspective refers to personal attributes and orientations that influence the readiness and ability of actors to self-directing learning and change projects in various contexts. Candy distinguishes between **personal autonomy** referring to the more philosophical strand of theorizing that focuses on individual freedom and **self-management** focusing on the willingness and capacity to conduct one’s own intentional change and overall education (Candy, 1991).

While we appreciate Candy’s outstanding attempt to clarify the conceptual landscape, we find the term of “learner control” for activity-oriented research on self-direction in the realm of formal instructional settings to be somewhat misleading. We thus simply speak of **self-directing intentional learning projects** (in formal educational settings) in higher education. Nevertheless, we do so from a decidedly activity-oriented perspective.

**Selecting appropriate networked tools and services for self-directing intentional learning projects in distributed settings**

It seems fair to say that most explicit attempts to foster self-direction in higher education focus exclusively on the shift of responsibility in relation to the scope, focus and depth of the subject matter studied; aspects of pacing and sequencing; and criteria and procedures of evaluation and assessment. It is usually neglected that many individuals in our increasingly networked societies, find themselves regularly collaborating within distributed activity systems (Fiedler & Pata, 2009) in which co-workers are not physically present all the time and in which activities are thus inevitably technologically mediated. Therefore, we suggest that contemporary conceptualizations of “self-direction” in education need to be expanded. Negotiating and making decisions on technological tools and services that are appropriate (or at least promising) for mediating particular activities needs to be considered as an important aspect or “expression” of self-direction in education.

Hiemstra (1994) has stated that taking personal responsibility in education refers to individuals assuming ownership for their own thoughts and actions, which does not necessarily mean control over personal circumstances or environmental conditions in all parts of life. Nevertheless, developing and fostering at least partial personal control over the technological means that mediate and support work- and study-activities seems to be an appropriate and timely educational objective for higher education. We suggest that taking initiative and responsibility for one’s own learning and change increasingly includes and requires the ability to select adequate mediating technologies to enrich a **personal learning environment**. We understand personal learning environment as a rather broad and subjectivist concept. A personal learning environment entails all the instruments, materials and human resources that an individual is aware of and has access to in the context of an educational project at a given point in time (See Fiedler & Pata 2009, for more detailed account on this aspect). Networked tools and services offer an ever-expanding variety of means to support, amplify, and enrich our personal environments for learning and change. The ability to gain access to, and choose selectively from a full range of tools, services and other resources thus needs to be considered as an important aspect and expression of self-directing intentional learning and change in education. Negotiating and selecting networked tools and services for collaborative action can help to externalize thought processes, understandings, and expectations within a group of actors that otherwise remain obscure and invisible. While trying out different tools and services actors can become aware of their thinking process, missing knowledge, and lack of understanding. Analyzing what kind of tool or service appear to be suitable for reaching a particular action goal under specific conditions, presumes participants reflecting upon perceived affordances, expectations, orientations, and so forth. In principle, this often requires a trial-error approach to find out how the potential of existing resources can be un-locked and utilized (Brockett, Hiemstra, 1991).
Learning contracts to support self-directing intentional learning projects

Though educational research has largely ignored the need to expand notions of self-directions in the direction of control over technological means to mediate one’s productive and conversational actions, it has produced some well researched and validated (cognitive) tools for fostering various aspects of self-direction in education. One such instrument is the personal learning contract (PLC). PLCs usually require some structured, written outline of what and how an individual (or group) intends to achieve within an intentional project of learning and change (Harri-Augstein, 1995). The creation, cyclical adaptation, and elaboration of PLCs is embedded in a set of conversational procedures that normally entail a facilitator and/or peers who help to clarify and explicate the essential components of such a “contract” with oneself. Thus, the main function of facilitators is to support participants drafting, refining, and revising their own contracts.

As we have mentioned above, in adult education the use of “personal learning contracts” that are embedded in a conversational coaching approach are a well documented and evaluated approach (Harri-Augstein, 1995). In its most simple form a learning contract consists of an explication of purposes that drive one’s project and that describe what one wants to achieve; a statement on strategies, explicates what activities one intends to carry out and what resources might be used; a statement regarding the desired or expected outcomes describes some criteria that would allow evaluating if or how successful a project was. What is actually carried out during the project is documented in records of action. The core statements on purpose, strategy, and outcome can, and indeed should be, revised and adjusted while the project unfolds. In a final review procedure the overall material is used to reflect and analyse the process that actually took place (Harri-Augstein, 1995).

