Report from Dagstuhl Seminar 13361

Crowdsourcing: From Theory to Practice and Long-Term Perspectives

Edited by

Tobias Hoßfeld[1], Phuoc Tran-Gia[1], and Maja Vukovic[2]

1 **Universität Würzburg, DE,** [hossfeld,trangia]@informatik.uni-wuerzburg.de
2 **IBM TJ Watson Research Center, US,** maja@us.ibm.com

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 13361 "Crowdsourcing: From Theory to Practice and Long-Term Perspectives". Crowdsourcing is a newly emerging service platform and business model in the Internet. In contrast to outsourcing, where a job is performed by a designated worker or employee, crowdsourcing means to outsource a job to a large, anonymous crowd of workers, the so-called human cloud, in the form of an open call. Current research in crowdsourcing addresses the following issues: crowdsourcing as a novel methodology for user-centered research; development of new services and applications based on human sensing, computation, and problem solving; engineering of improved crowdsourcing platforms including quality control mechanisms; incentive design and gamification of work; usage of crowdsourcing for professional business; theoretical frameworks for evaluation. The topic on crowdsourcing may have a huge impact on the Internet and its technical infrastructure, on society, and the future of work. In short, crowdsourcing will be a guiding paradigm and form the evolution of work in the next years. Therefore, this seminar helps coordinating research efforts in the different communities. In five presentation and discussion sessions, the diverse aspects of crowdsourcing were elaborated. The topics of the sessions covered (S1) crowdsourcing in general, (S2) industry use cases, (S3) crowdsourcing design and engineering, (S4) programming and implementing crowdsourcing, (S5) applications of crowdsourcing.

The major interests of the seminar participants were then focused in four different working groups on (W1) long-term perspectives & impact on economics in five years, (W2) theory – taxonomy and dimensions of crowdsourcing, (W3) industry use cases, (W4) crowdsourcing mechanisms and design. In parallel to this seminar, a topically related seminar on "Cloud-based Software Crowdsouring", organized by Michael N. Huhns, Wei Li, Martin Schader and Wei-Tek Tsal,(Dagstuhl Seminar 13362) took place. Therefore, a joint late night session was organized to discuss crowdsourcing with respect to ethics and its relation to social computation.

Seminar 01.–04. September, 2013 – www.dagstuhl.de/13361

1998 ACM Subject Classification H.1.2 User/Machine Systems, H.4 Information Systems Applications, H.5 Information Interfaces and Presentation, H.5.1 Multimedia Information Systems, H.5.3 Group and Organization Interfaces, J.0 General Computer Applications, K.4 Computers and Society, K.4.3 Organizational Impacts

Keywords and phrases Crowdsourcing, Human Computation, Human Cloud, Applications, Industry Use Cases, Crowdsourcing Design, Mechanisms, Engineering, Practical Experience

Digital Object Identifier 10.4230/DagRep.3.9.1

Crowdsourcing: From Theory to Practice and Long-Term Perspectives, *Dagstuhl Reports*, Vol. 3, Issue 9, pp. 1–33

Editors: Tobias Hoßfeld, Phuoc Tran-Gia, and Maja Vukovic

Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Executive Summary

Tobias Hoßfeld
Phuoc Tran-Gia
Maja Vukovic

Over the past several years crowdsourcing has emerged as a new research theme, but also as a new service platform and Internet for harnessing the skills of the large, network-connected crowd on-line. Whilst the research community has not just yet recognized crowdsourcing as an entirely new discipline, many research challenges remain open and need to be addressed to ensure its successful applications in academia, industry and public sectors. Crowdsourcing research intersects many existing domains and brings to the surface new challenges, such as crowdsourcing as a novel methodology for user-centered research; development of new services and applications based on human sensing, computation and problem solving; engineering of improved crowdsourcing platforms including quality control mechanisms; incentive design and gamification of work; usage of crowdsourcing for professional business; theoretical frameworks for evaluation. Crowdsourcing, as a new means of engaging human capital online is increasingly having an impact on the Internet and its technical infrastructure, on society, and the future of work.

With crowdsourcing gaining momentum and becoming mainstream, the objective of this Dagstuhl seminar was to lead coordination of research efforts in the different communities, especially in US currently leading the crowdsourcing market and in Europe. The seminar engaged experts from the different research fields (e.g. sociology to image processing) as well as experts from industry with a practical background on the deployment, operation or usage of crowdsourcing platforms. From industry, real-world problem statements, requirements and challenges, position statements, innovative use cases, and practical experiences are tackled and discussed. The collection and analysis of practical experiences of the different crowdsourcing stakeholders were key outcomes of the Dagstuhl Seminar. The seminar was structured so that the participants use existing use cases, as a driver in the discussion to envisions future perspectives of this domain. To move forward, we identified the need for a common terminology, classification and taxonomy of crowdsourcing systems, as well as evaluation frameworks; and have already proposed a blueprint of the same. The impact of crowdsourcing from different perspectives has been discussed, by participants' viewpoints stemming from societal, business, economic, legal and infrastructure perspectives.

From platform provider side, Nhatvi Nguyen (Sec. 3.11) showed the actual challenges in operating a crowdsourcing platform. As industry use case, the example of enterprise crowdsourcing was presented by Maja Vukovic (Sec. 3.14), where the rapid generation of a snapshot of the state of IT systems and operation is conducted by means of crowdsourcing. This allows for massive cost savings within the company by uncovering knowledge critical to IT services delivery. Crowdsensing is another industry use case presented in the seminar by Florian Zeiger (Sec. 3.15). Environmental sensing in the area of safety and security was discussed from industry point of view along with the challenges and open questions, e.g. user privacy, data quality and integrity, efficient and reliable data collection, as well as architectural decisions and flexible support of various business models. A concrete application for crowdsensing is radiation sensing as shown by Shinichi Konomi (Sec. 3.7).

Beyond this, there were also discussions on multimedia related use cases. Crowdsourcing can be efficiently used for describing and interpreting multimedia on the Internet and allows to better address other aspects of multimedia with meaning for human beings. Martha

Larson (Sec. 3.10) provided examples of these aspects like the emotional impact of multimedia content, and judgments concerning which multimedia is best suited for a given purpose. Klaus Diepold (Sec. 3.6) applied crowdsourcing to move subjective video quality tests from the lab into the crowd. The resulting ratings are used to train mathematical model for predicting subjective quality of video sequences. Multivariate data analysis tools are recommended to incorporate contextual information to further validate the mathematical model. Vassilis Kostakos (Sec. 3.8) showed that the data quality of appropriate subjective tests may be increased by using public displays and touch screens in cities compared to online surveys. While gamification pops up as buzzword aiming among others at increased data quality, Markus Krause (Sec. 3.9) mentioned that the player should be put first i.e. the desires of player are paramount. In particular, task and game ideas need to be able to be linked, while fun has to be the main motivator for the game.

General approaches to improve crowdsourcing and the resulting data quality were a topic of interest by several participants. Gianluca Demartini (Sec. 3.5) proposes to model workers in the crowd as basis for quality assurance mechanisms. Alessandro Bozzon (Sec. 3.2) demanded for better conceptual abstractions for crowd tasks and processes design and (automatic) generation; better understanding of crowds properties such as (soft and hard) skills, reliability, availability, capacity, precision; and better tools for measuring and driving worker engagement. Cristina Cabanillas (Sec. 3.3) considered the human resource management aspects starting from workflows to crowdsourcing. Abraham Bernstein (Sec. 3.1) discussed human computers as part of computational processes, however, with their own strengths and issues. The three traits on human computation, that are motivational diversity, cognitive diversity, and error diversity, are embraced as strengths instead of weaknesses. While the main focus of the seminar was on technical challenges, the potential impact and long-term perspectives were discussed from an interdisciplinary point of view too, given the social and human aspects of crowdsourcing. Those issues were also raised by Phuoc Tran-Gia (Sec. 3.13) and Joseph G. Davis (Sec. 3.4).

Overall there were 22 participants from 9 countries and 16 institutions. The seminar was held over 2.5 days, and included presentations by researcher and specific hands-on discussion sessions to identify challenges, evaluate viewpoints and develop a research agenda for crowdsourcing. While the abstracts of the talks can be found in Section 3, a summary of the discussions arising from those impulse talks is given in Section 7. Additional abstracts and research statements without any presentation in the plenary are also included in the report in Section 4. The different aspects of crowdsourcing were discussed in more detail in four different working groups formed during the seminar: (W1) long-term perspectives & impact on economics in five years, (W2) theory: taxonomy and dimensions of crowdsourcing, (W3) industry use cases, (W4) crowdsourcing mechanisms and design. The summary of those working groups can be found in Section 5.

