

Visualizations of User Data in a Social Media Enhanced Web-based Environment in Higher Education

<http://dx.doi.org/10.3991/ijet.v8iS2.2740>

K. Silius, A.-M. and M. Kailanto

Tampere University of Technology, Tampere, Finland

Abstract—In this article, we introduce the results of a study concerning user data visualizations in the context of the TUT Circle, a social media enhanced web service for learning, networking, and communication at the Tampere University of Technology (TUT). Well-designed visualizations can provide valuable insight into student activity and participation in an online course, and therefore support both teachers' and students' strategic decision making. We found social network analytics and visualizations useful in providing information to improve the pedagogical practices of online courses, as well as the students' overall engagement in their studies and study lives. However, to be useful, the visualizations need to have enough hints so that the users can interpret the different elements in them in the right way.

Index Terms—Learning, learning analytics, social network services, visualization of user data.

I. INTRODUCTION

Social media enhanced web environments have been a topic of increasing interest in recent years. Also, at the Intelligent Information Systems Laboratory (IISLab) at Tampere University of Technology (TUT), we have studied the use of social media enhanced environments in a higher educational context from several different viewpoints, such as the quality factors of social media enhanced web-based environments, students' motivation to use that kind of environment, the possibilities of supporting peer-to-peer learning, and networking among hypermedia students, as well as the collaboration between students with two different academic backgrounds [1] [2] [3] [4] [5] [6]. The results of these earlier studies were utilized to develop the TUT Circle¹, a content management framework built on Drupal.² The TUT Circle contains all the basic functionalities of a modern social media web service, including the possibility to publish several types of content, e.g. wiki pages, blog posts, news items, and events. The TUT Circle promotes the networking, collaboration, and communication of small groups by allowing users to form friendships with each other, send private messages, chat, create, contribute, and comment on content, exchange opinions, and share resources within and between the groups. When compared to the use of third-party social media platforms, the use of the TUT Circle, that is maintained by its developers, enables the collection

of rich data about students' activity and participation, as well as the development of new features based on these findings.

The development related to the analytics of the log data, collected from the TUT Circle, has been done by implementing tailored scripts and batch processes for exporting data directly from the TUT Circle database for further analysis [4]. Currently, the IISLab is studying and developing user data analysis and visualization possibilities. The target is to develop useful tools for teachers and students into the TUT Circle, to monitor and visualize the users' actions in an appropriate way. The aim of this article is to describe our recent research related to visualization. The research questions are: "How do students interpret different visualizations in a higher educational context," "What kinds of visualizations are considered useful," and "How much explanation is needed (if any) to help to interpret the different elements (lines, colors, arrows, etc.) in visualization?"

This study is part of a Finnish national project called the Campus Conexus, which is being carried out from 2009 to 2013, and is financed by the European Social Fund. Five Finnish universities are participating in the project with the aim to study how students engage in university studies, by enriching learning experiences and enhancing the students' personal, intellectual, collegial, and professional development [1]. Another aim of the project is to develop a social media enhanced web-based environment to support the students' collaboration, networking, and studying, which the IISLab is responsible for.

II. VISUALIZING INFORMATION

The use of web-based learning environments provides an opportunity to collect an enormous amount of data on students' communicative, collaborative, and participative actions. The main challenge for utilizing the tracking data collected is the quantity of data. Large amounts of data can be difficult for people to analyze, interpret, and understand, because processing data requires complex cognitive processes. However, if the data is packed and presented in correct, clear, and effective ways, humans can process an incredible amount of data. Furthermore, if the data presented is difficult to perceive and process, visual representations are critical [7] [8].

The visualizations can help users to understand data better. However, to be understood, the data used should be familiar and interesting to the users. Visualization can help by increasing the number of memory and processing resources available, reducing the search for information,

1 <http://www.tut.fi/piiri>
2 <http://drupal.org/>

enhancing the recognition of patterns, enabling perceptual inference operations, and using perceptual attention mechanisms for monitoring. From the users' perspective, a good visualization will confirm what they already know and show them several insights into the data [9].

