AFEL

Analytics for Everyday Learning

EC Project 687916

D4.1 - Report on the analysis of learning and cooperation

Deliverable Coordinator: Peter Holtz, Leibniz-Institut für Wissensmedien IWM (Knowledge Media Research Center)

Contributors: Joachim Kimmerle, Seren Yenikent, Ulrike Cress (IWM), Mathieu d'Aquin (OU), Stefan Dietze, Besnik Fetahu, Ujwal Gadiraju (L3S), Belgin Mutlu, Peter Hasitschka (Know-Center).

Reviewers: Ricardo Maturana (GNOSS), Ilire Hasani-Mavriqi (Know-Center)

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Change Log

Version	Date	Amended By	Comment
0.1	04/10/2016	Peter Holtz	First draft
0.2	10/10/2016	Peter Holtz	More elaborate draft
0.3	12/10/2016	Peter Holtz	Bookmarks and links
0.4	18/10/2016	Peter Holtz	First comprehensive draft
1.0	31.10.2016	Seren Yenikent/Peter Holtz	Consolidated draft
1.1	10/11/2016	Besnik Fetahu, Stefan Dietze, Ujwal Gadiraju, Mathieu d'Aquin, Belgin Mutlu, & Peter Hasitschka	Revised draft
1.1a	16/11/2016	Ilire Hasani-Mavriqi	Review comments and minor edits
2.0	30/11/16	Keyur Dave	Formatting and finalising report

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Executive Summary

Within this deliverable, we will first explain the theoretical underpinnings of the socio-cognitive models that guide research and development within the AFEL project and discuss previous empirical studies on key assumptions of our theoretical framework such as the need for productive friction in learning. Afterwards, we will present the ongoing and planned collaborations with members of the other work packages and discuss the implications of these research projects for the AFEL project.

Knowledge, traditionally defined as 'justified true belief' [OKC16], is often regarded in psychology as an individualistic and rather static phenomenon. Not only since the advent of the internet as a mass phenomenon, there were suggestions for a more social and dynamic concept of knowledge. Since the 1990s, the 'knowledge construction metaphor' [PLH04] has emerged as an important theoretical framework to analyse processes of learning and knowledge construction in digital environments. The 'co-evolution model of learning and knowledge construction' [CK08] treats individual learning and collective knowledge construction as structurally coupled processes and serves as our guiding theoretical framework for the AFEL-project. Our theories guided the design of the AFEL data source taxonomy in D1.1 and the development of the AFEL knowledge base schema for D2.1 together with colleagues from work packages WP1 and WP2. Our socio-cognitive models were taken up in WP3's deliverable D3.1 as well.

Together with project partners from the L3S (WP2), we are currently preparing a study on processes of collaboration in resource-centric environments such as Slideshare, Bibsonomy, or Wikipedia. In a first step, we are investigating which contributor and discourse characteristics are predictive of positive or negative developments in the collaborative construction of Wikipedia articles. An example for a negative development would be a violation of Wikipedia's 'neutral point of view policy' resulting from an 'echo chamber effect' [DBZPSCSQ16] caused by a lack of heterogeneity with regard to attitudes and previous knowledge on the side of the contributors. We are also currently planning to work together with work package WP3 in the visualization of fragile and contested knowledge graphs. Together with colleagues from work package WP2, we investigate the effectiveness of achievement primes in task-centric social networks such as Crowdflower and Amazon MT.

In all these cases, our studies aim on the one hand to increase our theoretical understanding of processes of learning and collaboration and the respective incentive structures in digital environments; on the other hand, we aim as well at using these insights for the development of software tools to support learners as well as administrators. For example, our findings on indicators of beneficial and detrimental processes in the collaborative construction of

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Wikipedia articles can serve as a model to develop ‘warning systems’ for other resource-centric websites such as Slideshare and Bibsonomy.

Table of Contents

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Introduction[1]

The main goal of AFEL-project is "... to develop methods and tools to understand informal/collective learning as it surfaces implicitly in online social environments." [afel-project.eu ; last retrieved Oct. 10, 2016]. Within this document, we will first discuss the theoretical background of the socio-cognitive models that are developed in work package WP4 and intended to describe and explain processes of informal and collective learning. These models guide the data collection, enrichment, and analysis in the work packages WP1, WP2, WP3, and WP4 and the intended application of research findings in form of theoretical insights and newly developed technologies in the GNOSS-platform (work package WP5). In this deliverable, we will focus on the social aspects of everyday learning and collaborative knowledge construction.

The second part of the deliverable summarizes empirical findings from previous studies on some of the key assumptions of the 'co-evolution model of learning and knowledge construction' [CK08], which serves as the theoretical backbone of our models. Within the third section, we will describe the research activities in form of collaborations between work package WP4 and the other work packages that are currently ongoing or planned for the duration of the AFEL project. In the fourth section, we will focus on the relations between our theoretical framework and the development and application of technologies to enhance and facilitate learning and knowledge construction in online environments before a few concluding remarks.

Knowledge in the Information Age

The "digital revolution" in the second half of the 20th century and the emergence of the Internet in the 1990s have profoundly changed people's daily lives and the social and economic world [Arn02]. In view of the fact that most of people's knowledge is premeditated by the media [e.g., HP13], it comes to no surprise that the Internet strongly affects our knowledge as well. Already in 2006, comedian Stephen Colbert coined the term "Wikiality": Much of our knowledge is mediated through the mass media – namely the Internet and most prominently the online-encyclopedia Wikipedia. And Wikipedia, as a collaboratively constructed knowledge corpus, is finally only what its users make of it. Hence, "together we can create a reality that we all agree on—the reality we just agreed on." If the declining number of African elephants makes us sad, why don't we change it – by just editing some numbers on Wikipedia (instead of costly and strenuous and possibly ineffective conservation measures)? What was of course meant to be a cynical joke touches one of the most important

[1] The introduction is based on the introduction of the following paper: Holtz, Cress, & Kimmerle (forthcoming). Analyzing and predicting massive learning and knowledge construction processes in digital environments.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

problems in philosophy: What is knowledge - and how does our notion of knowledge change in the information age?

Since antiquity, knowledge was often defined as 'justified true belief' [for a summary see e.g., OKC16]. In psychology, knowledge is often equated with memory structures that represent a justified true belief (ibid.). In both approaches, knowledge is treated as an *individualistic* phenomenon insofar as knowledge is situated within a person's mind: A person holds a certain memory that is considered to be knowledge in an epistemological sense, if it entails a justified and true belief about something. Here, *justified* refers to the sources of knowledge or the methods of knowledge acquisition (a guess that is true by chance alone would not be considered to be knowledge by Plato, e.g. in the Meno-dialogue), whereas *true* refers to observable facts or consequences in form of knowledge-based assumptions or hypotheses. From the fact- or outcome-based knowledge described here, other forms of knowledge such as procedural knowledge ('know how') and experiential knowledge ('knowing by acquaintance') can be discerned.

Not only since the advent of mass communication, there have also been attempts to conceptualize knowledge in a more social way: Do I 'know' (as an individual) that homeopathy is ineffective (or vice versa) or do I just partake in a culture that does not accept the homeopaths' justification procedures and that does not agree with the homeopaths' interpretation of empirical evidence (or vice versa)? Is the 'knowledge' of practitioners of an innovative artistic genre something that the artists acquired or rather something they created themselves? Many researchers (among many others, Karl Popper, Richard Rorty, Niklas Luhmann, Humberto Maturana) and social scientists (e.g., Kevin Gergen, Serge Moscovici) have questioned that it is finally possible to define objective criteria for 'justification' and (empirical) 'truth' beyond all doubt. Hence – at least partly – knowledge needs to be considered as a social phenomenon, if only because concepts of justification and truth rely—at least partly—on negotiated social consensus.

Metaphors of Learning

Somewhat similar to this discussion about epistemology in general is the discussion about the nature of learning that Anna Sfard's seminal paper "On two metaphors of learning and the dangers of choosing just one" [Sfa98] put forward. The article is based upon a previous debate within the *Educational Researcher* between representatives of a more cognitive approach to learning [e.g., ARS96] and endorsers of a situated learning approach [e.g., Gre97]. In a nutshell, Anderson and colleagues argued that situated learning is based on counterfactual claims such as that "knowledge does not transfer between tasks" (p. 6). Whereas there may be cases, in which a transfer of learning outcomes is difficult, overall, there is sound empirical evidence from cognitive psychological studies that transfer is possible. Greeno [Gre97] replied that to him Anderson and colleagues misunderstood situated learning—or in broader terms a more constructivist, discursive, or symbolic interactionist approach to learning—which is not meant to replace a cognitive approach; both approaches

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

try to answer different questions. For example, Greeno would rephrase the statement "knowledge does not transfer between tasks" as the question "how should the social conditions of learning be arranged?" (p. 8), which is more typical for situated learning research. Toward the end of his reply, he argued that a synthesis between the cognitive/individualistic strain of educational research and the other one, which is based upon "methods and concepts of ethnography, ethnomethodology, symbolic interactionism, discourse analysis, and sociocultural psychology" (p. 14), is needed for the field of educational science to move ahead.

Sfard [Sfa98] has based her attempt at a synthesis of the two positions on an analysis of the two underlying metaphors: The *acquisition metaphor* and the *participation metaphor*. The acquisition metaphor implies that learning consists in the accumulation of "basic units of knowledge" within the "container" (p. 5) of the human mind. These units exist independently from learners and teachers. Once acquired, knowledge can be developed, transferred, transmitted, and grasped similar to other physical goods. Even in "moderate to radical constructivism", the acquisition metaphor is used frequently. Though learning is conceptualized in these approaches as a "never-ending self-regulated process of emergence in a continuing interaction with peers", the "essence" (p. 6) of what is to be learned is still something that can be acquired, transferred, and developed.

Her analysis of the participation metaphor reveals that here frequently the noun *knowledge* is replaced by the verb *knowing*: Knowledge is not regarded as an object that can be developed, acquired, or transferred, it is an activity in which people can participate. These activities are connected to certain contexts within which they take place. Teachers can be understood as the "preservers of ... continuity" (p. 6) of the norms and customs of the respective community, every learner is a potential "reformer" (p. 6). Here, learning is a social process: The outcome of learning is the enculturation of the learner into a community of practice [LW91]. To Sfard, both metaphors have their strengths and weaknesses, and a successful educator will have to take both aspects of learning into account.