Please note that a related seminar on "Cloud-based Software Crowdsouring" (Dagstuhl Seminar 13362), organized by Michael N. Huhns, Wei Li, Martin Schader and Wei-Tek Tsai, took place in parallel to this seminar. We held a joint social event and a session on discussing research challenges and planned publications. In this late night session, on one hand ethical issues in the area of crowdsourcing were raised in a stimulus talk by Martha Larson (TU Delft). On the other hand, Munindar P. Singh (North Carolina State University) intended to provoke with his talk on the critique of current research in the area of social computing and crowdsourcing. A summary can also be found in Section 7.

A comprehensive list of open problems and challenges in the area of crowdsourcing as observed and stated by the participants is another key outcome of the seminar which is provided in Section 6.

2 Table of Contents

Executive Summary
Tobias Hoßfeld, Phuoc Tran-Gia, and Maja Vukovic 2

Overview of Talks

Programming the Global Brain with CrowdLang and CrowdOS – Challenges and Ideas
Abraham Bernstein 6

Crowdsourcing Engineering
Alessandro Bozzon 7

Human Resource Management: From Workflows to Crowdsourcing
Cristina Cabanillas 7

Human Computation and Crowdsourcing
Joseph Davis 8

Next-Generation Micro-task Crowdsourcing Platforms
Gianluca Demartini 8

Data Analysis Tools for Crowdsourced Video Quality Tests
Klaus Diepold 8

Supporting Exploration and Rapid Development by Citizens to Collect the Right Information through Crowd Sensing
Shinichi Konomi 9

Crowdsourcing Beyond the Desktop: Experiences and Challenges
Vassilis Kostakos 9

A Homo Ludens in the Loop
Markus Krause 10

Towards Responsible Crowdsourcing
Martha A. Larson 10

Crowdsourcing Challenges from Platform Provider's Point of View
Nhatvi Nguyen 11

Social Protocols as a Basis for Social Computing
Munindar P. Singh 11

Current Research Challenges in Crowdsourcing
Phuoc Tran-Gia 12

Enterprise Crowdsourcing in IT Service Delivery
Maja Vukovic 12

Crowdsourcing and Crowdsensing Application Scenario
Florian Zeiger 12

Overview of Additional Abstracts

Recommender Systems in Crowdsourcing Platforms
Kathrin Borchert 13

Technical and Human Aspects in Crowdsourcing Research
Matthias Hirth 13

Project Management Practices in Crowdsourcing Management
Deniz Iren 14

Crowdsourcing of Multimedia QoE Subjective Experiments
Christian Keimel 14

Motivation and Quality Assessment in Online Paid Crowdsourcing Micro-task Platforms
Babak Naderi 15

Working Groups

Working Group 1: Long-Term Perspectives & Impact on Economics in 5 Years
Christian Keimel 16

Working Group 2: Theory – Taxonomy and Dimensions of Crowdsourcing
Tobias Hoßfeld 16

Working Group 3: Industry Use Cases
Maja Vukovic 18

Working Group 4: Crowdsourcing Mechanisms and Design
Matthias Hirth 18

Open Problems and Challenges for Crowdsourcing 20

Panel Discussions 22

Presentation Session 1: Crowdsourcing in General
Matthias Hirth 23

Presentation Session 2: Industry Use Cases
Kathrin Borchert 24

Presentation Session 3: Crowdsourcing Design and Engineering
Christian Keimel 27

Presentation Session 4: Programming and Implementing Crowdsourcing
Deniz Iren 28

Presentation Session 5: Applications of Crowdsourcing
Babak Naderi 29

Late Night Session: Social Computing and Ethics of Crowdsourcing
Tobias Hoßfeld 30

Participants 33

3 Overview of Talks

In a madness session, all participants introduced themselves, stated their areas of interest, their expectations from the seminar, and their view on crowdsourcing science within at most five minutes. Before the seminar, the participants were asked to address the following questions. *What is your specific interest in crowdsourcing? What do you expect from this seminar? What are your objectives? Which challenges and specific questions would you like to address or see addressed? What do you have to offer? What do you expect from industry/academia?* Those questions were then addressed in more detail in the presentation session. The participants were asked whether they want to give a presentation and volunteers were given a 30 minutes slot for presentation including time for intensive discussions on their research. Based on the interests and the abstracts, the program of the seminar and corresponding presentation sessions were formed. A tag cloud of the most common words in the abstracts is visualized in Figure 1, while the abstracts itself can be found below.

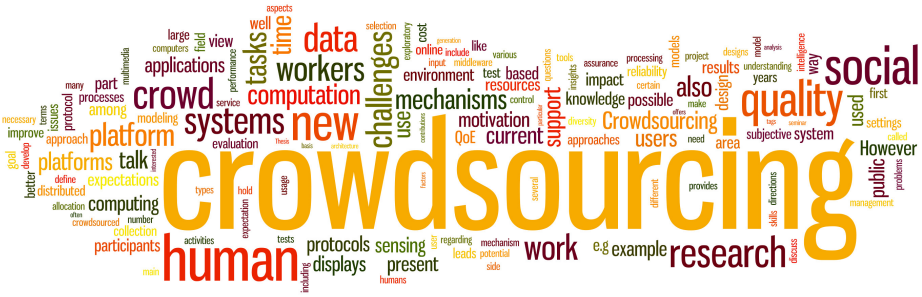

Figure 1 Tag cloud of the most common words in the abstracts.

3.1 Programming the Global Brain with CrowdLang and CrowdOS – Challenges and Ideas

Abraham Bernstein (Universität Zürich, CH)

Before the Internet most collaborators had to be sufficiently close by to work together towards a certain goal. Now, the cost of collaborating with anybody anywhere on the world has been reduced to almost zero. As a result large-scale collaboration between humans and computers has become technically feasible. In these collaborative setups humans can carry the part of the weight of processing. Hence, people and computers become a kind of "global brain" of distributed interleaved human-machine computation (often called collective intelligence, social computing, or various other terms). Human computers as part of computational processes, however, come with their own strengths and issues.

In this talks we take the underlying ideas of Bernstein et al. (2012) regarding three traits on human computation—motivational diversity, cognitive diversity, and error diversity—and discuss the challenges and possible solution approaches in order to embrace these traits as strengths instead of weaknesses.

3.2 Crowdsourcing Engineering

Alessandro Bozzon (TU Delft, NL)

While we are just starting to understand the societal and economical impact of crowdsourcing, available models and technologies for crowd organization and control are still in their infancy. The talk is intended to stimulate discussions about the need for 1) better conceptual abstractions for crowd tasks and processes design and (automatic) generation; 2) better understanding of crowds properties such as (soft and hard) skills, reliability, availability, capacity, and precision; 3) better tools for measuring and driving worker engagement.

3.3 Human Resource Management: From Workflows to Crowdsourcing

Cristina Cabanillas (Wirtschaftsuniversität Wien, AT)

The selection and allocation of human resources to activities has been increasingly researched in the field of workflow management in the last years. In particular, we have developed a language to define selection conditions on resources called RAL (Resource Assignment Language), we have presented mechanisms to automatically resolve RAL expressions and to analyze the business process resource perspective, and we are currently working on the prioritization of resources for allocation, specifically on the definition and resolution of preferences to generate a resource priority ranking.

We have evaluated a number of crowdsourcing platforms and we have found many differences among them regarding the selection and allocation of human resources to tasks. This makes us wonder whether our contributions on the field of BPM could be suitable in crowdsourcing platforms. In particular, we are interested in delving into the gap that should be bridged in order to improve, extend, or at least ease, the way in which human resources are managed in crowdsourcing systems.

References

1 C. Cabanillas, J. M. García, M. Resinas, D. Ruiz, J. Mendling, and A. Ruiz-Cortés. *Priority-Based Human Resource Allocation in Business Processes*. ICSOC 2013, pp. In press, 2013
2 C. Cabanillas, A. del Río-Ortega, M. Resinas, and A. Ruiz-Cortés. *CRISTAL: Collection of Resource-centrIc Supporting Tools And Languages*. BPM 2012 Demos, CEUR-WS, pp. 51-56, 2012
3 C. Cabanillas, M. Resinas, and A. Ruiz-Cortés. *Automated Resource Assignment in BPMN Models using RACI Matrices*. OTM 2012 (CoopIS'12), pp. 56-73, 2012
4 C. Cabanillas, M. Resinas, and A. Ruiz-Cortés. *Designing Business Processes with History-Aware Resource Assignments*. BPM 2012 Workshops (BPD'12), pp. 101-112, 2012
5 C. Cabanillas, M. Resinas, and A. Ruiz-Cortés. *Defining and Analysing Resource Assignments in Business Processes with RAL*. ICSOC 2011, pp. 477-486, 2011
6 C. Cabanillas, M. Resinas, and A. Ruiz-Cortés. *RAL: A High-Level User-Oriented Resource Assignment Language for Business Processes*. BPM 2011 Workshops (BPD'11), pp. 50–61, 2011

3.4 Human Computation and Crowdsourcing

Joseph Davis (The University of Sydney, AU)

URL http://kmrg.it.usyd.edu.au/node/587

Harnessing human computation through crowdsourcing and its integration with machine computation as a viable problem solving strategy for solving complex problems has emerged as an important research focus in recent years. I present the conceptual foundations of this approach and review some of the major contributions in this area.