A learning contract applied in such a way, guides an iterative process. Participants can draw parallels between their tentative plans and their actual study process and analyze the differences. This provides an opportunity to identify the direction of development and to formulate the next contract (Harri-Augstein, 1995). The main purpose of such a systematic, practical procedure is to empower participants to think positively and constructively about their study- and work skills and to be more aware of what and how they study (Harri-Augstein, 1995). Learning contracts enable individualisation and externalisation of a person’s thoughts and pursuits in respect to her goals and strategies.

In our teaching experiment we tried to make use of the general format of conversationally grounded PLCs within a landscape of social media tools and services to support the gradual shifting of the locus of control in regard to a variety of instructional functions that are normally provided by the representative of the formal educational system.

Case description

Overview of the course and landscape of social media

The Master’s level course “Self-directed learning with social media” in Tallinn University was designed to create challenging situations for participants to advance their dispositions for self-directing learning projects with the support of social media tools and services.

The course was designed and carried out by two facilitators who work at Tallinn University, Estonia, as researchers and lecturers in the field of educational technology and who are rather proficient users of social media. 26 students participated in a pilot course in autumn 2007. The background of the participants varied a lot. The majority of participants were active secondary school teachers, while the rest were full time master students, who predominantly had gathered some work experience before enrolling for the master program. However, ICT skills varied considerably among the participants. They ranged from being limited to regular use of email clients and Web browsers to high level programming skills.

The course lasted for eight weeks. In this period three full day face-to-face contact meetings were organized. The purpose of the meetings was to give an overview of the course
structure and its requirements, to provide some introductory insight to the theoretical concepts, and to provide a glance of a set of networked tools and services that participants might find useful for carrying out assignments in a distributed and mediated work setting. The remaining study activities were carried out from the distance, making use of a variety of networked tools and services.

The facilitators seeded a distributed technological landscape (see Figure 1) on the basis of social media, leaving aside any centralized and closed systems hosted by the institution (Fiedler & Kieslinger, 2006). The central core of this selection of loosely connected tools and services was a course Weblog (Wordpress), where participants were provided with an updated overview of ongoing course activities and necessary materials in the form of Weblog-posts and hyperlinks. In order to get a better overview of the participants’ progress and ongoing activities, a page was created on an open-access mash-up service (Pageflakes) to aggregate the Webfeeds of all participants’ personal Weblogs and to display all the resources that got bookmarked collaboratively on a social bookmarking service (Del.icio.us) for this course.

Furthermore, facilitators or participants could leave messages to the entire group on this mash-up service page. Synchronous communication tools like MSN messenger and Skype connected facilitators and participants for real time conversations. Beyond this pre-selected set of tools and services, participants enriched their personal landscapes with a wider selection of social media according to their individual needs and preferences.

![Figure 1. Landscape of tools and services used in the course](image)

**Course framework**

The general purpose of the course was to introduce different learning methods with the support of social media. Special attention was paid to the notion of self-directing learning and change projects. A range of techniques and assignments were given to participants with the purpose to stimulate their self-directed acts. Participants were asked to carry out two
assignments, one individually and the other one in groups. For both assignments participants had to think of an authentic work or study activity and come up with the real working technological landscape that supports this particular activity.

The idea of learning contract was used also in the course, but in a slightly modified way as described above. In order to help participants to exercise control over their study activities they were asked to draft an individual personal learning contract for each of the two assignments. Participants were recommended to explicate the core parts of their personal learning contracts each time before they were given one of the two major assignments during the course. The contracts were revised after the assignments had been carried out and then reflected more deeply at the end of the course. Reflection was done in an essay format, where participants were asked to review their personal learning contracts and their actual learning process. Participants reflected in the light of their initial plans and projections upon the actual activities they had carried out and the outcomes they finally had achieved. All individual and group work was supported and mediated by the tools and services the participants had selected. Participants were given complete freedom and full responsibility over their activities and the technological means for supporting their performance. They were encouraged to take control over both, the objectives and the means.