Usually, the designers of visualizations lean on things like the Gestalt theory of psychology in which pattern recognition is based on proximity, similarity, continuity, symmetry, closure, and relative size. This means that things close together, similar elements, or smoothly connected elements are perceptually grouped together, and also that symmetrically arranged visual elements are often perceived as a whole. The principles of relative size mean that smaller components tend to be perceived as objects, whereas large ones are a background. These kinds of basic manners used to recognize visual patterns make it easier to design a high quality visualization. Still, there are many kinds of differences among users in understanding the visualizations of large data. People's unique personal features, like cognitive capacities, motivational and cultural aspects, earlier experiences, as well as contextual factors, affect how they interpret visualizations and generate conclusions [8] [10].

According to the reference [11], disadvantages of visualization can be formed by the designer or the user. The designer can intentionally or unintentionally introduce mistakes in the visualization, or the user can make a wrong interpretation of it. Reference [12] also found out that there are three types of disadvantageous effects: cognitive, emotional, and social ones. Cognitive disadvantageous effects occur when the visualization distracts a person from the main goal, or when several items in a visualization are emphasized at the same time and confuse the user about where to focus. Emotional disadvantage occurs when some images cause emotional harm to the user because of their shocking or repellent content. Social disadvantage occurs when cultural and cross-cultural differences are induced by the heterogeneity of the users. For example, the meanings of symbols and colors are not universal [10] [12].

However, there exist mutual perceptual and cognitive processes, which help to design visualizations that can be interpreted generally. People have similar perceptual and cognitive processes, which means that data visualizations can help people with different backgrounds and knowledge to more easily get the idea of what the data has to say, especially because the information processing is externalized for perception rather than for cognition [8] [10]. People can be supported to assimilate information by providing a visualization with a reasonably clear purpose, that offers a suitable amount of information at the right time in a coherent way, avoiding the presentation of redundant information or distorting what the data has to say, revealing several (but not too many) levels of detail, and providing a broad overview of the fine structure [11].

III. EVALUATING DIFFERENT TYPES OF VISUALIZATIONS

In order to make network actions visible by developing metrics and visual tools for students and teachers to use in the TUT Circle, a pre-survey was needed. The IISLab organized a research study in the autumn of 2012 in order to evaluate the usefulness of different types of visualizations. In a preliminary research study organized as a part

of a workshop on visualizations in August 2012, ten visualizations (five static and five interactive) were given to 17 people, who represented different fields related to university education, e.g. coordinators, education developers, lecturers, assistants, and experts in online pedagogy. The study presented different types of visualizations such as networks, diagrams, and hierarchical presentations. The static visualizations were presented on paper, and the interactive visualizations were given on the laptop computer so that the participants could interact with them. The context of the visualizations was defined quite openly, if defined at all. The aim of the preliminary research was to see how people were able to interpret visualizations with little or no knowledge of the context, and gather ideas about where they could use them in a higher educational context. The ideas and the comments about each visualization were reviewed together in the workshop. The participants were also asked to evaluate how useful they thought each of the visualizations was in a higher educational context, by choosing a rating from the 7-level semantic differential that is the bipolar rating scale, with contrasting adjectives at each end (0=totally useless/ 6=very useful). Seven out of the ten visualizations were considered useful (the average points earned were over half of the maximum 6 points). The three visualizations that received lesser points had in common that the context was weakly presented in the visualization. The main results of the first phase were that more information about the context was needed to better evaluate the usefulness, and also help was needed with how they should interpret the different elements (such as line thickness, colors, arrows, etc.) in a certain visualization. The results of this preliminary research were used as a base for the more profound research that was done with the hypermedia students in the autumn of 2012.

In the second phase the data was collected from hypermedia students, regarded as early adopters in different hypermedia courses that took place in the autumn of 2012. The visualizations that were given to the 42 students were slightly altered from the ones that were used in the first phase, based on the feedback that was received. For example, the context for each visualization was presented more clearly to the students by providing a narrative context (i.e. by telling a short story of what each visualization is about). The information about the students' opinions and conceptions of the usefulness of the visualizations was gathered by using a web questionnaire where they could explore the ten visualizations. The students were asked how they interpreted the different visualizations in a higher educational context, what kind of visualizations were considered useful, if more explanation was needed to help to interpret the different elements in the visualization, and where the visualizations could be used in a higher educational context. The students were also asked to evaluate the usefulness of each visualization in a higher educational context by choosing a rating from the 7-level semantic differential similar to that used in the preliminary research.