3.5 Next-Generation Micro-task Crowdsourcing Platforms

Gianluca Demartini (University of Fribourg, CH)

At the eXascale Infolab in Fribourg, Switzerland we have been building hybrid human-machine information systems over the last 2+ years. Such systems involve both scalable/efficient data processing done by machines as well as effective processing of selected data done by humans. To build such systems it is important to have great control over the crowd performing micro-tasks. In my talk I will present current research directions in our group on how to improve current micro-task crowdsourcing platforms including quality assurance mechanisms by modeling workers in the crowd.

3.6 Data Analysis Tools for Crowdsourced Video Quality Tests

Klaus Diepold (TU München, DE)

Joint work of Klaus Diepold, Christian Keimel

We are active in moving Subjective Video Quality Tests from the laboratory space into the crowd. The ratings of a set of video sequences will be used to train a mathematical model in order to replace tests employing real eye balls for predicting the subjective quality of video sequences. To this end we propose to use tools from the multivariate data analysis toolbox, notably L-PLS Regression, to incorporate contextual information to further validate the mathematical model. This is a conceptual presentation void of results, featuring a list of expected outcomes which are up for discussion.

3.7 Supporting Exploration and Rapid Development by Citizens to Collect the Right Information through Crowd Sensing

Shinichi Konomi (University of Tokyo, JP)

To collect the information through crowd sensing, it is often required to define the goal and the method of data collection clearly in advance. However, it is not always possible to define them clearly in advance in many real-world situations. Therefore, I argue that there is the need to support exploratory activities in crowd sensing, in which participants can modify the relevant goal incrementally as they go about data collection activities. The Scene Memo system is an environment that supports exploratory crowd sensing based on a mechanism to share photos and tags in real time on smartphones and tablets. A small field study with the Scene Memo system suggested that shared tags can provide social cues and scaffold participants, thereby demonstrating the feasibility of a support mechanism for exploratory crowd sensing. We are also considering a rapid development environment of crowd sensing tools so that tools can be tailored easily to suit the changing goals of data collection.

3.8 Crowdsourcing Beyond the Desktop: Experiences and Challenges

Vassilis Kostakos (University of Oulu, FI)

Joint work of Vassilis Kostakos, Jorge Goncalves
Main reference J. Goncalves, D. Ferreira, S. Hosio, Y. Liu, J. Rogstadius, H. Kukka, V. Kostakos, "Crowdsourcing on the Spot: Altruistic Use of Public Displays, Feasibility, Performance, and Behaviours," in Proc. of the 2013 ACM Int'l Joint Conf. on Pervasive and Ubiquitous Computing (UbiComp'13), pp. 753–762, ACM, 2013.
URL http://dx.doi.org/10.1145/2493432.2493481

We discuss our work in moving crowdsourcing beyond the desktop. We present is the first attempt to investigate altruistic use of interactive public displays in natural usage settings as a crowdsourcing mechanism. We test a non-paid crowdsourcing service on public displays with eight different motivation settings and analyse users' behavioural patterns and crowdsourcing performance (e.g., accuracy, time spent, tasks completed). The results show that altruistic use, such as for crowdsourcing, is feasible on public displays, and through the controlled use of motivational design and validation check mechanisms, performance can be improved. The results shed insights on three research challenges in the field: i) how does crowdsourcing performance on public displays compare to that of online crowdsourcing, ii) how to improve the quality of feedback collected from public displays which tends to be noisy, and iii) identify users' behavioural patterns towards crowdsourcing on public displays in natural usage settings.

3.9 A Homo Ludens in the Loop

Markus Krause (Universität Hannover, DE)

Main reference M. Krause, "Designing Systems with Homo Ludens in the Loop," in P. Michelucci (Ed.), Handbook of Human Computation, Springer, 2014; to appear.
URL http://www.springer.com/computer/ai/book/978-1-4614-8805-7

For applications that are driven by input from a multitude of human raters, ensuring data reliability and organizing an interactive workflow constitute a complex challenge, especially when contributors are players. This talk introduces a novel approach to ensure data reliability of crowd-based and human computation systems. The proposed algorithm features the potential for direct feedback and interactivity in human computation games.

3.10 Towards Responsible Crowdsourcing

Martha A. Larson (TU Delft, NL)

Humans are the ultimate intelligent systems. Units of human work can be used to address the problems studied in the fields of pattern recognition and artificial intelligence. After years of research to crack certain tough problems, mere utterance of the phrase "human cycle" makes it seem like someone turned on a light in the room. Suddenly, we feel we are no longer feeling our way forward in darkness as we develop solutions. Instead, a bright world of new possibilities has been opened.

The excitement that crowdsourcing has generated in computer science is related to the fact that large crowdsourcing platforms make it possible to apply abstraction to human input to the system. It is not necessary to consider who exactly provides the input, or how they 'compute' it, rather the human processor can be treated as a black box. The magic comes when it is possible to make a 'call the crowd' and be sure that there will be a crowdworker there to return a value in response to that call.

However, crowdsourcing raises a whole new array of issues. At the same time that we excitedly pursue the potential of 'Artifical artificial intelligence' (as it's called by MTurk), it is necessary to also remember 'Human human computation'.

I am not an ethicist, and my first foray into crowdsourcing ethics is necessarily superficial. In fact, I started by typing the word "ethics" into my favorite mainstream search engine and picking a definition to study that seemed to me to be authoritative. However, I am convinced that the community of crowdworkers and taskaskers together form an ecosystem and that the main threat to this ecosystem is that we treat it irresponsibly.

In other words, we should not throw out everything that we have learned over centuries of human civilization about creating healthy and happy societies, stable economies and safe and fulfilled individuals in our quest to create new systems. Ultimately, these systems must serve humanity as a whole, and not disproportionately or detrimentally lean on the portion of the population that serves as crowdworkers.

The hopeful part of this undertaking is that it revealed many solutions to address ethical aspects of crowdsourcing. Some of them pose challenges that are just as exciting as the ones that motivated us to turn to crowdsourcing in the first place.

3.11 Crowdsourcing Challenges from Platform Provider's Point of View

Nhatvi Nguyen (Weblabcenter, Inc. – Texas, US)

Microworkers.com is an international Crowdsourcing platform founded in 2009, focusing on Microtasks. Until today over 450,000 users have registered at our platform from over 190 countries. This huge and diverse workforce is the key to the current success of Microworkers, but also imposed challenges both on the underlying infrastructure, as well as on support, e.g. handling service requests from or the resolving of disputes among users.

With my talk at the Dagstuhl seminar, I want to give some insights into the daily work of a Crowdsourcing platform provider. This can help to foster the discussion about new research directions. Here I would like to not only raise technical questions, but also social and legal ones, like "Can Crowdsourcing be used to support economies in developing countries?" or "Who is responsible for the healthcare of the workers?"'.

3.12 Social Protocols as a Basis for Social Computing

Munindar P. Singh (North Carolina State University, US)

Joint work of Munindar P. Singh, Amit K. Chopra

We are interested in investigating the foundations of social computing. We claim that today's social computing approaches provide a limited basis for modeling social relationships, which are the basis for the motivation and potential power of social computing. We propose an approach for understanding social computing as a means to enable the interactions among socially autonomous parties that involves a social protocol: the parties thus enact such a protocol.

A social protocol, in turn, is based on the expectations that it establishes between the participants. Examples of such expectations include domain-independent concepts such as commitments and authorizations as well as domain-specific concepts such as friendship. Associated with each expectation is a model of the trust that each participant may hold in the others, for example, whether they are competent to hold up the expectation and have the intention to hold up the expectation. We show how current social computing platforms and applications, such as Facebook and Reddit, may be understood in terms of the protocols that they embody.

The above applications incorporate the protocols into their design. However, our proposed way to thinking of protocols leads us to a simple architecture that would facilitate the development and maintenance of new applications. Specifically, the architecture would include social middleware with coverage for each family of social expectations used in a protocol. The middleware would support the interactions of the participants by carrying out the requisite bookkeeping.

This view leads to major research challenges including (1) modeling social protocols to verify their enactability and viability; (2) engineering middleware in a modular and composable manner to support new social expectations and new combinations of expectations in protocols; and (3) developing approaches for achieving agreement among stakeholders to lead to the specification of protocols that they can adhere to.

3.13 Current Research Challenges in Crowdsourcing

Phuoc Tran-Gia (Universität Würzburg, DE)

Joint work of Phuoc Tran-Gia, Tobias Hoßfeld, Matthias Hirth

The research of crowdsourcing in several aspects is part of the work of the Chair of Communication Networks – University of Würzburg. Within this area the current and previous research topics include, the design of mechanisms for crowdsourcing, analysis and modeling of platforms and their users, and the evaluation of new crowdsourcing use cases, e.g. in subjective QoE assessments. The questions of interest beyond are for example how to implement a quality control for tasks or how a platform will grow.