Knowledge Construction in Digital Environments

In 2004, Paavola, Lipponen, and Hakkarainen [PLH04] argued in favor of the need for a third metaphor of learning: The knowledge creation metaphor. In their seminal paper, they synthesize three different approaches to a more social and constructivist conception of knowledge: Nonaka and Takeuchi's [NT95] model of *knowledge creation*, Engeström's [Eng99] model of *expansive learning*, and Bereiter's [e.g., Ber02] model of *knowledge building*.

Nonaka and Takeuchi [NT95] arrived at their model from an analysis of innovation processes in Japanese firms. They found that for innovation the dynamics between tacit and explicit knowledge play an important role. Workers such as expert bakers build up a vast corpus of knowledge through experience, which usually remains tacit, unless it is communicated to

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

others. In order to use this knowledge, for example, for the invention of an innovative baking machine, a communication process between bakers and engineers must be facilitated: The tacit knowledge needs to become explicit in a dialogical process. The result of this dynamic interaction is the creation of new knowledge [see also KMCK15].

Engeström [Eng99] studied workplace contexts as well to arrive at his model of expansive learning. He applied the Vygotsky-inspired *Cultural-Historical Activity Theory* (CHAT) to organizational learning processes [see also Leo78]. CHAT examines the interplay of mediating artifacts and people in the context of dynamic human activity systems that also include the community that people belong to and that specify the rules and norms that are prevailing. In this model as well, the focus is on the verbalization of experiential knowledge and the consequential possibility to communicatively arrive at new knowledge.

Bereiter [Ber02; SB94] came from a more educational background. His focus is on the need to restructure schools as knowledge building communities to facilitate the acquisition of expertise. Very much like in scientific communities, schools should give students the possibility to discover knowledge in a social environment. By doing so, schools can make use of two strong motivators: Curiosity and the need for social belonging. Educational technology can play an important role in facilitating these processes of collective knowledge building.

One characteristic of all these approaches is the "pursuit of newness" [PLH04, p. 562]: Knowledge is not some *object* that can be acquired; it is neither the acculturation into an existing community of practice [WS00]. It is something that is created from interactions between persons in form of communication and/or interactions between persons and 'things', for example in forms of attempts at problem solving. Of course, newness does not equal chaos or randomness. Knowledge creation—be it in the classroom or in work and organizational contexts—builds upon existing knowledge and in many cases, knowledge will be re-created in view of the *constraints of reality* that constitute the limits of what *can* be created in a given situation. Hence, learners in a comparable situation will end up creating similar knowledge given similar tasks and comparable prerequisites.

Even a growth of knowledge can be conceptualized within a collective knowledge building approach: Karl Popper [e.g., Pop72] developed his evolutionary model of the creation of 'objective knowledge' from the assumption that a growth of knowledge is only possible through the proposal of solutions for problems and subsequent criticism and attempts at an improvement over these solutions. A growth of knowledge needs the interplay of "three worlds" [Pop79]: The world of physical objects (world 1), the world of mental states (world 2), and the world of ideas and abstract concepts (world 3). Problems arise from interactions between the first two worlds: Human beings try to solve problems that arise from the interactions with objects (in a very wide sense). To facilitate discussions, criticism, and finally growth of knowledge, mental states (world 2) need to be communicated in the form of abstract ideas (world 3). These ideas have a 'life of their own' insofar as what others make

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

of these ideas once they have been brought into the realm of communication, clearly transcends the mental states that brought forward the idea [see also Hol16]. Socially shared ideas (world 3), in turn, can have an impact on the physical world (world 1), when they are taken up by individuals and influence mental states (world 2). Whereas Popper's main objective was the specific case in which world 3 objects are increasingly refined in an evolutionary way by means of attempts at falsification, the Three Worlds Theory applies to all forms of *cultural knowledge*.

Apart from the pursuit of newness, Paavola, Lipponen, and Hakkarainen [PLH04] maintain as well that knowledge creation is necessarily a social process—without communicating and exchanging mental states, the creation of something new that transcends mere experience is not possible. To facilitate this exchange of ideas, mental states must be transferred into socially shared objects or "artifacts" (p. 566).

Hakkarainen and colleagues [HP09; Eng09] further developed the knowledge creation metaphor into their *trialogical approach to learning*. Here, the role of the artifacts for knowledge creation is elaborated upon in more detail. In collaborative knowledge creation, different types of artifacts make the exchange of ideas possible: *conceptual objects*, for example, in the form of questions, theories, and designs; *material objects* in the form of collaboratively written documents, for instance; and finally, *procedural objects* such as certain norms and behavioral scripts can facilitate the exchange between members of a knowledge creation community as well. These objects may mediate between individual and collective activities.

Most research on the trialogical approach (for example in the Knowledge Processes Lab, http://kplab.evtek.fi/) so far has focued on knowledge creation in small groups of learners. Nevertheless, we argue that the real potential of trialogical artifact-based interactions for knowledge creation lies in mass collaboration on internet platforms such as Wikipedia and Scratch. Here, the artifacts are the collaboratively created text resources and in some cases the additional communication channels that are provided by the platform. Analysis of learning processes can address different forms of knowledge creation: they can focus on changes within individual participants' representations (learning trajectories); they can examine the artifact itself and study the reification of the collaboratively created knowledge; finally, aspects and behavior patterns of the knowledge creation community itself can become the object of scientific inquiry.

The Co-Evolution Model of Learning and Knowledge Construction

In the co-evolution model of learning and knowledge construction, individual learning and collective knowledge construction both result from the interaction of two operationally closed systems [CK08]: On the one hand, the cognitive systems of human beings with the modus operandi of thinking; on the other hand, a social system with the modus operandi of

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

communication. The main function of both systems is to deal with the complexity of the matter at hand and to make the world intelligible and manageable (see figure 1).

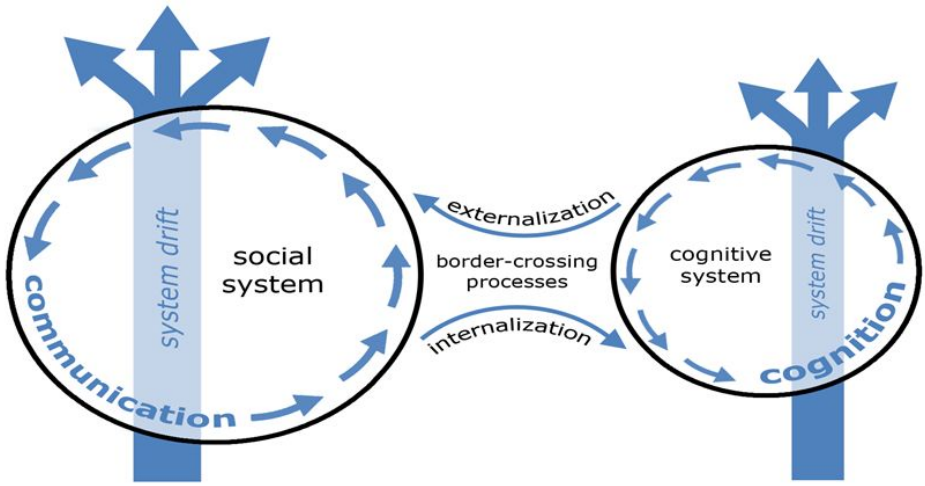

Figure 1: The dynamic processes of learning and knowledge construction [KMOC15, p. 128].

Operationally closed means that the processes within both systems are primarily sustaining the respective system in itself, although the systems are constantly involved in exchange processes with their environment. This process is called autopoiesis or self-organization [VMU74]. Both systems serve as environment for the other and trigger processes within the other system by means of irritating it. For example, a contribution to a wiki by person *A* can irritate the cognitive system of person *B* (when grasped and perceived by *B*; this process is called internalization) and trigger certain thought processes to ensure the systems functionality. On the other hand, an idea within person *C*'s cognitive system can—when expressed through externalisation—cause irritation in the social system and lead to processes of communication. These processes of dealing with irritations lead to a system drift over time: In case of the social system, this drift constitutes knowledge construction; in case of the cognitive system, the system drift can be called learning (see Figure 1). Both systems are structurally coupled [Mat75] insofar as they co-evolve toward greater capabilities in reducing complexity.

The key entities within the co-evolution of learning and knowledge construction in the context of everyday learning are persons and (digital) artifacts. Persons can be related to artifacts via authoring or editing them or via reading or consuming them. All forms of behavior that can at least potentially be perceived by others constitute acts of communication. People who participate in a knowledge resource in form of an electronic artifact via editing or consuming it constitute an online community or virtual community of individuals who share

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

an interest in the underlying topic(s) and who are willing to share their ideas with other community members or to consume other members' notions (see figure 2).

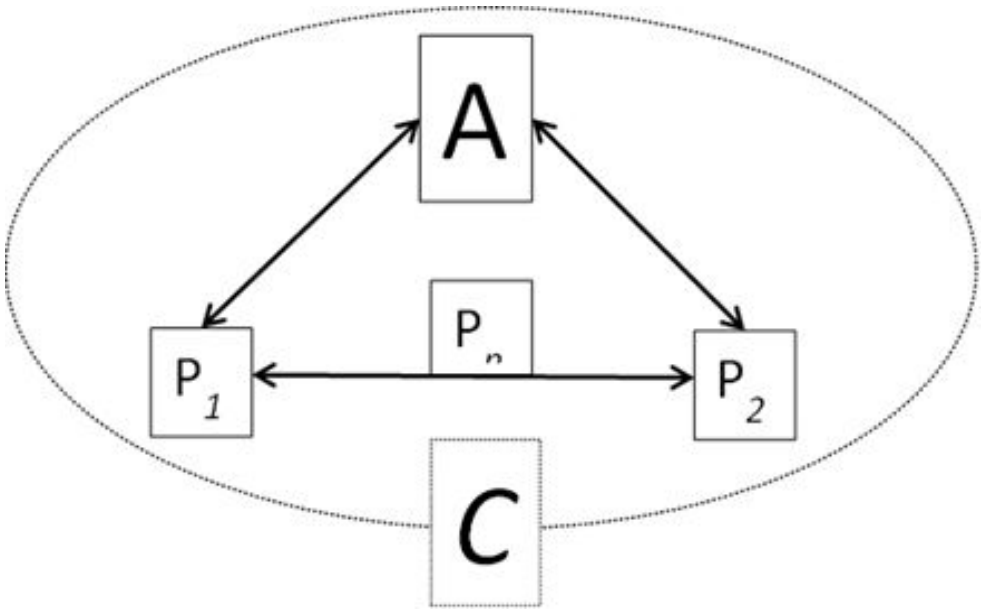

Figure 2: Features and Entities (A = Artifact; P = Person; C = Community; Arrows signify interactions).