The qualitative data from the open-ended questions was analyzed using content analysis, a method used to examine the meanings, themes, and patterns that were manifest or latent in the students' answers, and to identify the similarities and differences among answers. To ensure the validity of the analyses, researcher triangulation was used. It also has to be noticed that the semantic differential scale is an ordinal scale, and consequently the averages and

standard deviations do not have valid interpretations. However, they can still offer useful information for constructing an overall view of the phenomenon under study.

Five static and five interactive visualizations represent:

1. Network diagram. The network diagram represents the relationships between the courses available at the TUT based on the prerequisites. The area of the circle corresponds to the number of courses that require this course as a prior condition, and the color of the edge indicates if the prerequisite is mandatory or recommended.

2. Bubble chart. The bubble chart here represents course content. The area of the circle corresponds to the amount of content related to certain topics, and the color corresponds to the categories of the topics. The positions of the bubbles do not mean anything.

3. Circle packing. Circle packing is a method for visualizing a large amount of hierarchically structured data. Here the visualization represents the topics that the student groups have covered in their project works.

4. One-mode student-to-student network diagram. The diagram represents the relationships between the students based on the number of times a student had read or commented on the content produced by another student. The color corresponds to the group the student belongs to.

5. The activity of students' group in the TUT Circle is presented in three different ways: a) **Classic bar chart**, where the color corresponds to the group. b) **Star plot**, where each star represents one group. c) **List of faces** that represent the most active members of the groups.

6. Interactive histogram. The students were asked to imagine that the interactive histogram represents the activity of the students' group in the TUT Circle. The visualization allowed the students to explore their activity and participation history by using live histograms. The students could click and drag on the charts to filter the information by dimensions like time, unique viewers (i.e. number of single persons) per content, total views per content, and date.

7. Interactive Bundle-vision. The students were asked to imagine that the interactive Bundle-visualization represented the relationships between the topics of the learning material, and the topics of the additional learning materials. The users can investigate individual curves within the bundle by increasing the bundling strength in a highly cluttered display, and selecting an interesting curve in the bundle while the selected curves remain visible. Decreasing the bundling strength makes the individual curves related to the selected bundle visible for further investigation [13].

8. Matrix diagram. The students were asked to imagine that the matrix diagram visualizes the students' collaboration during the course. Each colored cell represents two students appearing to work together during the course. The darker cells indicate students that collaborate more frequently. The user can reorder the matrix for further investigation.

9. Narrative visualization. The students were asked to imagine the narrative visualization like an interactive slideshow that tells a story about the student' activity during the course. The structure of the visualizations follows a typical slideshow format, but allows the user to investigate a particular point of the presentation before moving ahead [14].

10. Interactive network diagram. The students were asked to imagine that the interactive network diagram represents how informal student associations, like students' clubs and guilds, are interconnected through the members' (i.e. associations are connected if they share members). The user can zoom the visualization for seeing a smaller subset of the data, and filter for seeing a subset based on the values of the following factors: most recently updated, most popular, and most connections.

TABLE I.
THE DIVISION PERCENT OF THE STUDENTS ACCORDING TO THEIR RATINGS ON USEFULNESS OF THE VISUALIZATIONS (N=42)

Visualization	Type	A	B	C	Av.	Stdv.
1. Network Diagram	static	49	19	32	2.53	1.53
2. Bubble chart	static	51	21	28	2.21	1.75
3. Circle packing	static	60	21	19	2.00	1.60
4. One-mode student to student network diagram	static	43	33	22	2.29	1.66
5.a Classic bar chart	static	10	19	71	4.21	1.34
5.b Star plot	static	14	17	69	4.00	1.43
5.c List of faces	static	79	14	7	1.60	1.23
6. Interactive histogram	interactive	24	29	48	3.43	1.45
7. Interactive Bundle - vision	interactive	43	19	38	2.76	1.59
8. Matrix diagram	interactive	38	24	38	3.00	1.55
9. Narrative visualization	interactive	17	17	66	3.98	1.51
10. Interactive network diagram	interactive	34	32	34	2.88	1.33

A=useless B=neutral C=useful

IV. RESULTS OF SURVEY

The overall results of the research were that the students usually liked the visualizations representing the activity of the users (except 5c, the face list visualization, which was useless in their opinion), while the visualizations representing the structure of the contents or the relationships between them were not that useful. In general, the students found the interactive visualizations to be more useful than the static ones, except the bar chart visualization representing the students activity in the TUT Circle.