New research directions are for example the investigation of new types of crowdsourcing, e.g. real-time crowdsourcing or mobile crowdsourcing. In the talk during this seminar, I give an overview of the evolution of work organization, the different layers of the crowdsourcing processes and of current and previous work in this area conducted at the chair. I also present some new research challenges which could be addressed in the future.

3.14 Enterprise Crowdsourcing in IT Service Delivery

Maja Vukovic (IBM TJ Watson Research Center, US)

To increase the value of IT, enterprises rely on insights obtained by extracting large volumes of tacit knowledge about processes, products and people. This knowledge is not systematically discoverable, as it is unstructured and widely distributed among the experts in an enterprise. Typically this 'non-discoverable knowledge' is gathered in semi-automated way, which at best provides crude estimates, and doesn't scale. In this talk I will present an enterprise crowdsourcing (self-)service based on principles of 'wisdom of crowd' to enable rapid generation of a snapshot of the state of IT systems and operation, enabling for collaborative and distributed approach to knowledge gathering. I will discuss a number of use cases that engaged over 50K employees through crowdsourcing to uncover knowledge critical to IT services delivery. I hope to learn more about the current research efforts in academia with respect to incentive mechanisms and quality assurancel and how open APIs will change the nature of crowdsourcing.

3.15 Crowdsourcing and Crowdsensing Application Scenario

Florian Zeiger (AGT International – Darmstadt, DE)

This talk gives some ideas on how crowdsourcing / crowdsensing can be used in the area of safety and security. In more details the application example of environmental sensing is described from industry point of view and several challenges and open questions are discussed. In terms of scale of crowdsourcing and crowdsensing applications initially need to motivate

participants and this motivation has to be kept up over time. Other challenges are present in the area of user privacy, reimbursement methods, data quality and integrity, efficient and reliable data collection, as well as architectural decisions and flexible support of various business models.

4 Overview of Additional Abstracts

4.1 Recommender Systems in Crowdsourcing Platforms

Kathrin Borchert (Universität Würzburg, DE)

I first got in contact with the topic crowdsourcing from theoretical point of view during my Master Thesis. Further, I approached this topic from a practical side during my internship at Microworkers.com which is a crowdsouring platform focused on microtasks. A typical crowdsourcing platform offers an unmanageable amount of jobs and tasks to the workers. They have to decide which kind of tasks fit their interests and skills based on a short description of the jobs. Thus, workers might accept tasks without being qualified. This could impact the quality of the results and the employers might reject their work. So the workers and employers are unsatisfied. The integration of a recommender system potentially improve the result quality and the satisfaction of the customers. There are two possible locations to integrate such a system in a crowdsourcing platform: on employer side to help them finding workers concerning their needs and on the worker side to recommend tasks with respect to their interests and skills. The main part of my Master Thesis will be the analysis of several recommender algorithms and approaches to find such tasks or workers.

4.2 Technical and Human Aspects in Crowdsourcing Research

Matthias Hirth (Universität Würzburg, DE)

Crowdsourcing offers a lot new possibilities in computer science. Suddenly not only algorithmic results are available in the computation process, but also subjective human judgments. This enables new types of services and fosters user centric evaluations like, QoE tests or network measurements on end-user devices.

However, in crowdsourcing systems human factors have a significant influence on the results. This leads to new challenges which require a holistic view on these systems including both their technical as well as their human aspects. Therefore the Crowdsourcing research at the Chair of Communication Networks aims to contribute to this holistic view by providing statistical analyses of crowdsourcing data, models of crowdsourcing platform, and, in collaboration researcher from psychology, a better understanding of the human factors, e.g. the trustworthiness of users.

4.3 Project Management Practices in Crowdsourcing Management

Deniz Iren (Middle East Technical University – Ankara, TR)

As a means to access a scalable and rather cheap workforce, Crowdsourcing has successfully been in use for quite some time, especially by risk-prone entrepreneurs. However utilization of Crowdsourcing in large organizational projects provides difficulties of management. There is no defined way of estimating the cost, time and quality of work when it is performed by a crowd. As a project manager and a researcher my goal is to develop mechanisms to make crowdsourcing measurable, estimateable and thus; more manageable.

I propose using the project management practices to solve this issue. The cost/time/quality of a crowdsourced job depends on a number of design decisions such. For starters I suggest focusing on the impact of the decision regarding which quality assurance mechanisms to use. We can model various types of quality assurance mechanisms to estimate the impact of their usage on the project cost, schedule and quality. Using these models with certain statistical methods we believe we can estimate this impact quite accurate.

4.4 Crowdsourcing of Multimedia QoE Subjective Experiments

Christian Keimel (TU München, DE)

The evaluation of multimedia Quality of Experience (QoE) in traditional laboratory-based experiments is time consuming and expensive. Using crowdsourcing to conduct these subjective tests online not only reduces time and cost, but also allows for the participation of a large and diverse panel of international, geographically distributed users in realistic user settings. But moving from the controlled laboratory environment to the uncontrolled crowdsourced online environment leads to new challenges. Some issues, as for example the assessment of the workers reliability, are similar to other crowdsourcing applications. Other issues, however, are more specific to the evaluation of multimedia QoE. Long established best practices for laboratory-based tests aiming at preventing biases or other unwanted effects in the quantification of the subjects' judgments, for example, can possibly not be adapted to the crowdsourcing test environment. Moreover, it is also unclear if all commonly used subjective test designs can be used or only a limited subset of the exiting designs. This raises the question, if it is perhaps necessary to develop completely new test designs for the crowdsourced multimedia QoE evaluation.

4.5 Motivation and Quality Assessment in Online Paid Crowdsourcing Micro-task Platforms

Babak Naderi (TU Berlin, DE)

During my PhD thesis, in Quality and Usability Lab – TU-Berlin, I am working on motivation and quality assessment in online paid crowdsourcing micro-task platforms. Although it looks that the main drive for workers is the money they earn, I am looking on different type of motivation and their effects on the quality of work. In addition I am trying to create a model for predicting worker's motivation, generated data quality based on workers' behaviors in platform.

5 Working Groups

The intention of the seminar was on the identification of a (inter-disciplinary) research agenda, common research methodology, joint (industry and academia) activities and collaboration for crowdsourcing. To this end, the participants were asked before the seminar to complete a short user survey indicating their main interests and potential discussion items. The results of this user survey can be found in Table 1. Some additional discussion topics were raised ('other') that are massive open online courses, social good applications, human computation, resource selection and quality assurance.

Table 1 Number and ratio of participants' interests in different research topics.

Discussion Item	Interest	Ratio
Practical issues and experiences	17	16%
Industry use cases for crowdsourcing	16	15%
Improvement mechanisms	13	12%
Interdisciplinary discussions, e.g. law, psychology	13	12%
Long-term perspectives	13	12%
Standardization and open APIs	9	8%
Non-for-profit and academic use cases for crowdsourcing	9	8%
Theoretical frameworks for evaluation	6	6%
Path towards deployment	5	5%
Other	6	6%

Based on the survey and the observations of interests, discussions, interactions among participants etc. during the seminar, four different working groups were formed during the seminar in order to discuss certain aspects in more detail, that are

(W1) long-term perspectives & impact on economics in five years,
(W2) theory – taxonomy and dimensions of crowdsourcing,
(W3) industry use cases,
(W4) crowdsourcing mechanisms and design.

5.1 Working Group 1: Long-Term Perspectives & Impact on Economics in 5 Years

Christian Keimel (TU München, DE)

Joint work of Gianluca Demartini; Klaus Diepold; Christian Keimel; Martha Larson; Nhatvi Nguyen; Phuoc Tran-Gia

The discussions in the focus group on long-term perspectives, ethical issues, social aspects & impact in 5 years on economics lead to the following conclusions: - Crowdsourcing is sustainable in the future but emerging public perception or legal regulation may come into play.

- Workers, requesters, lawyers and platform providers should work together to define a vision of crowdsourcing ahead of regulations.
- More crowd-building, resulting in specialized crowds.
- Crowdsourcing is opening new and thrilling aspects for empirical research, but commercial aspects need to be added.
- Microworking should strive towards a marketable skill development.
- Building economic capacity in geographical regions and social segments where needs exist.

5.2 Working Group 2: Theory – Taxonomy and Dimensions of Crowdsourcing

Tobias Hoßfeld (University of Würzburg, DE)

Joint work of Abraham Bernstein; Joseph G. Davis; Tobias Hoßfeld; Deniz Iren; Munindar P. Singh

Currently, we are in the early stamp collection phase of crowdsourcing which means that the landscape is not clearly known. Therefore, this working group aimed in analyzing the dimensions of crowdsourcing as a basic step towards a common taxonomy. The goal of the taxonomy is not to achieve automation (based on this taxonomy), but rather to have good guidelines for the design and development of crowdsourcing tasks and campaigns. According to the group discussions, the five major dimensions of crowdsourcing are for each task (1) who, (2) where (3) why, (4) what, (5) how. Their meaning is explained in the following with several examples to illustrate the dimensions properly.

