Mass collaboration – An emerging field for CSCL research

Article - January 2013

CITATIONS
4

READS
228

7 authors, including:

Ulrike Cress
Leibniz-Institut für Wissensmedien
378 PUBLICATIONS 6,349 CITATIONS

Brigid Barron
Stanford University
65 PUBLICATIONS 7,980 CITATIONS

Gerhard Fischer
University of Colorado Boulder
318 PUBLICATIONS 14,194 CITATIONS

Aileen Oeberst
FernUniversität in Hagen
67 PUBLICATIONS 635 CITATIONS

Some of the authors of this publication are also working on these related projects:

Information Processing of Sports Broadcasting View project

Collaborative Learning View project
Mass Collaboration – an Emerging Field for CSCL Research

Organizer
Ulrike Cress, Knowledge Media Research Center (KMRC), 72076 Tuebingen, Schleichstr. 6, Germany, u.cress@iwm-kmrc.de

Presenters
Brigid Barron, School of Education, Stanford University, barronbj@stanford.edu
Gerhard Fischer, Center for Lifelong Learning and Design, University of Colorado, Boulder, gerhard@colorado.edu
Ulrike Cress, Iassen Halatchlyiski, Aileen Oeberst; Knowledge Construction Lab, KMRC, Tuebingen
Email: u.cress@iwm-kmrc.de, i.halatchlyiski@iwm-kmrc.de, a.oeberst@iwm-kmrc.de
Andrea Forte, College of Information Science and Technology, Drexel University, aforte@drexel.edu
Mitchel Resnick, Media Lab, MIT, mres@media.mit.edu

Discussant
Allan Collins, collins@bbn.com

Abstract: Mass collaboration is a present-day Internet practice with far-reaching implications for education. The goal of this symposium is to establish the concept of mass collaboration as a relevant topic in CSCL by presenting related research done at different labs and with different focuses. Several presentations will provide insight into the current approaches to the complex and large-scale phenomenon. They will address a range of theories and methodologies, grounding the analysis of mass collaboration processes and outcomes, in order to identify examples of effective Web 2.0 settings and technological environments.

Introduction
Whereas in former times collaboration was mostly bound to smaller groups, the Internet tools of today provide various possibilities for the collaboration of masses of users. There is an almost unlimited variety of online communities where users share personal stories, experiences, or anything that can be expressed digitally. In wikis thousands of users collaboratively gather and organize knowledge. With social tagging systems users annotate and share online resources. The participants in such communities are not just a mass of learning individuals or passive consumers; they actively produce meaningful content and act as “prosumers” (O’Reilly, 2006; Tapscott & Williams, 2006). Most importantly, their activity develops outside the formal educational system (National Research Council, 2009).

This symposium will deal with those aspects of mass collaboration that have potential for education and learning and are thus of interest for the learning sciences (Cress, in press; Cress & Fischer, subm.; Fischer, 2011). The discussion will build on the symposium on long-tail learning at the CSCL conference in 2009 (Brown & Adler, 2008; Collins, et al 2009) and will extend the view to various forms of online interaction when masses of people learn together, collaborate and create new knowledge. The presentations will introduce relevant theoretical approaches to the process of mass collaboration; empirical studies will detail out different analysis methodologies, and examples of communities and Web 2.0 environments in formal and informal settings will ground the discussion of practical issues.

In sum, the aim of the symposium is threefold:
(1) to establish the concept of mass collaboration as a relevant topic of CSCL
(2) to present and interconnect existing research on the subject, and
(3) to give prototypical examples of learning communities and mass collaboration platforms

In the following we give a short introduction to the concept of mass collaboration, providing the background for the presentations in the symposium.

CSCL and Mass Collaboration
Since the emergence of CSCL as a research field the predominant approach has been to study small groups of students in a neatly arranged situation: The students engage in synchronous discourse around a problem-solving task, and the sequence of their interactions represents a major research interest with regard to meaning making and learning outcomes. Collaboration has been used in the sense of Roschelle and Teasley, (1995: p. 70) who defined it as “a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem”. Whereas CSCL in its early times mainly dealt with synchronous small group or classroom-settings, currently, CSCL research faces a much broader range of real-life situations, where people take part in decentralized communities, act asynchronously and do not necessarily all come to a
shared conception of a problem. In fact, complex knowledge phenomena involve long periods of time, large and changing numbers of people, and fuzzy-structured settings. In this spirit, any human achievement can be seen as a collaborative accomplishment – in terms of dwarves standing on the shoulders of giants. Extending the view on collaboration beyond small groups leads to a macro approach that considers the complexity of knowledge development across space, time, and collectives of people. CSCL research is just starting to address this global level of human learning and creation of knowledge (Kafai & Peppler, 2011; Collins et al, 2009). The increasing number of Web 2.0 communities and tools now bring this perspective to the foreground.