A. The most useful visualizations according the students

For developing a high quality visualization tool into TUT Circle we highlighted the most relevant visualizations from the students' point of view. Those are: 5. activity visualizations a) bar chart and b) star plot, 6. interactive histogram and 9. narrative visualization (See Table 1).

The most useful visualizations according the students' opinions were the activity visualizations like classic bar chart and star plot (Fig. 1). Seventy one per cent of the students thought that the bar chart visualization was useful (see #5a in Table 1). They mentioned that the most active group was easy to identify, and it was also easy to evaluate activity levels of each group. One out of ten (10 %) students considered bar chart as useless. There were no differences in interpretation of the visualization among the students in relation to their impression of usefulness of the bar chart visualization. All the students (41) interpreted

that visualization represent the activity of the five student groups.

“Traditional, you can see of it who have been active.” - Respondent 41³

Little bit more than two thirds of the students (69 %) found the star plot visualizations as useful (Fig. 1). The students mentioned that visualization represented clearly a division of the activity of each group during the year. They also mentioned that it was easy to see in which months did each group was most active, but it was not easy to interpret which group was the most active during the year. The 14 % of the students, who thought the visualization was useless, mentioned that the visualization was suitable for evaluating the activity of each student groups. They did not mentioned possibilities for comparing groups' activity or evaluating group activities in different months in contrast to those students who considered the star plot visualization as useful (See #5b in Table 1).

“You can see better than from bar chart monthly activity of the groups. Although interpreting is harder.” - Respondent 24

According the students bar chart and star plot visualizations can be used for different purposes, for example, for describing performance of the student groups (e.g. evaluating levels of student groups activity, number of content produced or read) or comparing the students group according their activity levels.

“To make peer stress. To follow progression of exercise. To define activity.” - Respondent 14

“To supervise if every group aren't working as much.” - Respondent 30

“To consider rating. To consider if members of group aren't working as much.” - Respondent 40

Two thirds of the students (66 %) considered the narrative visualization (Fig. 2) to be useful (See #9 in Table 1). According the students the way how this narrative visualization represented information was joyful, interesting and easy to understand. The visualization also offered a good overview of the evolution of students' activity during the course. For the teachers this kind of visualizations could be useful for monitoring and evaluating students' participation, and further, improve the instructional design of the course. The students also thought that the visualizations could offer them motivating information about the activity and participation of individual students.

“In TUT Circle most of students red unique materials and in total they red 3600 pages. Twenty percent of these 3600 pages were ridden by five students. For example student who red most, red 255 materials when mean was 56 and median 60.” - Respondent 34

“... if you could see your own position among the others, you could try to reach the top.” - Respondent 6

However, almost one fifth (17 %) of the students found the narrative visualization as useless. Some of them thought that information was not interesting at all or they did not understand what was the story the narrative visualization supposed to be represent. Two of the students mentioned that visualization made them feel guilty for laziness.

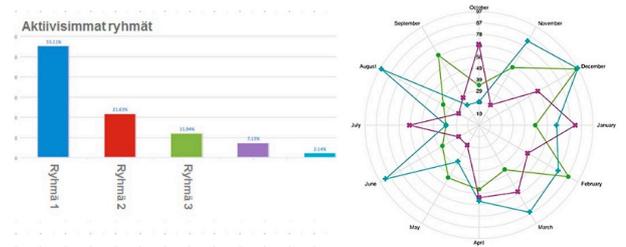


Figure 1. Activity visualizations a. bar chart and b. star plot.

These are the **people** who read resources related to the course on TUT Piiri.

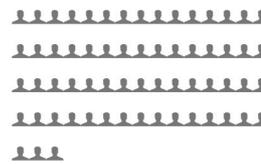


Figure 2. Interactive narrative visualization telling a story based on data.

“Once upon a time there was a course. At course there were lazy people and also people not that much lazy. Mean of the graph is to accuse lazy people by representing workload of not-lazy people and by showing how active people are on average.” - Respondent 32

The students were also asked to what kind of a story the visualization was presenting. Most of the students (39/42) had interpreted that the narrative visualization told a story with data about the activity and participation of individual students and evolution of them during the course with a purpose to motivate students to evaluate their own activity and participation and to make strategic decision for improving their performance, if needed.