Who? A more detailed question is what is the type of the crowd. Here we have to differentiate the actual performers conducting the task. A crowd required for a certain task can be described in terms of (i) anonymity, (ii) number of (reliable) performers, (iii) social structure of the crowd and prior relationships between performers, (iv) diversity of performers (with respect to cognitive, cultural, gender, or personal bias), (v) trust, (vi) and the communication structure within the crowd. But we also have to take into account additional individuals or parties involved in the crowdsourcing process, including (a) employers, task assigners, or promoters, (b) beneficiaries, or (c) platform providers. This means the entire stakeholder structure has to be taken into account. This brings us directly to the next question.

Where? Another dimension is the location or context of the task, the crowdsourcing platform and the performers. The context where the task is executed may be important, e.g. for crowdsensing. However, it has to be differentiated between a physical location and

context. For example, the context of this seminar is the spirit of Dagstuhl, while it does not matter if this spirit appears in Kaiserslautern or in Saarbrücken. The crowdsourcing platform used has to be considered too, since it brings in not only its own crowd, but also certain mechanisms, which may affect the crowdsourcing task design.

Why? The incentives provided by the task and its execution are relevant towards a taxonomy. Thus, motivational diversity must be taken into account. Examples are (i) money as provided by commercial platforms like Amazon Mechanical Turk or Microworkers.com, (ii) fame and reputation e.g. for open source implementation, (iii) learning, i.e. the performer wants to benefit from executing the task by gaining new knowledge e.g. for open source implementation, (iv) altruism like social responsibility as observed for wikipedia, (v) side-effects like ReCaptcha where the actual task does not bring any incentives, but the task has to be conducted e.g. in order to get access to download software.

What? The type of goal or problem which is to be solved by the task is captured by this dimension. This includes the following aspects, (i) size of the task itself which may vary from contests to micro-tasks and therefore may require different mechanisms and design approaches, (ii) type of activity covering human computation, sensing, information foraging, evaluation, creation e.g. software or logo design, (iii) stopping condition of a tasks, i.e. whether the employer knows when the campaign is really done[1], (iv) possibility of evaluating the quality and reliability of the submitted results[2], (v) level of engagement of the performer which may be high e.g. crowdfunding, medium e.g. do something, or low e.g. sensing which is done by the user's device in the background, (vi) degree of control.

How? This dimension spans a variety of aspects on how the task is actually designed and executed. Thereby we differentiate on a higher level between (A) the method, mechanisms, or solution architecture as well as (B) the types of coordination mechanisms. Examples for coordination mechanisms are (1) aggregation of users from different platforms, of user results, and of tasks to macro-tasks, (2) selection of users or (sub-)tasks, (3) scheduling of the (sub-)tasks, or (4) resource allocation in terms of cost, time, and human efforts. To be more precise, the coordination mechanisms can be separated according to (i) the functionality, (ii) time, e.g. synchronized vs. asynchronous task execution by performers, e.g. real-time vs. serial execution, (iii) governance, (iv) coordination science framework.

Finally, the methods, mechanisms, or solution architectures address (I) degree of specificity, (II) degree of control, (III) locus of control, and (IV) the actual approach of the task. Thereby, we can differentiate between transparent tasks and obtrusiveness, direct vs. side-effect, piggyback e.g. sensing, parasitic e.g. captcha, or deliberate approaches.

Those five dimensions **Who, Where, Why, What, How** provide a basic framework for a taxonomy of crowdsourcing which may lead to useful guidelines for the design and development of crowdsourcing tasks and campaigns.

---

[1] The stopping condition is sometimes a non-trivial task, as it may require real-time evaluation of the submitted crowdsourcing results or a continuous execution of a task is necessary.

[2] For tasks like subjective testing, it is difficult to evaluate quality or reliability of results, as no correct answer and gold data exists. Thus, other quality assurance and reliability mechanisms are needed.

5.3 Working Group 3: Industry Use Cases

Maja Vukovic (IBM TJ Watson Research Center, US)

Joint work of Kathrin Borchert, Shi'nichi Konomi, Vasilis Kostakos, Maja Vukovic, Florian Zeiger

This workgroup has tackled the challenges in identifying trends and understanding what makes crowd sourcing applications successful, with focus on, but not limited to industry domain. We started this discussion session by having each participant share their favorite example of the crowd sourcing application. These ranged from Netflix, GalaxyZoo, to Amazon Mechanical Turk applications. The team concurred that often, most successful examples are the ones with most significant publicity and exposure. At the same time, the team has observed that there is still lack of convergence towards a unified crowd sourcing platform. There are numerous custom platforms, each of which is a solution for different type of crowd sourcing application. Finally, we discussed the notion of quality of crowd sourcing tasks, and let of clear metrics. Perception of quality often depends on different stakeholders in crowd sourcing process.

We concluded the discussion by identifying following five key challenges:

1. Advanced research and development of best design practices for crowd sourcing
2. A framework for evaluating crowd sourcing design considerations: task complexity, quality, price, and reaction time.
3. How can we "pre-certify" crowds, so that they can be quickly deployed to work on critical tasks, while maintaining expected quality levels.
4. How can we bring back and integrated research results from related crowd sourcing efforts?
5. There is an opportunity to engage community to reduce the isolation of crowd sourcing research.

5.4 Working Group 4: Crowdsourcing Mechanisms and Design

Matthias Hirth (University of Würzburg, DE)

Joint work of Alessandro Bozzon; Cristina Cabanillas; Matthias Hirth; Andreas Hotho; Markus Krause; Babak Naderi

This report summarizes the discussion and the outcome of the group work on "Crowdsourcing Design and Mechanisms" of Alessandro Bozzon, Cristina Cabanillas, Matthias Hirth, Andreas Hotho, Markus Krause, and Babak Naderi during the Dagstuhl seminar on "Crowdsourcing from theory to practice". The discussion mainly focused on two questions:
1) *"How can crowdsourcing experiments be designed in such a way, that they their results can be reproduced and validated by other researchers?"*
2) *"How and which steps in crowdsourcing measurements can be abstracted or generalized to speed up the task design process and to later provide guide lines for task design?".*

Currently, experimental driven research papers in the field of crowdsourcing include detailed descriptions of the test setup, the proposed/used methodology, and the obtained results. In contrast, details about the actual participants of the test, their demographics, their hardware equipment, or their cultural background are often neglected. However even

the results of simple tasks, like the transcriptions of texts on images can be influenced by these characteristics, because workers understanding the content of the texts can correct characters or passages which are unreadable based on the context. This is not possible for workers which are unfamiliar with the texts' language. Therefore, the results of subsequent test repetitions might significantly differ depending on the participants, even if the task setup remains unchanged. To address this challenge, the group discussed about *how crowdsourcing experiments can be designed in such a way that they their results can be reproduced and validated (by other researchers).*

During the discussion, the group agreed that some preliminary questions have to be addressed in order to tackle reproducibility of task results. First, similarity metrics of tasks have to be defined in order to identify task groups, among which the result should be comparable. Second, ways have to be found to measure similarity of crowds and to make crowds comparable. This can help to consider crowd specific biases and how to remove them in order to compare the results from different crowd provider. And finally, it has to be considered which parameters of an experiment have to be reported in a publication. Is it enough to simply mention the used crowdsourcing platform, or is it required to include detailed demographics of the workers? Crowdsourcing experiments are more difficult to repeat than measurements in technical environments, because of the human factors and possible varying behavior of the users. However, similar issues also exist in the field of psychology or human computer interaction research. Methods applied in these files might be suitable for conducting reproducible crowdsourcing experiments, even if some adaptions might be necessary. But also new methods have to be developed to consider crowdsourcing specific issues, e.g. technical changes in the crowdsourcing platform which affect the users' behavior. A simple solution to overcome some of these issues is using only a single crowdsourcing platform and performing multiple iterations. However, it is doubtful if these results would be really representative for crowdsourcing in general.

Another big challenge in current crowdsourcing research is the design and implementation of the crowdsourcing tasks themselves. Due to the lack of a standardized task descriptions and design guidelines, many researchers perform redundant work. This challenge was also discussed in the group and the participants tried to answer the questions *"How and which steps in crowdsourcing measurements can be abstracted or generalized to speed up the task design process and to later provide guide lines for task design?"*

One step to reach this goal is finding general properties of tasks, which can be used to describe tasks in a standardized way. A possible solution could be a semantic description of tasks. This can enable an abstraction of crowdsourcing task creation, e.g. via crowdsourcing programming languages, which has again benefits and drawbacks. The abstraction would enable an easy setup of crowdsourcing experiments, but would not reduce the number of influence factors on the results. Instead the abstraction layer hides the influence factors and makes them harder to analyze.