Key Assumptions

The aforementioned theoretical considerations and in particular the co-evolution model of learning and knowledge construction serve as the guiding theoretical framework for the AFEL-project. We can summarize our main ideas as follows:

A1 Learning arises from encounters between individuals ('learners') and new and/or surprising information that is integrated into the learner's' cognitive system.

A2 Learning can take place intentionally (in form of planned learning activities) or unintentionally; learning can happen in a conscious or an in an unconscious way (see Figure 5).

A3 The structurally coupled epiphenomenon of learning on the side of individuals is knowledge construction on the side of communities.

A4 In all cases of intentional or unintentional and conscious or unconscious learning in digital environments, interactions between learners and digital artifacts are pivotal.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

A5 These interactions can be ‘passive’ for example in the form of a learner reading text on a Wiki or can be ‘active’ for example in the form of a learner editing a document.

A6 In both cases, the learners’ behavior leaves traces in form of for example log files or edit histories.

A7 Online activities that are related to one and the same topic—be they intentional or unintentional and conscious or unconscious—can in part be reconstructed in form of a learning path consisting of interactions between a learner and artifacts representing a certain topic.

A8 From this learning path, it is possible to reconstruct a learning trajectory comprising of “the learning goal, the learning activities, and the thinking and learning in which the students might engage” [SIM95, p. 133].

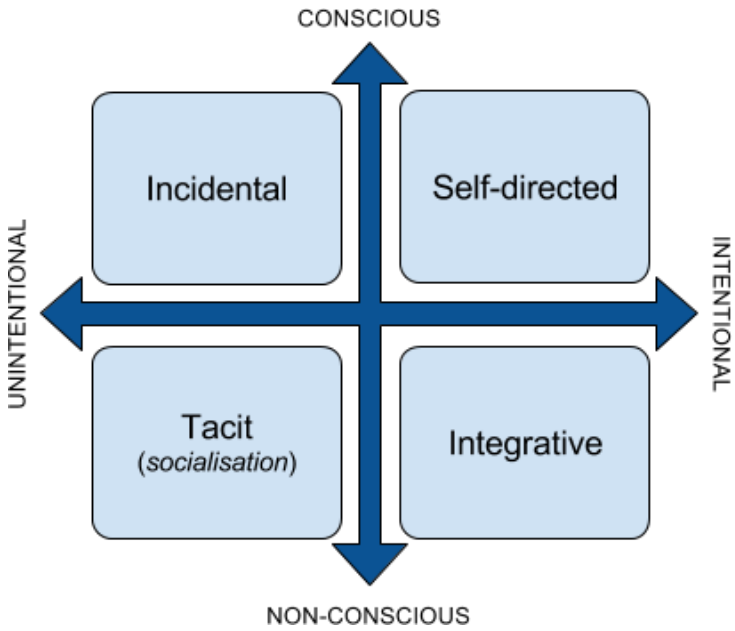

Figure 5: Bennett’s four-part model of informal learning [Ben12]; taken from D1.1

Implications for the AFEL - data source taxonomy

The aforementioned key entities—learners and digital artifacts—and their interactions form also the core of the AFEL data source taxonomy [D1.1, p. 22 ff.]:

Users: The data source categories ‘user data’, ‘user activity records’, ‘behavioral traces’, and ‘indicators of learning effectiveness’ all refer to ‘learners’ in the digital environments. From more elementary data such as log files and edits, more complex user characteristics such as

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

personality characteristics ('user profile') and activity patterns such as 'learning objectives' will be derived.

Previous research over the last years [e.g. KSG13] showed that it is possible to derive even complex and elusive user characteristics such as personality traits, intelligence, and sexual orientation from behavior traces such as Facebook status updates, postings, and likes. In the case of personality traits, the validities of the estimates based on internet behavior reached *r*s > .40[2] and hence surpassed even the validity of scores from ratings by trained observers [e.g. MPS94].

One challenge for the AFEL project is that in order to reliably assess hidden psychological characteristics, 'training data' in form of reference data (e.g. questionnaire or test scores that can be related to digital behavior traces) that were collected in a more 'traditional' way is needed. Whenever possible, we will try to enrich the data from work packages WP1 and WP2 with external data sources such as questionnaire data as well (see the paragraph '3.3 further studies' below). Another pivotal source of information for the modeling of successful learning trajectories are also the aforementioned 'indicators of learning effectiveness' such as (i) results of online tests and quizzes and (ii) behavioral traces such as the time needed to consume/solve documents/tasks and in some cases (iii) status indicators mirroring users' success in solving tasks.

Artifacts: The categories 'resource data' and 'system data' refer to characteristics of the digital artifacts the learners interact with. System data refers to more technical information on the platform that hosts the digital artifacts; resource data refers mainly to characteristics of the artifacts such as their complexity, controversiality, and potentially 'biasedness' as well as the artifacts' relations to other resources. Extracting as much information as possible on the resources in work package WP2 will be pivotal as means of extracting learning activities, paths, and trajectories, which all are related to activities within substantially 'similar' domains (see below the paragraph '... towards a formal definition of learning activities').

Groups and Communities: The co-evolution model of learning and knowledge construction [CK08] explicitly takes into account that learning is not only a cognitive activity that takes place within a socio-historical vacuum: learning as an individualistic phenomenon is always coupled to processes along the social dimension. Hence, one aim of the project is to also derive information on learning and knowledge building communities and hence the social processes that constitute the epiphenomenon of learning activities.

Social networks analysis and dynamic visualization will be important techniques within work package WP3 to investigate communities as temporal-dynamic structures of interrelated entities.

[2] These validity estimates represent the correlations between the Facebook-activity-based estimates and participants' scores in established personality tests.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

The AFEL-Glossary and steps towards a formal definition of learning activities

Based on the aforementioned theoretical considerations, an AFEL-glossary was created collaboratively by members of all work packages. This glossary comprises definitions of overarching key terms such as knowledge, learning, learning activities, learning scopes, and learning trajectories. The AFEL-glossary was designed to bridge the gap between the very wide conceptualizations of learning processes in the theoretical framework and the very narrow and concrete formal definitions of concepts and properties in the AFEL Knowledge Base Scheme.

Another approach towards translating the theoretical base of the AFEL-project into concrete operationalizations in the context of the development and application of learning application, underlies the paper "Towards defining learning activities to detect them in users' online activity streams" [DHAKDur], which has been submitted for the 2017 Learning Analytics & Knowledge (LAK) Conference in Vancouver, Canada. The paper addresses the problem of how complex activity patterns such as a 'learning trajectory' [see Sim95] can be derived from elementary user activities such as browsing activities. As a first step, it presents concise formalized definitions of terms such as 'learning trajectory', 'learning path', or 'learning scope'. In this paper, we assume that learning in the context of everyday learning can be defined as "... learners encountering at least partly new information in form of digital artifacts". Among all the user-artifact interactions that can be mined from a given platform, activities can be regarded as belonging to a certain 'learning scope', if the digital artifacts are related to the same overarching topics. The idea in this work is therefore to first group activities using specific resources according to the topics they cover and find the "central theme" of each group. We can then trace the usage history of each resource, extracting contributions from each activities to the learning of this theme. In the specific case used in the paper, activities are obtained from the learner's browsing history, and contributions are considered according to the way in which new activities introduce new subtopics in a given scope (group of activities). We used a Named Entity Recognition technique relying on the taxonomy of Wikipedia categories to extract topics of activities/resources, and a standard K-Mean-based process for creating groups/scopes as clusters of activities.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Previous Empirical Studies[3]

In the following section, we will provide a small overview of previous empirical studies that have investigated key aspects of the co-evolution model of learning and knowledge construction [CK08]. One such key aspect, which sets this model apart from other theories, is the importance of 'productive friction' in form of information that is at least partly inconsistent with the learner's expectations for learning success.

Laboratory Experiments

In the 1970s, Willem Doise, Gabriel Mugny and colleagues performed several psychological laboratory experiments which demonstrated the importance of 'socio-cognitive conflicts' [MD78] for learning. In one line of experiments [DMP75], they found that children did not only solve spatial coordination tasks more efficiently when working in teams of two children than when working alone; the children could also transfer the knowledge they acquired in a social coordination task to a Piagetian conservation of liquids task, which they were unable to solve beforehand.

In another series of experiments, Mugny and Doise [MD78] could demonstrate the importance of conflicts and their resolution via communication and social interaction for learning. Children were only better in learning how to solve the spatial coordination task from the previous study [DMP75], if they employed different problem solving strategies, which could be compared and discussed with the other participant: A socio-cognitive conflict seems to drive learning. Similar results could be obtained as well from experiments that concentrated on interactions between a child and an adult [DMP75]. Later studies [e.g., DDB07] showed that only an epistemic conflict in the sense of discrepancies between two persons' knowledge can produce this productive friction and have a positive impact on learning, whereas a relational conflict between two persons does not.

Processes similar to those analyzed in the aforementioned studies do supposedly also play a role in collaborative knowledge construction scenarios. In two experimental studies, Moskaliuk, Kimmerle, & Cress [MKC12] could show that productive friction does indeed contribute to successful online learning. Participants were in both studies first provided with some information on scientific studies on the causes of schizophrenia. In the first experiment, a cognitive conflict was operationalized through the redundancy or the conceptual overlap between information about schizophrenia that was previously provided to the participant and information they found in a fictitious wiki, which they were supposed to work on during the experiment. Learning success was measured by means of a questionnaire-based test afterwards. Participants in the medium-inconsistency condition showed significantly better

[3] Parts of this paragraph were as well taken from the paper: Holtz, Cress, & Kimmerle (forthcoming). Analyzing and predicting massive learning and knowledge construction processes in digital environments.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

learning outcomes than those in the low and high inconsistency conditions.