“Once upon a time there was a course in which were students and a few lecturers. At the course were a few hyperactive students and one of them was a superman. If that was not you, shame on you!” - Respondent 28

Almost half of the students (48 %) found the interactive histogram visualization (Fig. 3) to be useful (See #7 in Table 1). According the students this kind of visualization that allowed users to filter information according the time, unique viewers, views and date for further investigation could be useful for monitoring participation and activity of the students during the course. The students mentioned also that visualization could offer useful information about workload and use of time for improving a schedule of the course.

One fourth of the students (24 %) considered the interactive histogram visualization to be useless. Some of them thought that visualization could be useful in some way for monitoring performance of the group, while some could not find any meaningful use for the visualization. Most of the students (37/41) interpret that the visualization represent students participation during the course and offered possibility to investigate daily distribution of the performance and activity of the students.

3 The responses of participants were originally in Finnish, and they have been translated into English.

“To perceive the meaning of deadlines, to perceive if schedules of courses are even/uneven.” -Respondent 12

“To visualize/foresee the peak of workload. How students for example are doing their exercises (are everything done at the last moment etc.)” -Respondent 32

B. The most useless visualizations according the students

For avoiding poor design of visualization the most useless visualization were also analyzed more closely. Those were: 5 c. List of faces, 6. Circle packing and 2. Bubble chart (See Table 1).

The most useless visualization according the students was list of faces (i.e. pictures of the students) (Fig. 4), that majority of the students (79 %) considered as useless (See #5c in Table 1). The students thought that list of faces was presenting the most active students, but they did not considered that information as important or they did not understand what was to purpose of that list.

“Users of TUT Circle, relation to activity is unclear. Obviously all the most active ones?” -Respondent 15

Only 7 % of the students found the visualization useful. They seemed to think that any information about activity or participation of the students was useful in general. According the students the appropriate use for this kind of visualization was representing the most active students of the course or the students who were online.

“The latest log in users.”-Respondent 39

Circle packing visualization (Fig. 5) was useless according to over half of the students (60%). The students mentioned that they did not understand the visualization at all or the visualization was unclear and difficult to interpret without any explanatory texts. Some of these students thought that nested circles of different sizes represented division of the content topics into subcategories, but they could not find any appropriate use for them. However, one fifth of the students (19%) considered the circle packing visualization as useful, for example, for presenting the course topics (See #3 in Table 1).

“There are so many different kind of circles, so isn't easy to answer to this question.” -Respondent 7.

According the students the circle packing visualization could be useful for representing how the topics of course, learning material or project work etc. form entireties and how they distribute into subcategories. With the visualizations a teacher could also demonstrate to students what subjects have already dealt with. However, 12 students out of 41 did not find any appropriate use for the circle packing visualization.

“You can at least figure out that smaller circles include themes which small groups are talking about. It is still unclear if blue circles containing many blue circles are also small groups or is it wanted to collect together exercises of similar themes in visualization. And is there in visualization a teamwork, which would include only one theme talked about? How that could be marked? It could be also so that in a visualization is not divided themes according to groups but according to content of themes, when it could be also marked if theme was talked about or not. Necessarily it was not anyway marked. So in the end you can't figure out almost nothing at all.” -Respondent 25

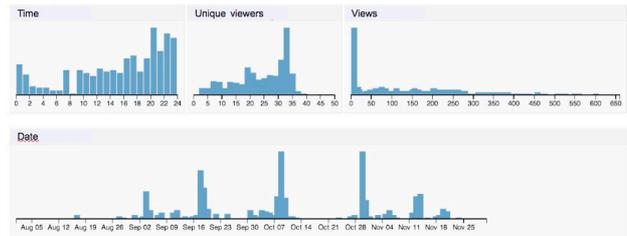


Figure 3. Interactive histogram visualization allows users to filter information according the time, unique viewers, views and date.



Figure 4. List of faces (i.e. pictures of the students).