Artifacts as Mediators of Mass Collaboration
The large-scale perspective raises the question of how intersubjective understanding and collaborative activities of a vast number of people are coordinated. Bearing in mind that most of the participants cannot interact directly, some form of coordination is needed. All the more so when the process is collaborative and fulfills the conditions that individuals act consciously following a common direction; that they take the perspective of the other participants into account; and that they contribute by building on the accomplishments of others. The solution to the coordination problem resides in artifacts that support the collaborative process by mediating common understanding (Bruner, 1996). Beyond making artifacts accessible for a large number of people, social software environments afford their collaborative creation, revision and remixing in a mass collaborative process. Collaborative artifacts represent crystallized knowledge that is preserved from past interactive situations and that can be built on in future to produce phenomena like scientific understanding, social practices or social norms (Engeström & Sannino, 2010; Paavola, Lipponen & Hakkareinen, 2004). An artifact is a means to an end and at the same time it is a desired product whose creation can even be the main goal of the collaboration process (see for example Kafai & Resnick, 2000). Theoretical and methodological approaches to mass collaboration need to be centered around artifacts as major elements in the complex process across space, time, and collectives of people.

Learning in Mass Collaboration
Whereas formal learning mostly takes place organized in smaller groups and classroom settings, mass collaboration by definition happens “in the wild”, and thus, mainly outside the educational system. Most of the Web 2.0 communities that share and create knowledge maintain an informal context. Activities in these communities induce individual learning, and at the same time they represent collective knowledge creation (Paavola, Lipponen & Hakkareinen 2004). In order to interact, people have to create or refer to artifacts. In this process of usage and construction of artifacts people deepen their understanding and enhance their individual knowledge (Kafai & Resnick, 2000). Thus, learning and knowledge creation in a collaborative setting are one continuous process of internalization and externalization (Nonaka & Takeuchi, 1995; Vygotsky, 1978).

During mass collaboration it is obvious that learning occurs not just at the level of individuals, but also at the community level: A community may enhance its knowledge base, and may deepen its understanding of a given topic. Through cumulated reciprocal referencing during a discourse some views in the group may become more pronounced than others. Some ideas stand the test of time, others fade away. This is obvious in the example of wikis: Some ideas remain in the collaborative text, while others are revised or deleted soon after they have been contributed. These processes are not just random. They show that a community represents a self-organized, autopoietic system (Maturana & Varela, 1987). Knowledge development within this system is an emergent process. It depends on individuals’ activities but is not reducible to the knowledge of independent individuals. The “learning” of a community and the learning of the participating individuals are intertwined and suggest a systemic understanding of mass collaboration (Cress & Kimmerle, 2008).

Mass Collaboration Environments
There are many examples of mass collaboration environments for learning and knowledge creation. Perhaps the most remarkable example is the online encyclopedia Wikipedia. It offers a unique field for studying large-scale open-ended knowledge processes from large amounts of data on the history of articles and authors’ contributions. Wikipedia is a dynamic knowledge artifact of interconnected articles deliberately produced by a complex system of individual collaborative activities. The explicit written content mediates shared understanding on a specific topic amalgamating perspectives and styles of expression of a multitude of authors into a coherent exposition. Although Wikipedia is not aimed at developing new knowledge or at providing a learning environment for the contributors, the processes that unfold there share some essential features of scientific and knowledge-building discourse (Forte & Bruckman, 2006; Swarts, 2009). Thus there is a vibrant field of research concerned with data from Wikipedia, and several presentations in this symposium (those by Aileen Oeberst, Lassen Halatchylishki and Andrea Forte) will deal with it.