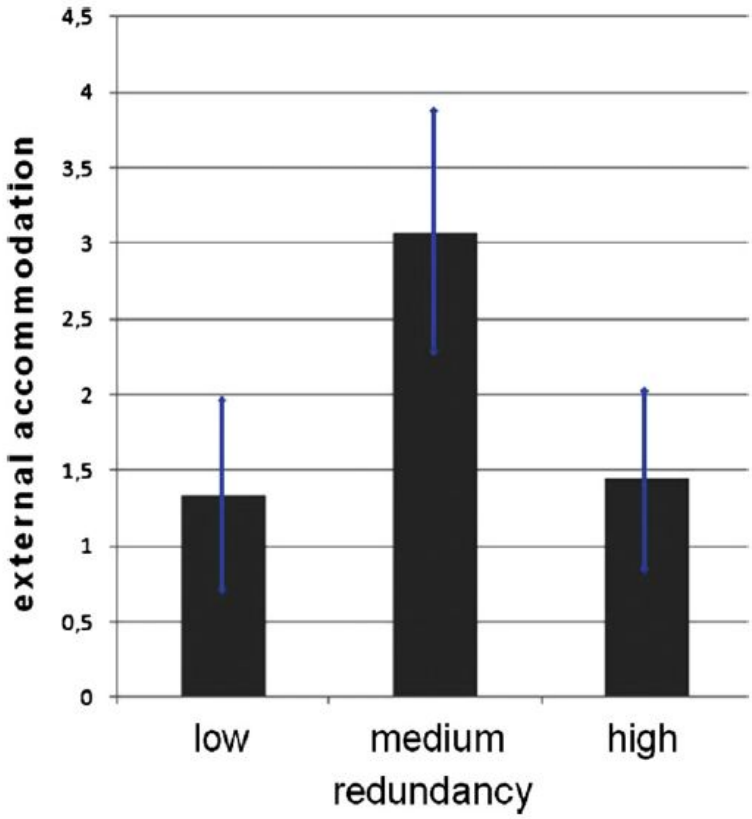

Figure 6: External accommodation depending on the level of redundancy (low vs. medium vs. high); MKC12, p. 1054.

In a second experiment, the polarity of the provided information was manipulated by means of providing the arguments in a similar wiki to the one in the previous study either with four arguments in favor of a biological explanation of the causes of schizophrenia or four arguments that focus on a more 'social' explanation (high polarity) or two arguments from each group (low polarity). Again, participants were to work on the wiki for some time. Afterwards, participants in the high polarity condition showed better learning outcomes than those in the low polarity condition.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

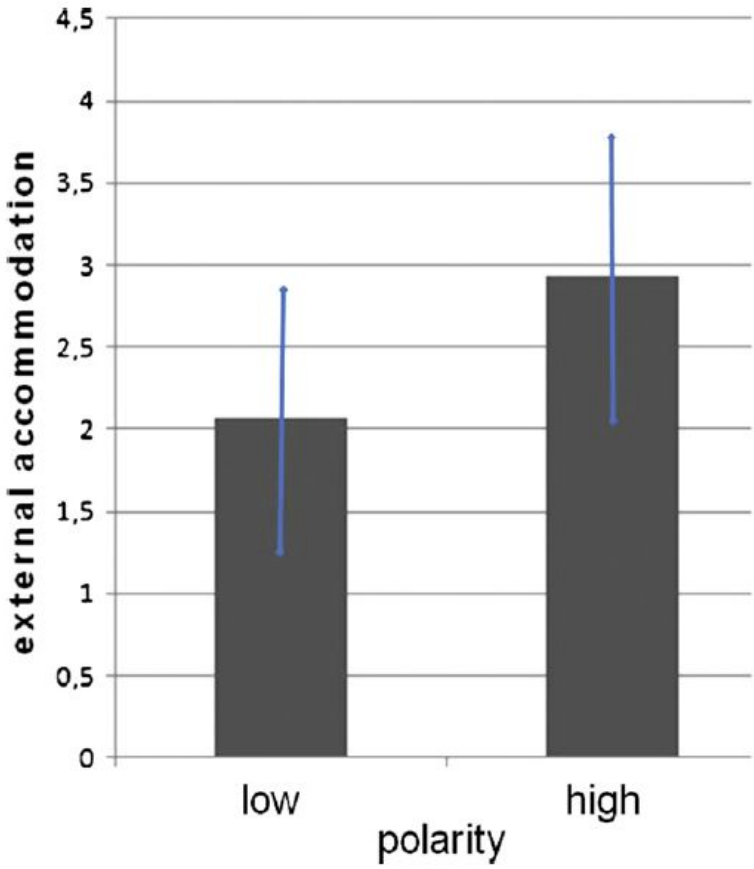

Figure 7: External accommodation depending on the level of polarity (low vs. high); MKC12, p. 1055.

Whereas the co-evolution model of learning and knowledge construction postulates that constructive friction is the driving force behind learning, another psychological process can then be supposed to stand in the way of learning: confirmation bias [JSFT01]. In the tradition of Festinger's [Fes57] classic work on cognitive dissonance, confirmation bias describes human beings' tendency to prefer information that is consistent with their own beliefs, attitudes, and knowledge over inconsistent information. Hence, one strategy to increase learning outcomes in online environments could be a 'forced exposure' of learners to inconsistent information, although a recent analysis of Facebook user's passing on of information [BMA15] showed that information that is inconsistent with the user's political ideology is very unlikely to be read and even more unlikely to be passed on to other users.

Schwind, Buder, Cress, and Hesse [SBCH12] devised two laboratory experiments to study in detail the effects of confirmation bias and friction on learning outcomes. In both experiments, participants first received information on the then relatively little known topic of 'neuroenhancement' and, after reading the material, they indicated their attitudes towards the topics. In a second step, participants were either recommended through a supposed

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

recommender system one out of eight arguments which was either in line with their own attitudes, that ran contrary to their attitudes, or they received no recommendation at all. In the first experiment, an online, experiment, participants had to indicate their attitude towards neuroenhancement again towards the end of the experiment; in the second experiment, a laboratory study, participants were asked to rate whether the recommended argument presented rather a minority or majority view; additionally, they had to recall in the end as many arguments as possible and they were asked to justify their opinion in a written statement.

In both experimental conditions, participants preferred reading arguments that were consistent with their own attitude. However, study 1 showed that those participants who were recommended an argument that went against their beliefs showed in the end more moderate attitudes to the topic of neuroenhancement. This findings resemble older studies by Shook and Fazio [SF09] who showed that surprising feedback reduces biases that result from over-certainty. Study two showed that preference-inconsistent recommendations lead to a more balanced recall of arguments and lead to more new arguments in the justification that was written by the participants; this finding can be interpreted as an indicator of a higher degree of divergent thinking. Overall, these experiments show that countering confirmation bias can be an effective means to stimulate learning processes.

The laboratory experiments in this section have their strengths and weaknesses. Their strength is certainly the possibility to control for unwanted side-effects in a laboratory environment: Apart from the elements that are manipulated and systematically varied (such as for example the degree of redundancy), all participants are presented with the same material under very similar circumstances. Of course, such data in turn suffers from the problem that the whole situation is artificial and the relevance of the findings for 'real-life' problems can be a problem. Another problem is that such laboratory experiments can only capture a relatively small time-span within a learning process, whereas usually learning is a rather lengthy endeavor. Furthermore, the co-evolution model posits that knowledge construction is the 'epiphenomenon' of learning: Whereas individuals try to resolve cognitive conflicts by means of communication, they do at the same time contribute to collaborative knowledge construction in online environments. This aspect of the co-evolution model is very difficult to study using laboratory experiments. Consequently, we will describe different approaches to 'big data' studies based on real behavior traces from knowledge building platforms such as Wikipedia in the following paragraphs.

'Big Data' Studies on Knowledge Construction

Confirmation bias and productive friction in Wikipedia. In the following paragraph, we will present different studies that studied learning and knowledge construction in Wikipedia using 'authentic' behavior traces such as the article's and users' history of previous edits. The first study [JKHHMCur], employed a 'semantic' approach insofar as users' edits were classified by means of a supervised machine learning algorithm as being in favor, against, or neutral

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

towards a given topic (alternative vs. conventional medicine). The other two studies employed different forms of social network analysis.

The Semantic Approach. In a recent study, Jirschitzka et al. [JKHHMCur] devised a 'big data' study to validate the findings from the laboratory research described beforehand [KMC11; SBCH12] in a realistic knowledge building environment: the Wikipedia platform.

The study focused on the struggle between alternative and conventional scholarly medicine in Germany. All edits of all articles within the categories of Alternativmedizin (alternative medicine) and Ernährung (nutrition) of the German language Wikipedia were sampled; all in all, more than 70.000 edits from 398 articles were crawled, processed and analyzed. For further analysis, a supervised machine learning algorithm was trained to distinguish all modifications into pro alternative medicine, pro conventional medicine, or neutral modifications. Based on these scores, a summary score for all edits (usually comprising of a number of modifications) was calculated that indicated in how far the edits displayed a view in favor or against alternative medicine or a neutral view.

Like in the laboratory studies, Wikipedia contributors did edit more frequently those articles that were in line with their view than those who weren't. Hence, confirmation bias does seem to have a stronger influence on information selection than the wish for productive friction. Interestingly, a higher ratio of non-neutral edits was predictive of a lower level of an imbalance in the sense of the final version of the article being more strongly in favor or against alternative medicine. Similarly, a higher number of contributors predicted more balanced articles - but only in case of pro-conventional medicine articles. In case of pro-alternative medicine articles, these articles were more balance in when they were edited by more heterogeneous contributors. Overall, it seems that friction and resulting discussions between contributors from different backgrounds and with different attitudes to medicine are a means to prevent a possible polarization of views [BMA15] or so called 'echo chamber effects' [DBZPSCSQ16].

These results mirror findings from the previous lab studies insofar as they show human beings' general preference for information that confirms their opinions, attitudes, and ideologies. Nevertheless, it still seems that productive friction between one's own views and others' views is necessary for learning and successful knowledge construction. Platforms such as Wikipedia enforce a certain degree of exposure to others' viewpoints by means of policies such as their 'neutral point of view policy'. Where does that leave us? There seem to be 'innate' human tendencies such as confirmation bias that apparently work against the goals of learning and knowledge construction - particularly in a digital world in which everybody is always just a mouseclick away from information that comforts his or her beliefs. Nevertheless, platforms such as Wikipedia seem to be rather effective in countering these tendencies [OCBN16]. One approach to further study the interplay between human beings' apparent 'laziness' in selecting preferentially confirming information and the necessity to be exposed to productive friction as means of learning something or of producing

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

knowledge, is the analysis of the history of 'actual' knowledge construction processes in Wikipedia using techniques of social network analysis. The following to studies fall into this category.

A social network analysis of knowledge trajectories in Wikipedia. The following study [KMHC10] is closely related to the previously discussed study by Kimmerle, Moskaliuk, & Cress [KMC11] and addressed as well the topic of possible social or biological causes of schizophrenia. The aim of this study was to visualize and to reconstruct the history of the respective Wikipedia articles by means of social network analysis [*SNA*; WF94]. For the analysis, the link structure of the Wikipedia article on schizophrenia and the two articles that according to experts' opinion represented best the different explanatory approaches (social vs. biological) was analyzed in a series of six annual cross-sectional 'snapshots' for the years 2003-2008. The Weaver software was used to calculate SNA traits such as centrality and density for all Wikipedia pages that were linked at the present time to the article about schizophrenia or the two other aforementioned articles. Furthermore, SNA allowed as well to calculate scores for the individual contributors with regard to the closeness or distance of their linking activities to the nodes that represent different views at the causes of schizophrenia.