Although participants had been given a final deadline for assignment completion, they were encouraged to follow their own pace while respecting general organizational constraints such as the overall duration of the course. The deadlines of the assignments were meant to function as indicators for planning activities within the organizational time limits.

To foster personal responsibility, facilitators also provided a variety of alternative study resources to participants. In addition to reading materials on various related topics (self-directing learning projects, social media, collaborative learning, learning management systems, and so forth) a set of social media were introduced from which participants later could choose according to their personal needs and interests.

Research design

This research followed elements of action research (Creswell, 2002). It tried to alter a rather traditional approach to course design, while observing the effects of these changes on participants' experiences (Breakwell et al., 2000). The purpose of the research was to bring about changes in course design and the overall learning/teaching process while trying to map and understand the consequences of these changes at the same time.

The overall change process consisted of several interrelated stages:

Stage 1: Definition of the changes made in the course design The first stage of the research was to identify and describe intended changes for the course design. Here we drew from research, debates and discussions generated in the extensive body of literature based on self-direction, aspects of learning environments and use of social media in education and the ongoing work within the iCamp (http://www.iCamp.eu) project.

Stage 2: Design and implementation of remedy in the course design. The course was redesigned to incorporate the changes and carried out with master level students.

Stage 3: Observation and data collection. The third stage refers to the actual learning/teaching process, where the changes were observed and data collected.

Stage 4: Analysis of the impact of changes and reflection. This stage investigated students’ perceptions of their experiences while participating in the redesigned course.

Research questions

The purpose of this research was to determine the possibility of applying social media for fostering and promoting self-directing intentional learning projects into a master level course design and to investigate students' responses to that learning situation. This research asked
the following questions:

1. What were the challenges for the students in this kind of course design?
2. How did the students perceive the concept of self-directing their intentional learning projects?
3. To what extent support personal learning contracts the self-directing of intentional learning projects from the students’ perspective?
4. What is the role of social media while self-directing intentional learning projects from the students’ perspective?

**Data collection**

26 students participated in this experimental course. Intrapersonal data was on focus, where cognitive and emotional aspects of the students were considered (thoughts, feelings, attitudes). A direct elicitation method was used for data gathering (Breakwell et al., 2000): in the form of students’ essays about their experiences and open-ended questionnaire.

**Data Analysis**

The framework for data interpretation was based on the research questions and the changes that were brought about the course design (personal learning environment, learning contract, reflective task, social media, different role of the facilitator). Data analysis was done qualitatively with the purpose to explicate perspectives of the participants in this course, to interpret and discover patterns within the students’ accounts.

Techniques of qualitative analysis recommended by Miles and Huberman (1994) were used to analyze the data collected from the students’ essays. The analysis involved a three-step process: data reduction, data display, and conclusion drawing and verification. The analysis was done with the assistance of HyperRESEARCH, a computer-based qualitative analysis program. Data from the essays was initially coded according its set of a priori codes that were derived from the research questions together with sub-themes that emerged within these categories.

The process of coding the data is summarized in Table 1. The codes were gradually elaborated by bringing in additional themes as sub-themes while working with the data.

In addition to the students’ essays, an open-ended questionnaire was conducted after the course had finished. Open-ended questions allowed students to respond using their own vocabulary and terms to describe their expectations regarding the overall course and the role of facilitators in particular; their challenges and difficulties in this course; their understanding of self-directing intentional learning projects; their opinion about learning contracts as a means for self-directing intentional learning projects; and their previous and prospective use of social media for study or work.

The answers to the eight questions (behavioural, background and opinion/experience questions) of the paper-based questionnaire were analyzed qualitatively. Questions that referred to similar aspects were analyzed together. This was done with questions 1 and 2, 4 and 5, and 6 and 7 respectively. The analysis followed a top-down approach in which the data was categorized according to a priori codes based on the research questions and the changes made in the course design. As the number of the students and the length of their answers were not that extensive, data analysis was done manually. 24 from 26 participants returned the questionnaire.