While Wikipedia has emerged as a self-organized community, a lot of mass collaboration platforms have been deliberately designed to support learning. One prominent example is the Scratch online community
(presentation by Michael Resnick), where children playfully create visual artifacts like animations, stories, and games. Learning and knowledge creation are again supported by and organized around user-generated artifacts. The environment affords not only sharing of own projects, but also building on existing ones (Resnick et al., 2009; Brennan, Resnick & Monroy-Hernandez, 2010; Monroy-Hernandez, 2012). Knowledge development in Scratch is supported by the possibility of remixing existing content. Analogous to Wikipedia the individual contributions may be interlinked, revised, and rebuild. Contributors use others’ contributions to create new artifacts, which then in turn represent the basis for future ones.

The third example of a mass collaboration environment in this symposium is Citizen Science (presentation by Brigid Barron). It stands not for one specific technology, but subsumes various projects where amateurs or nonprofessional scientists participate in scientific activities. Amateurs for example collect data or test natural phenomena. Here it is obvious that their activities lead to new knowledge. The main goal of the projects is not the learning of the individuals, but the creation of new knowledge by using scientific methods.

Structure of the Symposium
The symposium will demonstrate the relevance of mass collaboration for CSCL research by bringing together the latest theoretical and empirical endeavors in this research area. It is organized by the Knowledge Construction Lab at the KMRC, Tuebingen, where the Co-Evolution Model of Individual Learning and Collaborative Knowledge Building (Cress & Kimmerle, 2008) serves as a systemic framework for describing and analyzing mass collaboration. The symposium presents results from this line of research together with relevant work being done in other institutes. Together the presentations want to open a dialogue about the current challenges and opportunities for CSCL research in the context of mass collaboration.

Presentation 1: Cultures of Participation — Fundamental Transformations of Learning, Working, and Collaborating
The first presentation is an introduction to the topic, pointing out the current and societal relevance of mass collaboration. Gerhard Fischer, director of the Center for Lifelong Learning and Design (University of Colorado, Boulder), explores mass collaboration in the context of cultures of participation (Fischer, 2011; Jenkins, 2009). He states that social media enable a shift from consumer cultures, which are specialized in producing finished artifacts to be consumed passively, to cultures of participation, in which all people are provided with the means to participate and to actively contribute in personally meaningful problems. The participatory web supports moving away from a world in which a small number of people create artifacts, define rules, make decisions concerning many others towards a world in which everyone would be able to actively participate following their own interests and to make their voices heard. This shift has introduced unique and fundamental opportunities, challenges, and transformative changes for innovative research in CSCL. Fischer will explore in his presentation the theoretical foundations and system developments for understanding, fostering, and supporting cultures of participation. His framework is centered on the following aspects:

- **Meta-design** defines and creates social and technical infrastructures for cultures of participation in which new forms of collaborative learning and design can arise.
- **Social creativity** transcends the individual human mind by making sense of the variety of voices, in order to frame and solve complex problems. Shared artifacts enable relevant interdisciplinary collaborations.
- **Rich ecologies of participation will emerge** based on different levels of participation, expertise, interests, and motivation.
- Idiosyncratic interests and unique contributions by self-directed learners will lead to long-tail distributions of knowledge.
- **Drawbacks** of cultures of participation can be seen in aspects of “do-it-yourself” societies, in fragmented cultures in which people live in their own “filter bubbles”, and in accumulation of irrelevant information.

The framework is grounded in a variety of different application contexts (including: open source software, urban planning, assistive technology, energy sustainability, and formal education). The presentation will include initial design guidelines and explore the implications of these developments for future research and innovations in technology enhanced learning.

Presentation 2: The Co-Evolution Model as a Theoretical Framework for Describing Mass Collaboration
The second presentation will show the research done at the KMRC in Tuebingen. It consists of three parts with three different presenters: Ulrike Cress will present the Co-Evolution Model of Individual Learning and Collaborative Knowledge Building as a research framework for describing and analyzing learning processes in mass collaboration (Cress & Kimmerle, 2008; Kimmerle, Cress & Held, 2010). The model describes individuals as cognitive systems, who externalize their knowledge into an artifact. The community deals with the individuals’ contributions by interlinking, revising or even deleting them. In its activities the community acts as
a social system that deals with the information according to its own rules. This process may result in the creation of emergent knowledge. Users working with artifact (just through consummation, but much more through participation) internalize the rules and knowledge, which can lead to individual learning. Thus, on the one hand, a system makes use of an individual’s knowledge and creates collective knowledge; and on the other hand, an individual develops the own individual knowledge by actively participating in a community. This means that the analysis of mass collaboration learning should not only consider individuals as unit of analysis but also the social system, which is the whole community that comprises the totality of all users’ activities. Mass collaboration should not just be considered as an environment, where individual learning takes place, instead, it is an actor itself. It shapes the individual’s activities.