Overall, the network did continually get more complex from 2003-2008: the number of links between the pages increased continuously (figure 3). Relatively early on in 2005, two clusters of closely interlinked articles appeared: one cluster that was related to the psycho-dynamic theory of schizophrenia one that was related to a biological explanation of schizophrenia (figure 4). Over time, a third cluster representing the more socially oriented diathesis-stress-model emerged (figure 5). Nevertheless, the 'social' cluster is closely related to the 'biological' cluster via a number of 'boundary-spanning articles'. This structure mirrors the fact that the diathesis-stress-model explicitly tries to incorporate biological and social causes of schizophrenia into a single model. In comparison, the psychodynamic cluster remained separated and isolated throughout the observation span with relatively little boundary-spanning articles. For the authors, it could be found that the endorsers of social or biological explanations shifted more and more towards the integrative diathesis-stress-model, whereas the endorsers of the psychodynamic explanations kept editing primarily the articles within their respective cluster.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

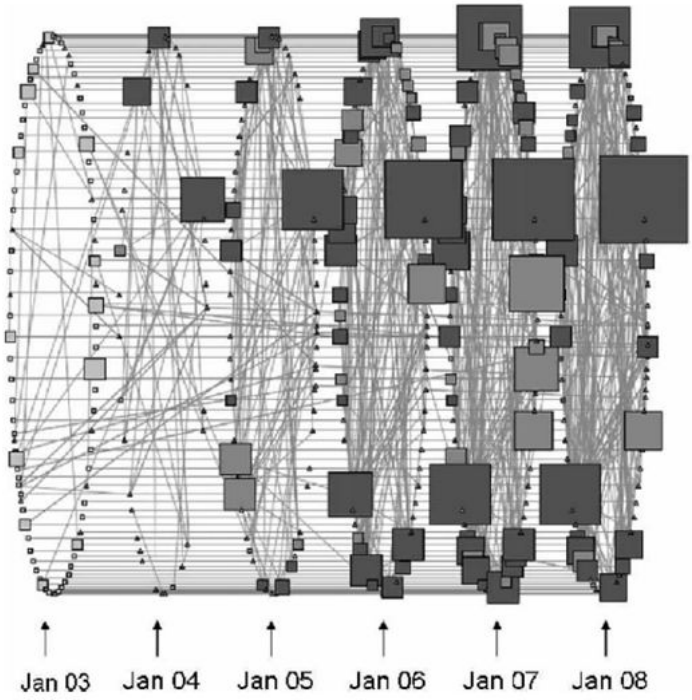

Figure 8: From KMHC10

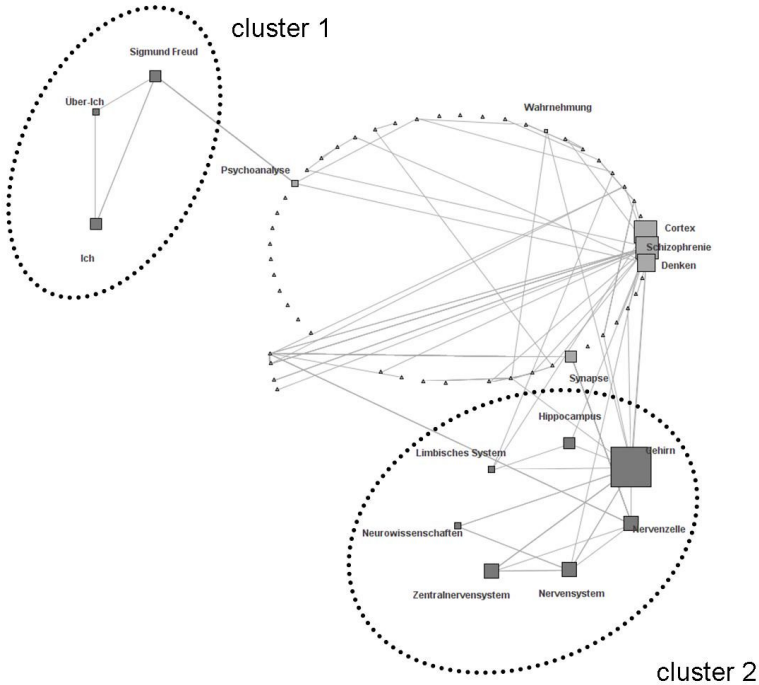

Figure 9: From KMHC10

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

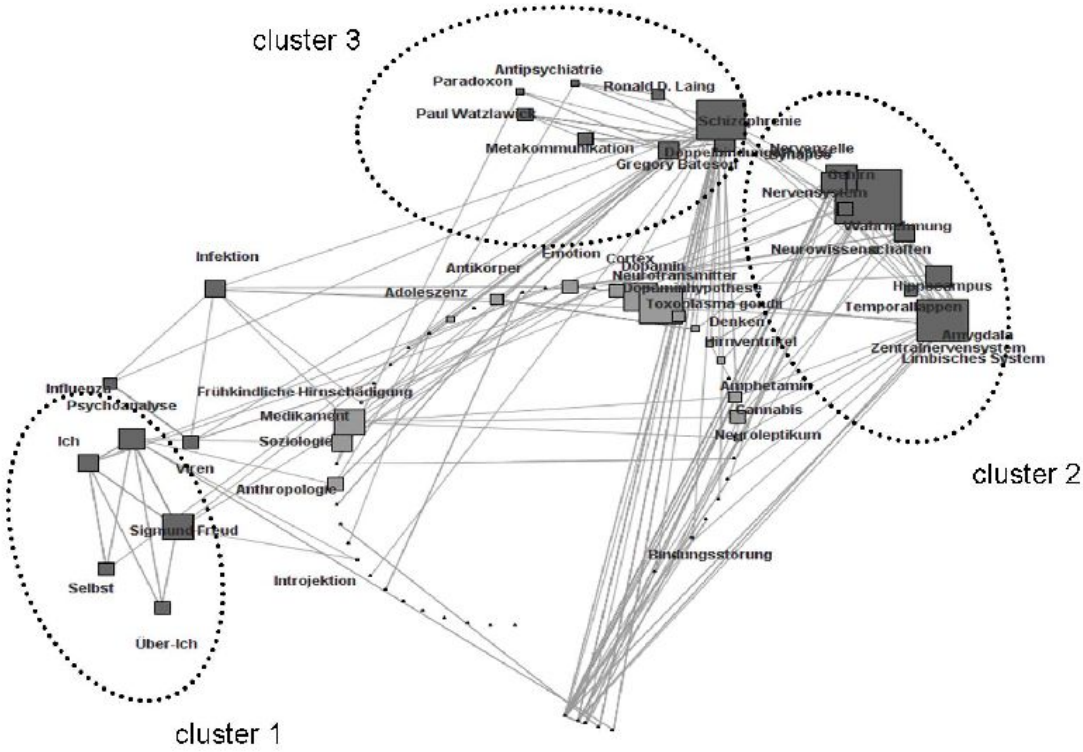

Figure 10: From KMHC10

Boundary Spanners and the Emergence of New Knowledge

Another series of studies [HC14; HMKC14] used a similar network analytic approach to identify the places where 'new knowledge' is created in Wikipedia and the persons creating this knowledge. The study focused on the Wikipedia domains 'psychology' and 'education' and again used snapshots of the link structure between related articles for the years 2006 to 2012. These two domains were chosen because although there is a substantial overlap between the two for example in form of articles on educational psychology, they do also represent clearly separate entities. All analyzed articles (all in all approximately 11.000) either fell into one of the two categories or both. The main hypothesis was that new knowledge—as a result of communication processes according to the co-evolution model—is primarily created in the vicinity of pivotal and important articles. Importance was measured for the articles categorized as either psychology or education articles via a network centrality measure; for articles that belonged to both categorizes, a betweenness measure was used as a proxy for the article's importance.

A first analysis [HC14] revealed that new knowledge in form of edits or newly created articles indeed centers either around articles with high levels of centrality within the respective domains or articles with high betweenness scores in cases of articles that fell under both categories. It should be noted though that whereas there was an exponential growth in

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

the number of new articles in the beginning of the observation period, a point of saturation was reached towards the end of the observation period.

Further analysis [HMKC14] revealed that both kinds of articles, those with high centrality scores and those with high betweenness scores, were authored by more experienced users (many previous edits). An analysis of the contributors' previous edits revealed that articles with high centrality scores, which were centerpieces of the respective domains of psychology or education, were primarily edits by 'specialists' who were only active in one of the domains. In comparison, articles with high betweenness scores, which spanned the two domains, were primarily authored by 'generalists' who regularly contributed to both domains. It seems that two kinds of knowledge creation processes can be observed here: on the one hand, there is the creation of conceptually new knowledge or the enrichment of existing knowledge in between the two domains by generalists who are well versed in both areas. On the other hand, an increase in very specialized areas at the center of the respective domains can be observed. This consolidation or refinement of existing knowledge is driven by specialists in the respective areas.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Empirical Studies Within the AFEL-Project

In the following section, we will provide an overview of ongoing and planned research projects in collaboration with the other work packages. On the one hand, these studies will have an impact on the development of technologies to facilitate learning and knowledge construction in digital environments; on the other hand, these projects will produce scientific insights into the learning processes in digital environments, which will be disseminated in scientific journals.

Effects of User Characteristics such as Heterogeneity on the Quality of Wikipedia Articles

This study continues the aforementioned line of studies on the role of productive friction for (everyday) learning - or the lack thereof. It will focus on one of the more negative consequences of the information age: The emergence of virtual 'echo chambers' wherein like-minded members of non-mainstream communities only share and perceive information that reinforces their beliefs; as a consequence, such communities are prone to radicalization effects. On the one hand, this project will enhance our understanding of the collaborative processes of learning and knowledge construction. On the other hand, we will develop software tools to estimate a topic's controversiality and to monitor discursive processes in collaborative knowledge construction scenarios.