To enable the development of generalize task properties and design guidelines, it is helpful to separate experiments into different layers, namely

- the design of the experiment workflow,
- the design of actual tasks, and
- the evaluation of the results using technical and algorithmic tools.

This decomposition into layers fosters the development of design patters for tasks. Here similar ides from distributed computations approaches (e.g. Hadoop) can be included in the design of crowdsourcing workflows and task design. In this area some research studies also

already available. To describe tasks in general, different dimensions are required. The group identified the following preliminary set of task dimensions, which can be used as a starting point but is definitely not complete yet.

- Task requirements
  - Time to complete
  - Required skills of the workers
  - Required diversity of workers
- Task properties
  - Decomposability
  - Possibility to complete multiple instances in parallel
  - Complexity
- Task output
  - Expected confidence level
  - Data properties
  - Data complexity
- Task input
  - Diversity of the task data in terms of complexity (e.g. skewness of the data, might be analyzed using machine learning)
  - Data properties (size, ...)
  - Data complexity

Further the impact of pilot studies on the result quality was discussed. Pilot studies can either be specialized tasks to evaluate the qualification of workers for the actual research study or shortened version of tasks to evaluate the task design. This approach is assumed to reduce the costs for the experiments, because pilot tasks are usually paid less and reduce the number of required test iterations, as the task design is already optimized based on the pilot task results. Further the quality of the experiments results is also supposed to be better, because unqualified workers are filtered based on their performance in the pilot studies. Nevertheless, pilot studies could also influence the experiment results in unexpected ways or even falsify them. For example, demographical properties of the test participants might be skewed due to filtering of the workers. This in turn can lead to biased results. Further, in subjective experiments, pilot studies might change or influence expectations of participants which can lead also to based results. During the discussion, the participants agreed that the influence of pilot studies on the result quality highly depends on the specific task and that no general conclusions can be drawn.

6 Open Problems and Challenges for Crowdsourcing

During the seminar, several open problems and challenges for crowdsourcing were discussed. Although most of those aspects are covered in the summaries of the working groups and the presentation sessions, an (unsorted) list of concrete questions is given below to provide a comprehensive overview.

- What is an appropriate taxonomy for classifying and analyzing crowdsourcing systems? How can crowdsourcing tasks be grouped based on their task complexity and along key challenges in successfully harvesting expertise of large human networks?

- Which use cases and applications will exploit the potential of crowdsourcing?
- How does the research community approach improved crowdsourcing mechanisms e.g. for quality and cost control or reliability of users and devices? Which requirements and challenges occur for particular operational conditions, like ubiquitous crowdsourcing due to the user mobility in time and space?
- How to design incentive schemes for coordinated problem solving of the crowd among individual humans with their own goals and interests? How to realize gamification of work for improved user engagement? How to identify expertise of users? How to implement such incentive schemes technically?
- How can the experiment and task design be standardized? Which kinds of APIs or templates are promising and useful in practice?
- What are the objectives to be fulfilled and the necessary capabilities of platforms towards the provision of Future Internet services built on top of crowdsourcing facilities?
- How can crowdsourcing systems be evaluated? Which common research methodologies are applicable? Which theories and models from a number various fields are applicable, including artificial intelligence, multi-agent systems, game theory, operations research, or human-computer interaction? How to include human-centric measures such as costs, availability, dependability and usability, including device-specific properties in evaluation frameworks?
- How does the research agenda for crowdsourcing look like in the next years?
- How can crowdsourcing experiments be designed in such a way, that they their results can be reproduced and validated by other researchers?
- How and which steps in crowdsourcing measurements can be abstracted or generalized to speed up the task design process and to later provide guide lines for task design?
- How will open APIs change the nature of crowdsourcing?
- How to motivate participants especially in terms of scale of crowdsourcing and crowd-sensing applications and how to keep the motivation up over time?
- How can we "pre-certify" crowds, so that they can be quickly deployed to work on critical tasks, while maintaining expected quality levels.
- How can we bring back and integrated research results from related crowd sourcing efforts?
- Can Crowdsourcing be used to support economies in developing countries?
- Who is responsible for the healthcare of the workers?

The following challenges are observed for crowdsourcing by the seminar's participants.

- Need for better conceptual abstractions for crowd tasks and processes design and (automatic) generation;
- Better understanding of crowds properties such as (soft and hard) skills, reliability, availability, capacity, and precision.
- Better tools for measuring and driving worker engagement.
- Advanced research and development of best design practices for crowd sourcing.
- A framework for evaluating crowd sourcing design considerations: task complexity, quality, price, and reaction time.
- Engaging community to reduce the isolation of crowd sourcing research.
- Commercial aspects to be added to current crowdsourcing research.
- Microworking should strive towards a marketable skill development.
- Managing human resources in crowdsourcing systems.

- Integration of crowdsourcing with machine computation.
- Improvement of current micro-task crowdsourcing platforms including quality assurance mechanisms.
- Modeling workers in the crowd for optimization purposes and improved mechanisms.
- Specialized crowds and building of specialized crowds for particular campaigns in contrast to wisdom of the crowd.
- Dynamic adaptation of crowdsourcing systems during operation, e.g. to suit the changing goals of data collection in environment sensing systems.
- Evaluation of new crowdsourcing use cases.
- Organizing and managing the different layers of the crowdsourcing processes.
- User privacy of crowdsourcing users, e.g. in crowdsensing campaigns.
- Appropriate reimbursement methods and flexible support of various business models.
- Architectural design of crowdsourcing and machine clouds.
- Better understanding of incentives for participation in crowdsourcing for improved motivational design of campaigns.
- Organization of interactive workflows and direct feedback in crowdsourcing systems for improving reliablity and data quality.
- Responsible usage of crowdsourcing and ethical aspects.
- Building economic capacity in geographical regions and social segments where needs exist.
- Modeling social protocols to verify their enactability and viability.
- Developing middleware in a modular and composable manner to support new social expectations and new combinations of expectations in protocols.
- Approaches for achieving agreement among stakeholders to lead to the specification of protocols that they can adhere to.
- Emerging public perception or legal regulation of crowdsourcing and the impact on sustainability.
- Interworking of employers, employees, lawyers, and platform providers to define a vision of crowdsourcing ahead of regulations.

7 Panel Discussions

Based on the abstracts of the talks and interests of the participants, the seminar was structured accordingly. In five presentation and discussion sessions, the diverse aspects of crowdsourcing were elaborated. A joint late night session was organized with the parallel seminar "Cloud-based Software Crowdsourcing" to discuss crowdsourcing with respect to ethics and its relation to social computation. The topics of the sessions covered

(S1) crowdsourcing in general,
(S2) industry use cases,
(S3) crowdsourcing design and engineering,
(S4) programming and implementing crowdsourcing,
(S5) applications of crowdsourcing,
(S6) social computing and ethics.

A summary of the discussions during those sessions can be found in the following.

7.1 Presentation Session 1: Crowdsourcing in General

Matthias Hirth (Universität Würzburg, DE)

The session was opened by **Phuoc Tran-Gia** giving a talk on *"Current research challenges in crowdsourcing"*. In his talk, he detailed on the changing granularity of work enabling the transition from traditional outsourcing to crowdsourcing. Further, the similarities between human and machine clouds were discussed, as well as possible interconnections between them. Finally, influencing factors on the data quality and current research challenges were presented.

In the following discussion Munindar Singh asked, whether the granularity of tasks influences the quality of the resulting data. Phuoc Tran-Gia mentioned that there are some correlations, but a suitable selection of the workers can help to minimize the impact of this factor.

Thereafter the question was raised, what defines crowdsourcing. Joseph Davis mentioned that the one key aspect is that the workers are not chosen for a specific task, but can the tasks is distributed in an open call to an anonymous crowd. However, Martha Larson replied that the current anonymity of the employers and workers it no a mandatory property of the crowdsourcing system, but results from the lack of appropriate features in the current platforms. Tobias Hoßfeld had a similar point of view as Joseph Davis, as he stated that a relevant characteristic of crowdsourcing is that the workers select their work. Maja Vukovic introduces the payment system as criteria for crowdsourcing, because in crowdsourcing payment is usually based on a competition approach and the workers are only paid based on their results. Vassilis Kostakos asked whether it is important at all to have a common definition of crowdsourcing and Deniz Iren proposed that there can be different types of crowdsourcing. However, a taxonomy of these different types is required develop different management strategies. In reply to the analogue of human and machine clouds, Martha Larson stated that it is important that people know their applications are human powered. Klaus Diepold replied that in his opinion, not all user of cloud solutions are interested about the detailed realization of the product. Abraham Bernstein agreed with Klaus Diepold and suggested to foster ethical standards for crowdsourcing providers, similar to fair trade initiatives.

Munindar Singh stated that for some tasks, customers have to be explicitly notified that the used solution is crowdsourcing based, e.g. if the tasks includes the processing of confidential data. In the end a few statements about the future drivers of crowdsourcing were made. Phuoc Tran-Gia mentioned that crowdsourcing can be used as instrument to provide work in developing countries. Martha Larson drew an attention to the ethical challenges of this approach. Alessando Bozzon stated that crowdsourcing will be successful, because it enables fast and scalable problem solving.