<table>
<thead>
<tr>
<th>Analysis process</th>
<th>Rationale</th>
<th>Software</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Reduction process:</strong> cutting words and phrases that were not relevant for the current analysis; segmenting data</td>
<td>Reduction of the data in order to discard, sort and organize the data in a way that allows to collect reasonable segments of data</td>
<td>Scrivener</td>
</tr>
</tbody>
</table>

**Table 1: Steps of the data analysis**
<table>
<thead>
<tr>
<th>Coding: coding text according to <em>a priori</em> codes determined by the research questions and the changes made in the course design as well as inductive codes which emerged from the data</th>
<th>Coding of the data in order to focus, organize and process data</th>
<th>Hyper-RESEARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-coding: codes were revised and compared within the data collected from the essays as well as with the data collected from the questionnaire with the purpose to merge them into categories based on their relationships</td>
<td>Sub-coding of the data in order to abstract and condense the data for permitting conclusion drawing</td>
<td>Hyper-RESEARCH</td>
</tr>
<tr>
<td>Ordering and displaying: themes were determined and generalizations were made.</td>
<td>Taking the reduced and coded data and displaying it in an organized, compressed way through such means as selection, paraphrasing and subsuming in a larger pattern that permits conclusion drawing</td>
<td>Hyper-RESEARCH</td>
</tr>
<tr>
<td>Conclusion drawing: conclusions were made and written up</td>
<td>Decisions about the meaning of data and testing validity of findings by noting regularities, differences and similarities, explanations and propositions were made</td>
<td>Scrivener</td>
</tr>
<tr>
<td>Verifying: conclusions were verified by referring back to the original data</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The overall data analysis was initially done separately for the essay and the questionnaires. After the coding system had begun to consolidate, the data from both instruments was merged.

**Results**

**What were the challenges for the students in this kind of course design?**

One of the purposes of this course design was to create challenging and authentic situations for the students regarding the technological support of their own learning environments and taking initiative and responsibility for their activities.

Although it was mentioned in some students’ essays that they had experienced some kind of problems, the questionnaire results showed that 5 students out of 24 claimed they had had no major difficulties and challenges in this course. However the main challenges were the following:

1. Challenges related to terminology
2. Challenges related to learning contracts
3. Challenges related to assignments (individual and group work)
4. Challenges related to tools and services

**Challenges related to terminology**

The analysis of the questionnaire and essays showed that 11 students experienced information overload in the beginning phase of the course. Insufficient explanation about new terms and concepts resulted in some students feeling frustrated. They claimed they had received too much new information and new terms at the same time, which made it complicated to make sense out of this oversupply. The students said: „The terminology was new and not understandable for me“ „In the beginning I thought I need a dictionary, because so many new terms were mentioned“

It is ineluctable that the introduction of new concepts, new tools and services as well as an
unusual course design carries terminology that is not necessarily familiar to the students. One student found the whole course design with its activities quite challenging while two other students claimed that they had only experienced problems understanding the concept of self-directing* their learning projects. The students said:

„The subject and the structure of the course were new to me“
„In the beginning the whole course appeared like rocket science“

*Challenges related to learning contracts*

Challenges related to the learning contracts refer mainly to the early stages of formulating the different parts of such a contract. Seven students claimed they had experienced problems trying to explicate their goals, strategies, tools and evaluation criteria. The students said:

„It was complicated to formulate the evaluation criteria. My whole life others have decided that“
All of the sudden I had to think through my whole learning process — what exactly am I going to do now? What do I want to achieve?“
Very few students had heard about learning contracts and for some of them it was unclear why it had to be done, especially since it did not form a major part of the final grading. The student said:
I didn’t understand the need for creating a learning contract... why do I have to do this?“

It was obvious that most students were not ready to take initiative and responsibility for their own learning. The main reason seemed to be a lack of experiences and rationale in this regard.

*Challenges related to assignments*

19 students out of 26 claimed that the first assignment was difficult, confusing and unclear. The main reason for this appears to be the new terminology used by the facilitators and the unfamiliar distributed course environment.