The Echo Chamber Effect. Enlightenment philosophers (e.g. Diderot) believed that the availability of a comprehensive easily accessible collaboratively created and administrated knowledge corpus in form of an 'encyclopedia' should bring the end of superstition and ignorance and would consequentially lead to the disappearance of radicalism and violent conflicts. In the late 20th / early 21st century, this optimism has apparently faded away quite a bit. We observe that the huge increase in available knowledge resources via the internet has not led to such unanimously positive consequences. We still have wars and radicalism and we arguably have more issues with man-made catastrophes than at any time before. Why is that so? One possible reason may be related to what psychologists call "confirmation bias" [JSFT01], that is the rather ubiquitous tendency of human beings to be much more open (willing to take up, process, and pass on) toward information that is confirming one's beliefs and that is compatible with one's previous knowledge than one is open toward incompatible and disconfirming information. Another problem could result from human's tendency to bond together with 'like-minded' others and to form communities based on mutual understanding and shared knowledge resources [BG99; BG08]. As a result, the vast availability of knowledge and the technical possibilities for collaborative knowledge construction could (in the worst case) instead of mutual understanding of all humanity only cause the formation of more and more isolated and radical 'echo chambers' [DBZPSCSQ16]: Like-minded individuals form knowledge communities within which only knowledge that is compatible with the group's beliefs and ideology is taken up and echoed in comments and discussions.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

From the perspective of our model of individual learning and collective knowledge construction [e.g. CK08], such echo chambers are very unfavorable because of a lack of irritation of both the cognitive and the social system.

Objectives and hypotheses. The main objective of this study is to shed light on such processes of isolation and in contrast openness to others' by using the edit histories of different Wikipedia articles. We aim at understanding the differences between more closed and more open communities of people who contribute to an article. We will analyze as well other antecedents (predictors) of positive and negative developments in the construction of Wikipedia articles. We hypothesize that more openness prevents radicalization and leads to a larger growth of knowledge and to higher quality knowledge. This may be the case in particular, if a topic is controversial.

More concretely, we hypothesize that Wikipedia articles with more diverse contributors (in the sense that they contributed to a larger number of different Wikipedia domains before) grow faster, last longer (less risk of being closed by admins), and produce higher quality knowledge, at least whenever the topic is controversial in the sense that there are different opinions or contested areas of knowledge within the respective domain. We will use Wikipedia template messages or 'tags'[4] as proxies for an article's quality: Wikipedia articles that are tagged as 'neutral point of view violation' or that are marked for 'speedy deletion', can be regarded as low quality articles (at least at the point in time, when they were tagged); in contrast, articles that receive recommendations and distinctions constitute cases of high quality articles.

We already know from previous studies that the number of contributors at time 1 predicts the number of contributors at time 2 [e.g. Per14]: The more people contribute to a Wikipedia article or a knowledge domain, the larger is the probability that the number of contributors will grow in the future (this is sometimes called the 'Matthew-effect'). In contrast, the effects of heterogeneity among authors on the growth and survival of article or domains at Wikipedia was to our knowledge not studied extensively before.

Method and Operationalizations. In our study, we will (for a start) focus on the articles within the Wikipedia domain 'health and fitness'[5], tying in with some of the aforementioned studies [e.g., HC14]. Within these articles, we will identify the respectively most recent occurrence of the template messages 'neutrality and factual accuracy [violation]', 'contradiction and confusion', and '(speedy) deletion' as indicators of an article's low quality; as indicators of a high quality, we will in the same way identify the respectively most recent occurrence of the tags 'featured content' and 'featured and good article' in the article's edit history. In all cases, we will then study a certain number (e.g. 300) of edits that

[4] https://en.wikipedia.org/wiki/wikipedia:template_messages ; last retrieved Oct. 11th, 2016.
[5] https://en.wikipedia.org/wiki/Portal:Contents/Health_and_fitness ; last retrieved Oct. 11th, 2016.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

immediately precede the application of these tags. From these tags, we will then extract a number of contributors and discourse characteristics such as:

- The total number of individual contributors.
- The previous experience of the contributors (the average number of contributions to all Wikipedia articles up to this point in time among the individual contributors).
- The homo- or heterogeneity of the contributors with regard to the totality of edits they made in all Wikipedia articles up to this point in time; more homogeneous groups of contributors have primarily contributed to articles within the same Wikipedia domains.
- The ratio of 'productive edits' (in the sense of adding text etc.) and 'destructive edits' (deletions) as a proxy for discussions and 'edit wars'.
- The temporal density of edits (edits/time) as a proxy for an article's centrality and popularity.

The articles' controversiality will be assessed by comparing linguistic similarities between digital artifacts on the topic and by studying the controversiality of contribution on this topic on the webpage www.reddit.com by means of studying the contributions' up- and down-votes.

Analysis and outlook. At the moment (November 2016), members of work package WP2 are developing software tools to analyze a topic's degree of controversiality; we are also already running pre-studies for example on the prevalence of the aforementioned in order to devise a concrete sampling scheme. First analyses of actual data will take place in early 2017.

The predictive power of the aforementioned contributor and discourse characteristics with respect to the aforementioned tags will be assessed using logistic regression analysis. We expect for example that user experience as well as the total number of individual contributors are more frequent in the edits preceding awards and distinctions than in those preceding negative template messages. At least for articles on more controversial topics, a certain degree of heterogeneity among the contributors should be predictive of positive tags, whereas a high degree of homogeneity should be typical for negative template messages.

On the one hand side, this study addresses issues that have been centerpieces in the academic debate on social media over the last years [e.g., BMA15; DBZPSCSQ16]. On the other side, this study does also have implications for the design of resource-centric digital learning environments such as also Bibsonomy and Slidehare. For example, based on the analyzed predictors of positive and negative developments in collaborative knowledge construction, warning mechanisms in the form of software tools can be developed to provide contributors and administrators with feedback on current developments. The software tools developed by the L3S to determine a topic's degree of controversiality, will also be helpful for designers and administrators of Websites for collaborative knowledge construction. They are as well a pivotal for another planned study, in which we will together with colleagues from work

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

package WP3 attempt to visualize different ‘textures’ of knowledge (please see the following paragraph).

Visualizing Textures of Knowledge

The following project takes up our previous reasoning on the need to take into account social processes in the negation of meaning into account when studying everyday learning. Together with colleagues from work package WP3, we will develop visualization tools that provide learners in digital environments with feedback on the digital artifacts they are using.

Textures of knowledge. As we pointed out in the introduction, many philosophers and social scientists doubt that it is possible to define ‘objective’ criteria for (empirical) ‘truth’ beyond all doubt (the old dream of the philosophical ‘positivists’). Hence – at least partly – knowledge seems to be a social phenomenon, if only because concepts of ‘justification’ and ‘truth’ rely on negotiated social consensus .

If we adopt a (moderate or non-radical) constructivist position [Gad10], we may consequently believe that there are different ‘textures’ of knowledge – often within the very same domain or topic: Some knowledge elements are uncontested and shared by (almost) everyone within a ‘knowledge community’ (one could think about using the term ‘fact’ for these cases). On the other hand side, there is contested knowledge: Different social groups regard different beliefs as true and justified. In contrast to what is usually called an ‘opinion’ or an ‘attitude’, these beliefs do refer to factual claims (e.g. homeopathy’s efficacy has / has not been proven). On the other hand, some knowledge may be fragile in the sense that at the present time, there is no generally agreed consensus because of a lack of data (under-determination) or contradictory data (over-determination). Finally, more abstract and comprehensive systems of beliefs (for example in form of ideologies) can come into play. These beliefs are not characterized through factual claims, but refer to truth statements on a more abstract experiential or ‘philosophical’ level. Metaphorically speaking, some knowledge is ‘hard’ and ‘solid’ whereas other knowledge is viscose like clay or deceptive as quicksand and other forms of knowledge may be elusive as smoke or thin air.

In parallel, some knowledge elements are more central or peripheral for a given topic to others. Hence, a two- or three-dimensional model of knowledge could arise from these assumptions. The model will take into account on the one hand the centrality of a certain knowledge element for a given topic and on the other hand the knowledge element’s fragility/controversiality and (maybe or if possible) its relatedness to overarching ideologies. The software tools for assessing a topic’s controversiality, which were developed by members of work package WP2 for the aforementioned project on the ‘echo chamber effect’, will be used in this project as well.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Objectives and time frame. We aim at a comprehensive visualization of the 'texture of knowledge' on a given topic, which takes into account both forementioned dimensions: a knowledge element's centrality and its fragility. In the case of fragility, we will also analyze ways to visualize the relations between certain knowledge elements and overarching ideologies and other belief structures.

As of end of the 2016, members of work packages WP2 are currently developing software tools to measure controversiality and fragility. Some theoretical groundwork for this project was also presented by members of work package WP3 in deliverable D 3.1. During the year 2017, more concrete steps will be taken to realize this project and a concise schedule for the next steps will be set up.

Improving Learning Using Achievement Priming in Crowdsourcing Microtasks

In this project, we support work package 2 (WP2) in studying the effectiveness of achievement priming in crowdsourcing marketplaces such as Amazon's Mechanical Turk[6] or CrowdFlower[7]. Crowdsourcing platforms can be considered as a novel learning environment with unique constraints as shown in previous work [GFK15]. Achievement priming is the activation of participants' achievement motivation by means of subtle semantic or pictorial cues [e.g., Eng09]. Achievement motivation can be defined as 'the need for success or attainment of excellenc' [GHDur]. In a first series of studies, different forms of achievement primes were used to improve worker retention and performance in information finding tasks with varying levels of task difficulty. For example, in some experimental conditions, crowdworkers were just passively shown inspirational quotes such as "Winning isn't everything, but wanting to win is. - Vince Lombardi", whereas in other conditions participants had to actively find out and enter for example Vince Lombardi's middle name (Thomas).

The results show that most of the crowdworkers are rather achievement motivated; high 'chronic' levels of achievement motivation (high trait achievement motivation) are correlated to higher retention rates and better quality work. Overall, achievement priming had a positive effect on work quality and retention rates. We also found that a result of higher retention rates was improved learning among workers. We are currently planning further studies in which we will apply the achievement priming paradigm to other environments such as collaborative learning websites. On the theoretical side, studies on the effects of motivational dispositions allow for better understanding of the incentives in collaborative digital environments.

[6] http://www.mturk.com/
[7] http://www.crowdflower.com/

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Further Research

In addition to the two aforementioned project, which are as of now (November 2016) already ongoing or in a rather concrete stage of planning and preparation, we are also discussing a number of other possible research projects with our project partners. These projects are still in the early stages of conceptual planning.

One of these projects addresses the extraction of personality traits from behavior traces in digital learning and knowledge construction environments. As explained beforehand, research on Facebook activities such as status updates, comments, and likes has shown that it is very well possible to develop models to capture psychological characteristics such as personality traits, intelligence, and sexual orientation, if it is possible to relate the behavioral data to established data sources, for example in form of validated personality tests [KSG13]. Hence, we already know that short informal communication such as postings and status updates contains information on the contributors personality; but how is the situation with more formal contributions such as edits of Wikis and collaborative knowledge construction environments? To our best of knowledge, this research question so far has not been investigated in a systematic way.