The second talk about *"Human Computation and Crowdsourcing"* was given by **Joseph Davis**. Joseph Davis illustrated that the idea of human computation has a long tradition. Already in 1794, Gaspard de Prony hired unemployed hairdresser who knew only basic calculations for producing logarithmic tables. Joseph Davis introduced Crowdsourcing as a way to integrate human and machine computation, which enables the combination of the power of computers and human intelligence to solve complex problems that are beyond the scope of existing AI algorithms. This especially applies to problems involving conceptual thinking or perceptual skills. However, it is still challenging to combine and aggregate

the results obtained from different workers. Further, a good understanding of the applied aggregation mechanisms is required, because different mechanisms generate also different final results.

After the talk a discussion about Open Innovation arose. Tobias Hoßfeld asked, whether it is better to have larger number of homogeneous users or a smaller but more diverse group.

Joseph Davis replied that the optimal crowd composition depends on the tasks, but in most cases a bigger diversity is more important. Phuoc Tran- Gia addressed the issue of intellectual properties and rewards in crowd contests and Joseph Davis agreed that this is still an open problem. Maja Vukovic further mentioned the problem of unintended outcome of large scale crowdsourcing task. Even if user data is anonymized before publication, it might still be possible to track individuals and de- anonymization user profiles.

7.2 Presentation Session 2: Industry Use Cases

Kathrin Borchert (University of Würzburg, DE)

The first presentation was given by **Maja Vukovic** about "Enterprise Crowdsouring in IT Services Delivery". She described a use case of crowdsourcing in the background of enterprise IT transformation. In this context there are different tasks and challenges to solve and the required knowledge is often available in the enterprise crowd. An example is to find the hosting servers of specific applications or to define the production server out of an amount of registered. In their crowdsourcing model the first step is defining tasks and then finding the specific crowd with the knowledge and the skills to send them requests. The requested people can devide the tasks in subtasks and delegate them to other participants. After completing the whole task the workers get points as payment. The first question was raised by Phuoc Tran-Gia: "What was the incentive for the people to participate?". Maja Vukovic answered that the gathered data is disclosed to the participants, so they can also use it. She explained that it took 3 month to collect the information about the application owners, the last 1% took quiet long. Here Phuoc Tran-Gia asked for the quality assurance techniques for the results. The first 20% of the results have been cross validated by asking the application owners on which server the applications are hosted and later asked the server owners which applications are hosthed on their servers, was the answer of Maja Vukovic. Then Munindar P. Singh liked to know if there are any savings by using the crowd approach. Maja Vukovic replied that even with considering the overhead of the crowdsourcing efforts a huge amount of money was saved. The next question was requested by Vassilis Kostakos: "What is the impact of the approach on the business itself? People do no longer come together, just work remote?" This approach offers a possibility to find knowledge experts was the response. Furthermore, Vassilis Kostakos asked for side effects and if workers do more microtasks than their regular work. Maja Vukovic replied that nowadays, worker often distributed they work to others instead of doing it by themself. This might be a general development. Phuoc Tran-Gia liked to know if they also tested the quality of experience? But they did not analyzed that yet. The same answer was given to the question "How much workforce for the regular work did you loose by asking people to perform microtasks?" raised by Abraham Bernstein. The panel starting a discussion about the point system and what is their impact. Maja Vukovic told them that the points are only shown to participants. Then Abraham Bernstein liked to know what the usage of the points is. They are just a leaderboard was the explanation. Deniz

Iren asked the next question: "What happens to inaccurate results?" The answer was given by Maja Vukovic. She said that 80% of the collected data was correct. So the wrong part was not significant. After that, Vassilis Kostakos asked for legal and general issues. Maja Vukovic replied that there are some internal issues as privacy, spamming due to the high amount of email and employment status of participants. Nhatvi Nguyen focused back to the leaderboard by requesting a question concerning the implementation of its. He asked if they added penalties if a participant did not work accurate. But they did not. Martha Larson liked to know if the task was more about gathering data or updating the information. Maja Vukovic responded that the crowd helped gathering the data but then they were no longer interested in it. Further, she explained that the other crowdsourcing projects in IBM are crowdbased translation and crowd coder as Phuoc Tran-Gia asked for them. The last question concerning the presentation was raised by Tobias Hoßfeld: "Are there now innovate use cases research at the moment?" Maja Vukovic negated.

The topic of the second talk was "Crowdsourcing/Crowdsensing Application Scenarios" and was given by **Florian Zeiger**. He presented the usage of crowdsensing by explaining a project about environmental sensing concerning air pollution. Here, the participants get sensor devices which are configurable by a smartphone via bluetooth. The goal is to detect pollution events, hot spots and to localize the source of them. Tobias Hoßfeld asked the first question about the realization of the source localization. Florian Zeiger answered that if the model is not able to explain the pollution it asks for more data. Tobias Hoßfeld continued asking: "How can you trust the measurement results?" Florian Zeiger responded that they have good ways to verify data of fixed sensors, but it is different for crowd-measurements. It is difficult to manage that in an open system because the motivation of the users differ and it is possible for them to manipulate the results. He continued his talk with the presentation of open challenges and questions, e.g. scale of the system, privacy of the users, the support of different business models, the data quality and the reliability. The next question was raised by Joseph Davis: "40 million people downloaded an app for managing traffic status and maintaining map. Is there a community based approach, where a group maintains the data instead of the company?" Florian Zeiger answered that it depends on the ecosystem which shall be developed. There are possible scenarios in which this approach might be applicable, but there is no general answer. Joseph Davis also liked to know what the motivation of the users is and if there a community beyond the system? There is a system to register the people replied Florian Zeiger and the main motivation to participate is the interest in how the own environment is. But if it is good they lost interest in it. Then Babak Naderi asked for the experience on the usability of the system. The system is currently in a very early stage, limited to this area and there are no clear results, was the given answer. Then Klaus Diepold requested if there are a desired sensor resolution. Florian Zeiger responded that it depends on the geographical setup. In New York it would be great to have a sensor every few hundred meters which are measuring every 30 seconds. On german cities it is easier. That means it is not necessary to install sensors in the same high resolution. "Can the platform be bought?", was the next question of Andreas Hotho. Florian Zeiger explained that its only a prototype and the system is in the evaluation phase. So only user have access to the data and the measurement results are not sold. Someone asked if the platform is webbased. Florian Zeiger negated. The system is closed in the company. Abraham Bernstein brought up the question why these sensors are not deployed to public workers or soldiers to increase the sensor resolution? Florian Zeiger expounded that this is like the chicken-egg problem. Measurement probes are required to estimate the gain, but companies only buy if they know the possible gain. The last question concerning the presentation was raised by Gianluca

Demartini. He liked to know if the results are different than the results of the fixed sensors. Florian Zeiger affirmed.

Nhatvi Nguyen gave the last talk of this session. It was about crowdsourcing platforms from an operators point of view. The focused platform was Microworkers.com which mostly provids microtasks. Nhatvi Nguyen presented an overview of daily challenges which he as provider has to solve and explained the basic functionalities of the platform. One part is the competition between the workers. That means there are no guarantee to submit the completed task because the workers do not like task lock in the past. Martha Larson asked if Microworkers.com set up on the completion approach? That means workers will not get their money if they are not fast enough? Nhatvi Nguyen explained that there are two setups. The basic approach is race based and the other campaign type uses a time based approach. Cristina Cabanillas liked to know who decides if the workers get paid. Nhatvi Nguyen responded that in general the employer has the decision. But there is also a review system to prevent the employers from mis-rating tasks. That means Microworkers.com might interfere if the ratings are unfair. Babak Naderi revisited the topic of task completion by asking if the quality is better using the race approach or the time based approach. Race based campaigns are only used for simple tasks. To enable the time based setup the employers add a second layer. So the time based approach offers the request of more difficult tasks was the response of Nhatvi Nguyen. Babak Naderi continued asking: "Do the workers argue that they get no money?". The speakers affirmed: Microworkers.com will change the system to time based. He continued with the presentation and explained the payment system. The payment varies between the countries. So it is important to verify the location of the workers. He also told us that to improve the result quality Microworkers.com has a reviewing system for the campaigns. Here Joseph Davis liked to know if Microworkers.com is doing better than Mturk because of the manual reviews. Nhatvi Nguyen answered: "Yes, we put effort in this". He explained that Microworkers.com also runs cheat detection to improve the quality of the results. Then Martha Larson raised the question: "Does the black and white judgment, cheater/ not-cheater, of worker help or stand in your way?" Nhativ Nguyen gave the answer that the review feature helps to overcome this. Martha Larson also liked to know: "Using crowdsourcing for the first time is like playing a video game. Workers are offended by the term cheater, employers just want to get rid of cheater?". Microworkers.com gives multiple changes to the workers before classify them as cheaters and kick them out was the reply of Nhatvi Nguyen. The next asker was Florian Zeiger. He requested if there is an influence of the price on the quality in respect to the explained price model. Joseph Davis referenced to a paper and said that there is an optimal value of wages. Then Alessandro Bozzon changed the topic to the ethical aspects of crowdsourcing, e.g. the fair treatment of the workers including fair prices. He asked: "Should social aspects be addressed by Microworkers.com or by the employers?" Nhatvi Nguyen responded that Microworkers.com should foster it. Abraham Bernstein raised up a discussion with his question if the listed problems or issues are really crowdsourcing issues or real life issues? The discussion is focused on the topic which party is the employer and who is responsible for the workers. There is no clear answer to this. After that discussion Andreas Hotho liked to know how Microworkers.com reacts on jobs concerning grey level areas, for example downloading iPhone apps and requesting positive reviews. This kind of campaigns are removed on request told us Nhatvi Nguyen. Then Joseph Davis requested how Microworkers.com calculates the minimum wages, because it is difficult to set the price and to define what is a fair price. Nhatvi Nguyen explained that by manually reviewing the campaigns the price is set to a meaningful way. So if the price is too low, they ask the employer to raise the price.