While the first assignment was confusing for most students, the second assignment that focused on group work was readily understood. However, some other challenges occurred.

Eleven students claimed that group work on the distance was complicated because of unreliable technological tools and services, such as synchronous editing of web-based documents or schemes. The students found it very time consuming to find a common understanding among group members, to communicate, and to regulate the group’s activities without meeting others face to face. The students said:

„Group work was especially difficult for me as I had never done it in this way“
„I must say that it is very difficult and time consuming to carry out group work on the distance“

Four students found it difficult to find common time frames for the group work to discuss issues synchronously. The students said:

„It is very difficult to find a common time frame that is suitable for everybody“
„Unfortunately the members of the group had so different time schedules and therefore taking common action was rather limited“.  

However it is interesting to note that the students did not encounter problems choosing the right tools and services for carrying out their group assignments on the distance.

In addition, a couple of students considered it unusual presenting home assignments in their personal Weblogs in a format that was public and easily accessible. The student said:
„But the idea to put everything to my Weblog didn’t put a smile on my face“
Challenges related to tools and services

The questionnaires and essays showed that all the students were familiar with e-mail services prior to the course. Some of the students had had experiences with video- and photo-repositories as well as Weblogs. Weblogs had been mainly used in other course settings. However the majority of the students were unfamiliar with the tools and services introduced and used during the course (see Table 2). Not surprisingly the main challenges that emerged from the learning process were related to new tools and services as well as the learning environment as a whole.

Table 2: Tool use before and after the course

<table>
<thead>
<tr>
<th>Type of tools and services</th>
<th>Tool use before the course</th>
<th>Potential tool use after the course</th>
<th>Possible explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social bookmarking (Delicious)</td>
<td>6</td>
<td>11</td>
<td>This tool was part of the course environment and the students were obliged to create an account in order to be able to find their Weblogs</td>
</tr>
<tr>
<td>Video repositories (Youtube, Google video)</td>
<td>10</td>
<td>1</td>
<td>Due to the prior use of these tools and services the students might forget to mention them again</td>
</tr>
<tr>
<td>Photo repository (Flickr)</td>
<td>7</td>
<td>0</td>
<td>Weblogs were used in the course as one of the obligatory tools</td>
</tr>
<tr>
<td>Weblog (Wordpress, Blogspot)</td>
<td>9</td>
<td>17</td>
<td>An aggregator was part of the course environment</td>
</tr>
<tr>
<td>Collaborative writing and drawing (Google docs, Gliffy, Vyew, Bubble)</td>
<td>4</td>
<td>9</td>
<td>Their applicability were tested and proved while doing group work</td>
</tr>
<tr>
<td>Web-based office (MS Office)</td>
<td>2</td>
<td>3</td>
<td>These tools were not directly used in the course, neither by facilitators nor the majority of students</td>
</tr>
<tr>
<td>Wiki (Pbwiki)</td>
<td>1</td>
<td>4</td>
<td>This tool was used by the facilitators during the face to face meetings</td>
</tr>
<tr>
<td>Presentation repository (Slideshare)</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

Six students from 24 claimed that they had difficulties with the large number of different tools and services introduced during the course. This meant it was challenging to keep up with understanding the purpose and use of these tools and services. Furthermore, the students considered the registration processes, getting oneself acquainted with tools and services, as well as remembering all the login details (mentioned by five students) as challenging. The biggest challenges for two students were finding and combining tools and services, and getting an overview of the overall landscape of tools.

Furthermore, two students claimed they had had not enough time to go through all the tools and services in depth, since their level of interest in the tools and services and the overall course pace was not aligned.

In conclusion, the main challenges for students in this course were related to the assignments they had to carry out, since they differed from the type of assignments usually found within traditional course designs. Another major area of challenge was the array of technological tools and services used to support personal learning environments.

How did the students perceive the concept of self-directing their intentional learning projects?

To answer the second research question, the questionnaire was designed to capture the students’ insights and opinions regarding the notion of self-directing intentional learning projects. The questions aimed at eliciting positive and negative aspects of self-directing
intentional learning projects from the students’ perspective.