Within this research project, we will first identify resource-centric platforms that allow for the collection of different types of user data, for example in form of contributions to Wikis or other knowledge resources and forum postings and other acts of informal communication. In a second step, we will offer users of the platform money in exchange for their willingness to fill in a validated personality questionnaire and to allow for the connection of this data with their activity data that can be mined from the platform. In a third step, the algorithms developed by Kosinski, Stillwell[8], and colleagues will be used to extract personality features from the participants' activity data. These scores will be correlated with the actual personality questionnaire scores. We expect that personality features that were extracted from informal communication are more highly correlated to actual personality test scores than more formal contributions such as edits of Wikis and other knowledge resources. Still, it may be possible that even these very formalized acts of communication bear linguistic traces of a contributor's personality. More concrete planning and preparation is planned for early 2017.

In another project, we focus on the motivational factors which influence the knowledge construction processes in online communities. Within this framework, motivational factors are operationalized based on the Terror Management Theory [TMT; GPS86]. Here, motives are considered to be intrinsic factors because they do not originate from external rewards. TMT suggests that awareness of psychological and physical vulnerability leads individuals to act more conservatively as a consequence of high levels of terror and discomfort. Most often, mortality salience is manipulated in TMT experiments. When people are reminded of their

[8] http://applymagicsauce.com/ , last retrieved Oct. 12th, 2016.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

mortality, they tend to stick to a more familiar environment [JA12]. Research on the effects of web-induced mortality salience shows that participants who were exposed to death-related content, displayed different preference patterns for online content compared to a control group [CE14; FFPD08; Y08]. Our aim is to extend this theoretical framework to the context of everyday online learning to get more insight into online community users' behaviors including topic selection, willingness to contribute, and learning outcomes.

Another theoretical perspective which we would like to take into account afterwards addresses the interplay of motives and traits. Socio-motivational models, like *Motivated Social Cognition* [JGKS03] and the *Dual-Process Motivational Model* [DS10] identify the motives of TMT as the underlying elements of ideological personality traits (e.g. authoritarianism). We will use these state-trait models to explain patterns in everyday online learning behavior.

We have already initiated this project with a lab experiment, where we measure the effect of motivational factors on Wikipedia topic preferences. Preliminary findings show that mortality induction significantly affects the willingness to contribute to an online community and that it leads to particular patterns of preferences in terms of topic familiarity and controversiality. Further lab studies will include other factors, such as affective reactions, traits, and need for certainty; in these studies, we will focus on outcome behaviors such as collaborative contribution and actual learning activities. Findings of these studies will provide a solid analytical ground, which would allow us to translate into technologies to support online learning activities.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Applications for Digital Collaborative Learning and Knowledge Construction Environments

Apart from the aforementioned theoretical insights, we will take into account the application side of our research as well. Whereas we have hinted towards possible applications of our ideas throughout the document, we will in the following paragraphs present a summary of the main 'practical' benefits we will derive from our research.

Possible applications of our Research on Echo Chamber Effects

In the course of the aforementioned research project on possible echo chamber effects in Wikipedia, members of work package WP2 will develop several software tools, which can be of use for the contributors and administrators of collaborative knowledge construction environments.

Our research on antecedents of positive (distinctions) and negative (deletion; violation of Wikipedia rules such as the neutral point of view policy) trajectories of Wikipedia articles will allow us to identify overarching abstract models of an article's trajectory based on user characteristics such as their experience and heterogeneity with regard to attitudes and previous knowledge on the one hand side and discourse characteristics such as the intensity of discussions on the other side side. We aim at testing these algorithms at other resource-centric websites such as Slideshare and Bibsonomy as well to evaluate their usefulness in providing the creators and administrators as well as the users of digital learning environments with feedback on ongoing beneficial and detrimental developments with regard to the collaborative construction of digital artifacts.

Members of work package WP2 are currently developing algorithms to identify the controversiality of topics by means of a linguistic comparison of the (dis-)similarity between resources relating to the topic. Based on these models, it will be possible as well to estimate the biasedness of a given artifact concerning a controversial topic by analyzing in how far all conflicting viewpoints are represented in the artifacts. Information on the potential biasedness of digital artifacts can help learners in digital environments to identify relevant and helpful resources; the same algorithms can as well serve as a warning system for the administrators of resource-centric digital learning environments.

Our measures of user homogeneity based on similar topic preferences can be used as well to bring together learners who are interested in similar topics via recommender engines. In the context of collaborative knowledge construction, it can be helpful as well to connect people with diverging views of a topic. In this case, recommendations will be based on homogeneity regarding topic preferences, but heterogeneity with regard to certain contested knowledge elements.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Using Visualization Techniques to Facilitate Cooperation in Computer-Supported Collaborative Learning Environments.

Whereas the aforementioned research project on potential echo chamber effects in Wikipedia is already in a very concrete stage of planning, our ideas on visualizing the texture of knowledge are still in the early stages of development.

Here, our aim is to use the power of state of the art visualization tools [D3.1] to provide learners and operators of digital learning environments with powerful tools to facilitate them in understanding complex and controversial knowledge elements with regard to a given topic. These tools will make it easier to discern consensual knowledge from contested knowledge, which is held to be true within certain social communities, whereas other communities oppose this view.

These software tools could be helpful as well in structured and planned learning activities such as massive open online courses (MOOCs). It is for example imaginable that interactive dashboards [D3.1] can be used to make learners aware of controversies, ambiguities, and potential sources of misunderstandings.

Further Applications

The aforementioned research on achievement priming in crowdsourcing tasks can on the one hand show ways to improve the retention rates in and the quality of crowdsourcing tasks. Apart from that, priming in general can be an interesting means of supporting processes of learning and knowledge construction in general. We will investigate ways to implement similar procedures in other task-oriented platforms.

Our planned research on the extraction of personality characteristics from behavior traces in digital environments could as well play a role for example in devising person-centered and tailormade recommender systems and customized dashboards in learning environments [D3.1].

The project focusing on the impact of motivational factors on the learning behaviors will be beneficial for the personalization of recommender systems and dashboards. For example, the findings of experiments can help in the long run to optimize the search engine of a platform to suggest interesting and relevant topics for the users. With the help of a big data model covering motivational and social factors consolidated by experimental findings, it would be possible to design platforms to best fit to the users` interests.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

Conclusion

The co-evolution model of learning and knowledge construction [e.g. CK08] serves as the theoretical framework for the AFEL-project. So far, it proved to be helpful in the definition of the data source taxonomy and the AFEL data scheme. It also serves as the theoretical background for our attempts towards a formal definition of learning activities [DHAKDur].

In the co-evolution model, learning takes place whenever a learner's cognitive system is irritated by means of novel or unexpected information and manages to integrate the new information into existing cognitive structures; the information comes from the social system, which serves as an environment for the cognitive system. Resulting changes in the cognitive system can—when communicated to the social environment—trigger in turn processes of collaborative knowledge construction. Hence, individual learning and collective knowledge construction are structurally coupled.

When applied to online learning environments, digital artifacts play a pivotal role in mediating between the cognitive system and the social system. This very broad and abstract concept of learning allows us to integrate as diverse online activities as taking part in structured online courses, self-studying a topic by reading Wikipedia articles, or learning how to knit a sweater by watching instructional videos under a single framework. In all cases, one challenge for work packages WP1 and WP2 will be to induce and reconstruct more abstract constructs such as 'learning trajectories' from concrete interactions between learners and artifacts in form of for example browsing activities or edits of a digital artifacts.

In continuation of previous research on the role of 'productive friction' in form of incongruities between learners' existing knowledge and the information they encounter in the process of learning, we are currently planning several empirical studies that will shed light on the applicability of our theoretical framework in the context of everyday learning in digital environments and which will further deepen our understanding of learning processes in general. A first collaborative project with members of work package WP2, which is already in the later on the way, investigates contributor and discourse characteristics that are predictive of positive and negative developments in collaboratively constructed Wikipedia articles. In particular, we will investigate whether so called 'echo chamber effects', which result from a too high degree of homogeneity in terms of e.g. knowledge, interests, and attitudes among Wikipedia contributors can have a negative influence on the quality of articles on controversial topics.

Our research will result in a number of applications in the form of software tools that will support learners in digital environments and the creators and administrators of digital environments alike. The insight from the aforementioned project on potential echo chamber effects in Wikipedia articles can be used to create software that monitors contributor and discourse characteristics in collaborative knowledge construction environments as means of

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

providing them and the site's administrators with feedback about potential detrimental and positive developments within the respective articles. As explained beforehand, members of work package WP2 will develop a procedure to assess the controversiality of topics and the 'biasedness' of digital artifacts relating to these topics based on text mining and data mining procedures.

Another important topic that will be addressed in the later stages of the project will be the visualization of fragile and controversial knowledge with the overarching goal to establish ways to visualize the different 'textures' of knowledge. Apart from the scientific merit, here as well, we aim at developing software tools that can be used for learners, contributors, and administrations in/of digital learning environments.

Our research on the effectiveness of achievement priming in crowdsourcing tasks together with the work package WP2 will help to shed light on possible incentives in task-centric digital environments. Here as well, a wide range of applications with regard to enhancing the productivity and quality of crowdsourcing tasks will be explored.