The following two presentations give an example how processes on this system level can be analyzed. The first study is a qualitative in-depth analysis of a special Wikipedia article, while the second is a quantitative large-scale analysis basing on thousands of interlinked pages.

Aileen Oeberst will present a study of the development of the Wikipedia article about the nuclear power plant in Fukushima during the first nine days after the nuclear disaster on March 11, 2011. She will show how the rules of the social system Wikipedia guided individual contributions to a topic where—at that time—no verified knowledge existed. People had to deal with a flood of novel, highly specific and mainly uncertain incoming information and they had to construction meaning out of different information sources. Many authors were involved in this process. Their activities and coordination were only mediated by Wikipedia’s norms, that valuable and thus accepted contribution should be verifiable and written from a neutral point of view. This norm effectively shaped individuals’ edits and enabled authors without much domain expertise to construct an article, which nuclear experts judged to be of high quality. These findings bring not the individuals as relevant contributors but the rules of the knowledge system to the fore. The system’s definition of knowledge implicitly coordinated the knowledge construction process, and shaped the activities of its contributors. In sum, the presentation shows that mass collaboration is much more than sharing and accumulating knowledge across individuals. It is an emergent process, where the system makes use of individuals to create system-accepted knowledge.

Iassen Halatchlyiski will present an approach to studying mass collaboration that can encompass large-scale and long-term characteristics: the macro level of the phenomenon (Halatchlyiski, Moskalik, Kimmerle & Cress, 2010). His study deals with pivotal knowledge within Wikipedia. He asks (1) what pivotal knowledge in Wikipedia is, and (2) who its authors are. In his analysis he combines the level of artifacts and the level of individuals in order to investigate authors’ experience as explanatory factor of the contribution of pivotal knowledge at the system level. In his analysis he considers two adjacent knowledge domains, each covering thousands of articles (articles of the categories Psychology and Pedagogy in the German Wikipedia).

He bases his empirical study on the network concept. This has already been used to describe knowledge organization at different levels, e.g. the semantic memory of individuals (e.g., Collins & Loftus, 1975), or meaning-making processes within a group discourse (Stahl, 2006). It is now used to describe knowledge processes in mass collaboration environments like Wikipedia, where knowledge resides in interlinked artifacts. In accordance with the network perspective, he operationalizes “pivotal knowledge” as articles with a central position at the macro level and distinguishes two types of pivotal knowledge: knowledge that is central in a single domain or knowledge that relates two different domains and resides at the border between them.

The presentation shows that social network analysis (SNA) is an appropriate methodological approach to go about the complex large-scale and long-term patterns of creation of pivotal knowledge.

**Presentation 3: Large-Scale Collaboration and Cultures of Accountability**

The third presentation deals not only with the aspect of production, but also with the aspect of consumption of knowledge and information provided in mass collaboration environments. Andrea Forte from Drexel University’s iSchool will talk about the intersection of critical consumption and production of participatory information resources. She will discuss her research group’s ongoing efforts to understand how contributing to online information resources can not only yield content knowledge, but also expose learners to cultures of accountability and equip them to engage critically with information sources.

Research on collaborative environments involves understanding who learns what, when, and how. In Wikipedia and other wiki systems, collaborative writing supports editors in learning about a wide range of topics and the characteristics of this learning have been the topic of many studies, including Forte’s and others in this Symposium. However, there are other kinds of learning taking place as people become socialized in communities of collaborative authorship.