Ten students pointed out that the most important aspect of self-directing one’s own learning is the freedom to choose one’s goals, strategies, means, and resources. It is interesting to note that one student focused on the importance of self, self-consciousness and responsibility. The student said:

„When the self and my purposes become more important than what the others expect from me...“

Four students understood the concept of self-directing intentional learning as a constant and conscious development based on one’s intrinsic motivation. Three students thought that it is about defining one’s needs and interests, and ways how to learn accordingly. The students said:

„Keeping a diary, which means organizing oneself “
„Motivating oneself, reflecting on one's activities and outcomes and defining future direction“
„Analyzing failures and success of one’s activities“

The positive aspects regarding the concept of self-directing intentional learning projects were the following: ten students thought that the freedom to choose is the main advantage. This included the possibility to plan one’s learning process and topics based on one’s needs and interests, the possibility to choose resources, and to follow one’s own pace. Furthermore, the possibility to combine work, study and home were considered equally important.

However, the essays showed that too much freedom and lack of structure can create chaos and can be seen also as an inhibiting aspect for the learning process. Setting up one’s goals independently can be seen as positive and negative, since most students experienced that it is not that easy to clearly define one’s goals. Some of the negative aspects that were brought up by four students were motivation and responsibility while carrying out self-directing intentional learning project.

Three students pointed out that the unlimited opportunities to acquire new knowledge and information were a positive aspect. One interesting aspect brought out by one student was the following:

„I have to think how to make the best out of the resources that I have at a certain point in time“

Three students thought that the lack of feedback and lack of others support were negative aspects. Wavering from the initial goals, lack of self-belief and fear of failure were mentioned by four students. One student considered the misunderstanding of the assignment as a drawback. Another student was afraid of making wrong choices in terms of supporting tools and services

To what extent support personal learning contracts students who are self-directing their intentional learning projects?

In general 21 students from 24 found learning contracts useful and supportive for self-directing their intentional learning projects. Students gave quite different answers in regard to the usefulness of the learning contracts in one’s learning process. The following aspects were pointed out:

- 20 students said that it was needed to document activities since it gave a clear overview of what, when and why.
- one student claimed that it offered a clear structure for learning and it helped to develop knowledge
- ...it became personal and it motivated
- ...it supported the achievement of better goals and outcomes
- two students thought that it helped the evaluation of development and the measurement of one’s achievements
- two students claimed that it helped to concentrate and choose the focus point, and to coordinate and direct one’s activities
- ...it helped to determine what the important activities were
Two students claimed that the learning contracts did not make sense for them and were useless. The students said: „This is just an expression of forming one’s thoughts and goals“ However one of them saw the general importance for other learners as a support for concentrating and focusing on important aspects in one’s learning process.

**What is the role of social media while self-directing intentional learning projects from the students’ perspective?**

As it is seen from the table above (Table 2), the perceived probability of using social media in the future was quite high. 18 students claimed that they intend to make use of some of these tools and services in the future to support their leisure, study or work related activities. Some of the students reported that this course encouraged them to investigate different social media. The main reason for a continued use and exploration of new tools and services is the perceived simplicity and variety of choices and functionalities found in social media. Social media enable for example to provide continuous feedback, to carry out reflective tasks, to draw schemes either individually or in groups synchronously, to mediate group communication, and to work together on common artefact. The students said: „Social media give me an opportunity to carry out my tasks despite of the location and the nature of the tasks“ „Social media applications are very effective, not only for the execution of self-directing intentional learning projects but also useful in very different context and conditions“ Despite of these perceived benefits the students wished to receive more supervision and practice for using different social media. From their perspective too many tools and services were introduced in a rather short period of time. Due to the limited time frame they found that their understanding of the tools’ nature and their ability for using them remained somewhat superficial.