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

References

[OKC16] Oeberst, Aileen, Joachim Kimmerle, and Ulrike Cress, What is knowledge? Who creates it? Who possesses it? The need for novel answers to old questions, In Mass Collaboration and Education, pp. 105-124, Springer International Publishing, 2016

[PLH04] Paavola, Sami, Lasse Lipponen, and Kai Hakkarainen, Models of innovative knowledge communities and three metaphors of learning, Review of educational research 74, no. 4 (2004): 557-576

[DBZPSCSQ16] Del Vicario, Michela, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H. Eugene Stanley, and Walter Quattrociocchi, The spreading of misinformation online, Proceedings of the National Academy of Sciences 113, no. 3 (2016): 554-559

[CK08] Cress, Ulrike, and Joachim Kimmerle, A systemic and cognitive view on collaborative knowledge building with wikis, International Journal of Computer-Supported Collaborative Learning 3, no. 2 (2008): 105-122

[Arn02] Arnett, Jeffrey Jensen, The psychology of globalization, American psychologist 57, no. 10 (2002): 774

[HP13] Happer, Catherine, and Greg Philo, The role of the media in the construction of public belief and social change, Journal of social and political psychology 1, no. 1 (2013): 321-336

[Sfa98] Sfard, Anna, On two metaphors for learning and the dangers of choosing just one, Educational researcher 27, no. 2 (1998): 4-13

[ALS96] Anderson, John R., Lynne M. Reder, and Herbert A. Simon, Situated learning and education, Educational researcher 25, no. 4 (1996): 5-11

[Gre97] Greeno, James G., On claims that answer the wrong questions, Educational researcher 26, no. 1 (1997): 5-17

[LW91] Lave, Jean, and Etienne Wenger, Situated learning: Legitimate peripheral participation, Cambridge university press, 1991

[Eng99] Engeström, Yrjö, Expansive visibilization of work: An activity-theoretical perspective, Computer Supported Cooperative Work (CSCW) 8, no. 1-2 (1999): 63-93

[NT95] Nonaka, Ikujiro, and Hirotaka Takeuchi, The knowledge creation company: how Japanese companies create the dynamics of innovation, Oxford University Press, New York, USA (1995): 304

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

[Ber02] Bereiter, C, Education and mind in the knowledge society, Mahwah NJ (2002)

[KMCK15] Kump, Barbara, Johannes Moskaliuk, Ulrike Cress, and Joachim Kimmerle, Cognitive foundations of organizational learning: re-introducing the distinction between declarative and non-declarative knowledge, Frontiers in psychology 6 (2015)

[Leo78] Leontiev, Aleksei Nikolaevich, Activity, consciousness, and personality, (1978)

[SB94] Scardamalia, Marlene, and Carl Bereiter, Computer support for knowledge-building communities, The journal of the learning sciences 3, no. 3 (1994): 265-283

[WS00] Wenger, Etienne C., and William M. Snyder, Communities of practice: The organizational frontier, Harvard business review 78, no. 1 (2000): 139-146

[Pop72] Popper, Karl Raimund, Objective knowledge: An evolutionary approach, (1972): 81-106

[Pop79] Popper, Karl Raimund, Three worlds, Ann Arbor, University of Michigan, 1979

[Hol16] Holtz, Peter, How Popper's 'three worlds theory' resembles Moscovici's 'social representations theory' but why Moscovici's 'social psychology of science' still differs from Popper's critical approach, Papers on Social Representations 25, no. 1 (2016): 1.13-1.24

[HP09] Hakkarainen, Kai, and Sami Paavola, Toward a trialogical approach to learning, Transformation of knowledge through classroom interaction (2009): 65-80

[KMOC15] Kimmerle, Joachim, Johannes Moskaliuk, Aileen Oeberst & Ulrike Cress, Learning and collective knowledge construction with social media: A process-oriented perspective, Educational Psychologist, 50, (2015): 120-137

[VMU91] Varela, Francisco G., Humberto R. Maturana, and Ricardo Uribe, Autopoiesis: the organization of living systems, its characterization and a model, In Facets of systems science, pp. 559-569, Springer US, 1991

[Mat75] Maturana, Humberto R., The organization of the living: A theory of the living organization, International journal of man-machine studies 7, no. 3 (1975): 313-332

[Sim95] Simon, Martin A,, Reconstructing mathematics pedagogy from a constructivist perspective. Journal for research in mathematics education(1995): 114-145

[Ben12] Bennett, Elizabeth E., A Four-Part Model of Informal Learning: Extending Schugurensky's Conceptual Model, (2012), Adult Education Research Conference, Paper 3

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

[KSG13] Kosinski, Michal, David Stillwell, and Thore Graepel, Private traits and attributes are predictable from digital records of human behavior, Proceedings of the National Academy of Sciences 110, no. 15 (2013): 5802-5805

[MPS94] Mount, Michael K., Murray R. Barrick, and J. Perkins Strauss, Validity of observer ratings of the big five personality factors, Journal of Applied Psychology 79, no. 2 (1994): 272

[DHAKDur] d'Aquin, Mathieu, Peter Holtz, Alessandro Adamou, Stefan Dietze, and Joachim Kimmerle, Towards defining learning activities to detect them in users' online activity streams, *submitted for* the 2017 conference on Learning and Knowledge Construction (LAK) in Vancouver, Canada

[MD78] Mugny, Gabriel, and Willem Doise, Socio‐cognitive conflict and structure of individual and collective performances, European journal of social psychology 8, no. 2 (1978): 181-192

[DMP75] Doise, Willem, Gabriel Mugny, and Anne‐Nelly Perret‐Clermont, Social interaction and the development of cognitive operations, European journal of social psychology 5, no. 3 (1975): 367-383

[DDB07] Darnon, Céline, Sébastien Doll, and Fabrizio Butera, Dealing with a disagreeing partner: Relational and epistemic conflict elaboration, European Journal of Psychology of Education 22, no. 3 (2007): 227-242

[MKC12] Moskaliuk, Johannes, Joachim Kimmerle, and Ulrike Cress, Collaborative knowledge building with wikis: The impact of redundancy and polarity, Computers & Education 58, no. 4 (2012): 1049-1057

[JSFT01] Jonas, Eva, Stefan Schulz-Hardt, Dieter Frey, and Norman Thelen, Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information, Journal of personality and social psychology 80, no. 4 (2001): 557

[Fes57] Festinger, Leon, Cognitive dissonance theory, Newbury Park, California, Sage Publications (1957)

[BMA15] Bakshy, Eytan, Solomon Messing, and Lada A. Adamic, Exposure to ideologically diverse news and opinion on Facebook, Science 348, no. 6239 (2015): 1130-1132

[SBCH12] Schwind, Christina, Jürgen Buder, Ulrike Cress, and Friedrich W. Hesse, Preference-inconsistent recommendations: An effective approach for reducing confirmation bias and stimulating divergent thinking? Computers & Education 58, no. 2 (2012): 787-796

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

[SF09] Shook, Natalie J., and Russell H. Fazio, Political ideology, exploration of novel stimuli, and attitude formation, Journal of Experimental Social Psychology 45, no. 4 (2009): 995-998

[JKHHMCur] Jirschitzka, Kimmerle, Halatchliyski, Hancke, Meurers, & Cress, under review, A productive clash of perspectives? The interplay between articles' and authors' perspectives and their impact on Wikipedia edits in a controversial domain

[OCBN16] Oeberst, Aileen, Ulrike Cress, Mitja Back, and Steffen Nestler, Individual versus collaborative information processing: The case of biases in Wikipedia, In Mass Collaboration and Education, pp. 165-185, Springer International Publishing, 2016

[KMC11] Kimmerle, Joachim, Johannes Moskaliuk, and Ulrike Cress, Using Wikis for Learning and Knowledge Building: Results of an Experimental Study, Educational Technology & Society 14, no. 4 (2011): 138-148

[KMHC10] Kimmerle, Joachim, Johannes Moskaliuk, Andreas Harrer, and Ulrike Cress, Visualizing co-evolution of individual and collective knowledge, Information, Communication & Society 13, no. 8 (2010): 1099-1121

[WF94] Wasserman, Stanley, and Katherine Faust, Social network analysis: Methods and applications, Vol. 8, Cambridge university press, 1994

[HC14] Halatchliyski, Iassen, and Ulrike Cress, How structure shapes dynamics: Knowledge development in Wikipedia-A network multilevel modeling approach, PloS one 9, no. 11 (2014): e111958

[HMKC14] Halatchliyski, Iassen, Johannes Moskaliuk, Joachim Kimmerle, and Ulrike Cress, Explaining authors' contribution to pivotal artifacts during mass collaboration in the Wikipedia's knowledge base, International Journal of Computer-Supported Collaborative Learning 9, no. 1 (2014): 97-115

[BG99] Bauer, Martin W,, and George Gaskell, Towards a paradigm for research on social representations, Journal for the theory of social behaviour 29, no. 2 (1999): 163-186

[BG08] Bauer, Martin W., and George Gaskell, Social representations theory: A progressive research programme for social psychology, Journal for the theory of social behaviour 38, no. 4 (2008): 335-353

[Per14] Perc, Matjaž, The Matthew effect in empirical data, Journal of The Royal Society Interface 11, no. 98 (2014): 20140378

[Gad10] Gadenne, Volker, Why radical constructivism has not become a paradigm, Constructivist Foundations 6, no. 1 (2010): 77-83

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org

[Eng09] Engeser, Stefan, Nonconscious activation of achievement goals: moderated by word class and the explicit achievement motive? Swiss Journal of Psychology 68, no. 4 (2009): 193-200

[GFK15] Gadiraju, Ujwal, Besnik Fetahu, and Ricardo Kawase. "Training Workers for Improving Performance in Crowdsourcing Microtasks." In *Design for Teaching and Learning in a Networked World*, pp. 100-114. Springer International Publishing, 2015.

[GHDur] Gadiraju, Ujwal, Peter Holtz, and Stefan Dietze, Prime time! Improving worker retention in information finding microtasks through achievement priming, paper currently under review for the 2017 ACM CHI Conference on Human Factors in Computing Systems in Denver CO, USA

[GPS86] Greenberg, Jeff, Tom Pyszczynski, and Sheldon Solomon, The causes and consequences of a need for self-esteem: A terror management theory, In Public self and private self, pp. 189-212. Springer New York (1986)

[K06] Kuznetsov, Stacey, Motivations of contributors to Wikipedia, ACM SIGCAS computers and society 36, no. 2 (2006): 1

[JA12] Jost, John T., and David M. Amodio, Political ideology as motivated social cognition: Behavioral and neuroscientific evidence, Motivation and Emotion 36, no. 1 (2012): 55-64

[CE14] Chopik, William J., and Robin S. Edelstein, Death of a salesman: Webpage-based manipulations of mortality salience, Computers in Human Behavior 31 (2014): 94-99

[DS10] Duckitt, John, and Chris G. Sibley, Personality, Ideology, Prejudice, and Politics: A Dual‐Process Motivational Model, Journal of Personality 78, no. 6 (2010): 1861-1894

[FFPD08] Fransen, Marieke L., Bob M. Fennis, Ad Th H. Pruyn, and Enny Das, Rest in peace? Brand-induced mortality salience and consumer behavior, Journal of Business Research 61, no. 10 (2008): 1053-1061

[JGKS03] Jost, John T., Jack Glaser, Arie W. Kruglanski, and Frank J. Sulloway, Political conservatism as motivated social cognition, Psychological Bulletin 129, no. 3 (2003): 339

[Y08] Yoo, Chan Yun, Unconscious processing of web advertising: Effects on implicit memory, attitude toward the brand, and consideration set, Journal of Interactive Marketing 22, no. 2 (2008): 2-18

This document is part of the AFEL project funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687916 - see http://afel-project.org