As the world’s largest and most widely accessed collection of reference works, quality is a critical concern for Wikipedia projects. Each language edition has developed complex policies and social norms that help contributors construct high-quality artifacts and that guide discourse on the site (Butler et. al., 2008;
Kriplean et al., 2007); this has bred a culture of accountability among contributors. Analogously, when asked to contribute to a public information resource as part of their classwork, high school students report a sense of responsibility to their potential readership for the quality of their written work (Forte & Bruckman, 2009). Even when they work individually, this sense of responsibility is manifest in student strategies for citing, organizing content, and looking for information. Moreover, when they have experience producing information online, some students begin to leverage these experiences to assess the quality online information resources they use. Forte’s research group is expanding this work to better understand how experiences with large-scale online collaboration affect people’s understanding and assessment of online resources both on Wikipedia and in more specialized collaboratively produced resources like Ancestry.com or Findagrave. Early findings suggest that experience in contributing to participatory information sources yield sophisticated strategies for engaging with online sources.

Whereas the first three presentations mainly considered theoretical aspect that are relevant for mass collaboration environments in general, the last two presentations deal with specific environments and show how they enable learning and knowledge building.

**Presentation 4: Learning through Remixing**

**Mitchel Resnick** from the MIT Media Lab will discuss his group’s research on mass collaboration in the context of *Scratch* (http://scratch.mit.edu), an online community in which young people (ages 8 and up) program interactive stories, games, animations, and simulations — and share their creations with one another online (Resnick et al., 2009). Since the launch of Scratch in 2007, young people around the world have shared roughly 3 million projects in the online community, adding (on average) two new projects every minute. The collection of projects is incredibly diverse: interactive newsletters, science simulations, virtual tours, animated dance contests, interactive tutorials, and many others, all programmed with Scratch’s graphical programming blocks.

An important aspect of participation in the Scratch online community is the ability to remix other people’s projects. Members of the community not only interact with and comment on one another’s projects, they can also modify, extend, and repurpose the programming code and media elements underlying the projects. All projects in the Scratch community are covered by Creative Commons Share-Alike license, so community members are free to build upon the sprites, scripts, images, and sounds of other people’s projects. Roughly one-third of all projects in the Scratch community are remixes of other projects.

In his presentation, Resnick will analyze and discuss:

- Different approaches to remixing within the Scratch community
- What people learn (and don’t learn) as they remix one another’s projects
- How ideas spread through the community through remixing
- Community attitudes towards remixing — and factors underlying those attitudes
- Design strategies for supporting and encouraging a culture of remixing

Many of the ideas and examples in the presentation will draw upon the research of Andres Monroy-Hernandez, who recently finished his PhD dissertation focusing on remixing in the Scratch community (Monroy-Hernandez, 2012).

**Presentation 5: Long-tail learning and access to external resources**

**Brigid Barron** and Caitlin Martin from Stanford University will report findings from a study of a genre of cyber-enabled massively collaborative activity known as *Citizen Science*. Citizen Science projects capitalize on the interest and efforts of non-scientist collaborators who join forces to contribute data that helps address problems of concern. Networked technologies have dramatically changed the potential of such projects. With mobile GPS enabled data collection devices, data contributions can easily be shared. Applications like Google Maps make it easy to share location based information and online databases allow data contributions to be reviewed by professional scientists and community members. Our research project was based in the premise that we need to understand Cyberlearning as a human-technical system and that to advance design relevant knowledge we need to attend to both the social community and the ways that the technology supports learning within the community. We chose to study a Citizen Science project with a significant focus on education called *Vital Signs*. Vital Signs is a citizen science networked system located in the state of Maine, linked statewide to schools and accessible not only to teachers but to anyone who want to learn. Because the state has a longstanding laptop program started in 2002, all middle school students have access to their own iBook, which they use at home and at school. Middle school teachers were also provided with laptops, technical assistance, and professional development. We will present findings from our research organized to address three main questions:
1) What patterns of personalized learning can we identify among a diverse group of teachers/learners who vary in prior experiences and community socioeconomic status?

2) When do Vital Signs opportunities sustain engagement beyond the classroom (for example by sparking family-based learning activities through a school-based project thereby bridging across formal and informal settings)?

3) How can member contributions to networked communities be mined and harvested for formative assessment data that designers and teachers can use to improve learning processes and outcomes?

**Discussant**

As discussant Allan Collins will consider what these projects tell us about whether people are becoming smarter and what they are learning by working in these mass collaborative environments. What are the gains and losses from participation in such environments? Are people thinking more deeply about important matters or are they flitting from topic to topic in the shallows that Nicolas Carr (2010) bemoans?

**References**