**Discussion**

The findings suggest that our re-designed course indeed created some challenging situations for students. The personal reflections on the activities varied a lot among students. They ranged from an experienced match of the activities with predefined personal goals and strategies, to explicit dissatisfaction. A major source of dissatisfaction with the process and outcome appeared to be a feeling of information overload in the beginning of the course and confusion around the given assignments and terminology. Quite many students claimed that in the end of the course they were able to write down their goals and how to reach them. On one hand, this is a fairly common pattern for novices who try to explore a new domain of knowledge and skill, on the other hand this might indicate that the instructions for the assignments were not clear enough. Some students found it hard to obtain an overall image of the course and its specific assignments and components

The students predominantly claimed that they had progressed a lot during the course and that this had changed their initial goals and understanding. They understood that the social media they worked with are also applicable in other contexts beyond educational settings, since these tools and services afford to carry out many different activities.

The essays showed that most of the students were rather self-critical in their evaluation of their capacity to identify their own needs, to develop personal learning contracts to meet these needs, and to achieve the goals that they had described within the contracts. It was obvious that the externalization of one’s own thoughts and strategies was something rather novel and challenging for all students, since most of the courses in higher education largely ignore students’ own learning goals and personal learning environments. In our re-designed course students found the personal learning contract procedure very useful. One reason for its perceived usefulness seemed to be the instrumental value of the learning contract material for writing a reflective essay on their overall experience and process at the end of the course period. The personal learning contract procedure was
described as a means and tool to keep them on track, to structure their own activities, and to monitor their success. Furthermore, students considered the learning contract procedure as a good way of documenting their ideas and thoughts, coordinating their activities and providing means to measure their achievements. They interpreted the learning contract as their own personally constructed “instruction” or guideline for their activities. Some of the students found it motivating, as they were recommended to write down in detail what, how and when to work and study. This provided them with a compact overview of the direction to pursue and the rationale behind their activities. Some students reported that the contracts helped them to concentrate and focus on their activities and not to deviate from the initiated path. Writing down goals and needs turned them into something personal and important for the author, giving her a feeling she was in control of her own activities.

After completion of the assignment many students understood that their predefined evaluation criteria did not make sense and were not really measurable. Furthermore, evaluating themselves seemed to be a new and rather challenging task. It was obvious that the first assignment was quite time-consuming and demanding, due to numerous new terms and social media. This certainly added to the initial problems that students experienced while they were trying to specify detailed goals and strategies. It seemed easier for most students to draft their second learning contract in relation to the group assignment, after they had explicated their goals and purposes, strategies, and intended outcomes already once before in the context of the individual assignment.

However, it appears that different tools and services and new concepts and terms should be gradually introduced over the course period. All students claimed that they had a rather positive experience regarding the acquisition of useful theoretical knowledge and practical skills in respect to the use of social media tools and services and self-directing their own learning projects within formal educational settings and beyond.

**Conclusions**

It is obvious that outside of formal educational settings individuals (and groups) cannot rely on educational authorities and formal instructional systems to structure and support their activities. We assume that formal education for adults should be designed in ways that allow all students to actually execute and advance their dispositions for self-directing intentional learning projects in general, and within distributed and net-worked settings in particular.

This paper described a redesign of a master’s level course called “Self-directed learning with social media” that intended to foster the ability of students to self-direct intentional learning projects in distributed settings. Thus, a significant aspect of this course design was the provision of opportunities to practice the selection of social media for mediating particular activities.

The reflective essays of the participants on their individual learning processes and the administered questionnaires were analyzed to gain an insight into how students experienced their own ability and effectiveness to plan, organize, and manage their own work- and study-activities. The reports of the students and the digital traces of their activities showed quite clearly that they gained considerable knowledge and skills regarding the use of social media for supporting a range of activities. It can be concluded that students indeed acquired some expertise regarding the selection and meaningful combination of a diverse set of social media for their own purposes.

However, it is important to note that rather isolated and short-lived interventions that are constrained by the academic semester rhythm make it difficult to observe any significant changes of students’ readiness and capacity for self-directing their own learning and change. Nevertheless, we believe that the course design presented in this paper offers a promising and feasible approach to foster the advancement of a set of dispositions (knowledge, skills, orientations) for self-directing intentional learning projects in distributed settings that are viable for coping with many authentic (educational) challenges in today’s increasingly
networked and mediated life.

References


Brockett (Eds.), Overcoming resistance to self-direction in adult learning, San Francisco: Jossey-Bass, 5-12.


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.