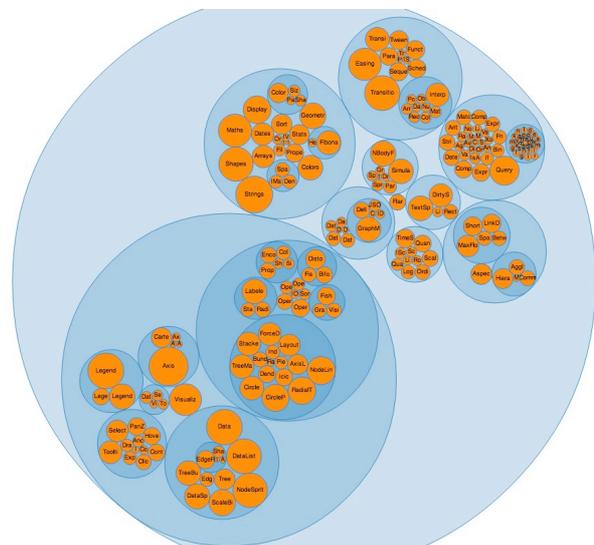


Figure 5. Circle packing visualizations. Source: <http://bl.ocks.org/4063530>

Bubble chart visualization (Fig. 6) was considered as useless by half of the students (51 %), while 28 % of them found the visualizations to be useful (See #2 in Table 1). Most of the student (30 /41) thought that visualization in Fig. 6 was representing the course topics. According to some of them the visualizations showed core issues of the course. However, some students thought that the visualization was difficult to interpret if you were not familiar with the subject. Some of them mentioned that there was

- ing," *Int. J. of Emerging Technologies in Learning*, vol. 7, issue 3, pp. 35-42, 2012. <http://dx.doi.org/10.3991/ijet.v7i3.2173>
- [6] A.-M. Tervakari, J. Marttila, M. Kailanto, J. Huhtamäki, J. Koro, and K. Silius, "Developing Learning Analytics for TUT Circle," in *Proc. of Open and Social Technologies for Networked Learning (OST'12)*, 30. July – 3. August 2012 in Tallinn, Estonia.
- [7] Elias, T., *Learning Analytics: Definitions, Processes, and Potential*. <http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf>
- [8] M. Berg, *Human abilities to perceive, understand, and manage multidimensional information with visualizations*. Doctoral Dissertations 30/2012. Helsinki: Unigrafia Oy, 2012. <http://lib.tkk.fi/Diss/2012/isbn9789526045498/isbn9789526045498.pdf>
- [9] J.-D. Fekete, J.J. van Wijk, J.T. Stasko, and C. North, "The Value of Information Visualization," in *Information Visualization - Human-Centered Issues and Perspectives*. LNCS 4950, A. Kerren, J.T. Stasko, J.-D. Fekete, C. North, Eds. Springer, 2008, pp. 1-18.
- [10] S. Bresciani and M. J. Eppler, *The Risks of Visualization - A Classification of Disadvantages Associated with Graphic Representations of Information*. ICA Working Paper # 1/2008. University of Lugano, February 2008. <http://www.knowledge-communication.org/pdf/bresciani-eppler-risks-visualization-wpaper-08.pdf>
- [11] E. Tufte, *The visual display of quantitative information*. 2nd edition. Cheshire, Connecticut: Graphics Press, 2001.
- [12] P. Buergi, and J. Roos, "Images of Strategy," *European Management J.*, vol. 21, no. 1, pp. 69-78, 2003.
- [13] D. Holten, "Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data," *IEEE Transaction on Visualizations and Computer Graphics*, vol. 12, no. 5, 2006.
- [14] E. Segel and J. Heer, "Narrative Visualization: Telling Stories with Data," *J. IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1139-1148, November, 2010. doi: <http://dx.doi.org/10.1109/TVCG.2010.1>

AUTHORS

K. Silius is with Intelligent Information Systems Laboratory, Department of Mathematics at Tampere University of Technology, Tampere, Finland (e-mail: kirsi.silius@tut.fi).

A.- M. Tervakari is with Intelligent Information Systems Laboratory, Department of Mathematics at Tampere University of Technology, Tampere, Finland (e-mail: anne.tervakari@tut.fi).

M. Kailanto is with Intelligent Information Systems Laboratory, Department of Mathematics at Tampere University of Technology, Tampere, Finland (e-mail: meri.kailanto@tut.fi).

This article is an extended and modified version of a paper presented at the EDUCON2013 conference held at Technische Universität Berlin, Berlin, Germany from March 13-15, 2013. It was supported in part by European Social Fund, North Ostrobothnia Centre for Economic Development, Transport and the Environment. Submitted 06 May 2013. Published as submitted by the authors 25 July 2013.