7.3 Presentation Session 3: Crowdsourcing Design and Engineering

Christian Keimel (TU München, DE)

Markus Krause: *Designing Systems with Homo Ludens in the Loop*

The classification in the proposed gaming applications is only possible if the ground truth is available and the conclusion by the proponent was that games are useful human computation tools, but do not elicit intrinsic motivation, as the payment is still important. With respect to "magic recipes", the contributor remarked that "gamification sucks" and that the player should be put first i.e. the desires of player are paramount. In particular, task and game ideas need to be able to be linked and also the supporting infrastructure needs to be sufficient for the number of gamers. The complexity of the games did not change over time and the workers were provided with a button in the game, leading to a description of the scientific purpose of the game, but were not told directly the game's purpose in order to avoid the puppet master effect, as fun should be the main motivator for the game. Moreover the data was anonymized. The contributor recommends gamification, but suggests that one should be careful, and there should be nothing to lose for players, especially that there should be nothing dangerous for the players. Another question raised was if there are only certain human intelligence tasks that are suitable for a game or if there are limitations on possible human intelligence tasks. In particular, how much of a game the game should be for crowdsourcing tasks with gamification. A-priori tests in a laboratory can be used to identify overall issues with the gamified task design before a larger deployment in the crowd. In the provided example game for image annotation, between 240-1000 users participated.

Alessandro Bozzon: *Crowdsourcing Engineering*

The design of the underlying data is important for the task design and the data needs to be understood, possibly with crowdsourcing, before the design can be made accordingly. Also it was mentioned that current crowdsourcing platforms mostly focused on questionnaires. The experiments with the contributor's framework suggest that in crowdsourcing with social networks it is important how much the task requester's connections "care" about the requester and the more people "care" about the requester, the faster the task is finished. For traditional crowdsourcing it takes usually longer and corresponding data is already available in literature. Also it may be sensible to retain qualified workers that performed well in past tests and build task specific communities. Such task specific communities are different to groups established by mediator platforms, as these platforms do not know the details and purpose of a task as well as the requester. It was suggested that it may be necessary in future to build a real relationship between workers and requesters in order to address potential ethical and social issues.

Shi'ichi Konomi: *Citizen sensing, applications, transportations, rapid prototyping, collaboration, exploration*

One of the main conclusions by the contributor is that one size may not fit all and that crowd sensing can be improved by designing the right support tools e.g. explorations and rapid development tools for crowd sensing applications, allowing for a rapid adaptation to novel crowd sensing tasks. In the study of the contributor, about 30 users were used for trialling the wearable prototype in the initial crowd sensing deployment.

7.4 Presentation Session 4: Programming and Implementing Crowdsourcing

Deniz Iren (Middle East Technical University – Ankara, TR)

Abraham Bernstein: *Programming the Global Brain with CrowdLang and CrowdOS – Challenges and Ideas*

"Before the Internet most collaborators had to be sufficiently close by to work together towards a certain goal. Now, the cost of collaborating with anybody anywhere on the world has been reduced to almost zero. As a result large-scale collaboration between humans and computers has become technically feasible. In these collaborative setups humans can carry the part of the weight of processing. Hence, people and computers become a kind of 'global brain' of distributed interleaved human-machine computation (often called collective intelligence, social computing, or various other terms). Human computers as part of computational processes, however, come with their own strengths and issues.

In this talk we take the underlying ideas of Bernstein et al. (2012) regarding three traits on human computation—motivational diversity, cognitive diversity, and error diversity—and discuss the challenges and possible solution approaches in order to embrace these traits as strengths instead of weaknesses."

Following issues were raised in the discussion part:
Martha Larson: How do you define success in so called global brain?
– Both Linux and Wikipedia are examples of a successful global brain. The successful examples exists however we do not yet know how to develop them systematically.
Vassilis Kostakos: How would you deal with the fact that people contribute from different countries with different laws and culture?
– As researchers (of computer science) we are not the first people in the history to encounter the globalization problem. Humanity faced this problem before and overcome it. We just have to find the ways to deal with it according to the labor laws.
Alessandro Bozzon: Does global brain have to be designed or can it develop by itself organically like cities do?
– There are cities built with detailed planning. Depending on what you are trying to accomplish, hierarchical design may not be a bad thing.
Joseph Davis: Can you specify abstract model of the global brain?
– I would like to do that but I don't know how. Not yet. I would like to use the diversities between human and machines to produce a better solution. But in the end we may take an approach similar to software engineering.

Gianluca Demartini: *Next-Generation Micro-task Crowdsourcing Platforms*

"At the eXascale Infolab we have been building hybrid human-machine information systems over the last 2+ years. Such systems involve both scalable/efficient data processing done by machines as well as effective processing of selected data done by humans. To build such systems it is important to have great control over the crowd performing micro-tasks. In my talk I will present current research directions in our group on how to improve current micro-task crowdsourcing platforms including quality assurance mechanisms by modeling workers in the crowd."

Following issues were raised in the discussion part:
Phuoc Tran-Gia: How do you know who is online to make real-time contributions? How can you count on their real-time contributions?

– It depends on the budget and the task completion time. One increases when the other one decreases.
Cristina Cabanillas: Some quality assurance mechanisms are needed. Selecting the most trusted individual based on Facebook data can be misleading.
– I agree. It is still a work in progress.
Phuoc Tran-Gia: Crowd selection has been done by other crowdsourcing platforms before. What is new here?
– We aim at crowd selection in an algorithmic manner in order to optimize it.

Cristina Cabanillas: *Human Resource Management from Workflows to Crowdsourcing*

"The selection and allocation of human resources to activities has been increasingly researched in the field of workflow management in the last years. In particular, we have developed a language to define selection conditions on resources called RAL (Resource Assignment Language), we have presented mechanisms to automatically resolve RAL expressions and to analyze the business process resource perspective, and we are currently working on the prioritization of resources for allocation, specifically on the definition and resolution of preferences to generate a resource priority ranking.

We have evaluated a number of crowdsourcing platforms and we have found many differences among them regarding the selection and allocation of human resources to tasks. This makes us wonder whether our contributions on the field of BPM could be suitable in crowdsourcing platforms. In particular, we are interested in delving into the gap that should be bridged in order to improve, extend, or at least ease, the way in which resources are managed in crowdsourcing systems."

Following issues were raised in the discussion part:
Gianluca Demartini: Can standard resource modelling techniques be used for crowdsourcing?
– Modeling the resources within an organization is straightforward however doing it in crowdsourcing is not.

7.5 Presentation Session 5: Applications of Crowdsourcing

Babak Naderi (TU Berlin, DE)

Vassilis Kostakos: *Social Media, Public Displays, Crowdsourcing*

Most of discussions were around the study: Rogstadius et al 2011 "An Assessment of Intrinsic and Extrinsic Motivation on Task Performance in Crowdsourcing Markets". To evaluate the intrinsic motivation authors published two equal task with different description regarding to organization whom is going to use the result, one profit and the other non-profit company. They conclude extrinsic motivators increase workers' willingness to accept a task & the speed at which a task is completed. However intrinsic motivators improve output quality. The discussion was focused on morality & ethical issues in designing the experiment. Questions were if the participants have been informed afterwards that the images and tasks were fake or how do you expect that workers do not cheat when you are manipulating task description. Another interesting part was deploying tasks on touch interaction screens in the campus which leads to engage more workers than similar task running in MTurk meanwhile. However due to the time limit this part was not discussed.

Klaus Diepold: *Data Analysis Tools for Crowdsourced Video Quality Tests*

Klaus Diepold has presented application oriented perspective to crowdsourcing with the goal of optimizing video services for visual quality. Their effort for replacing subjective visual quality testing in laboratory with crowdsourcing approach and their mathematical model for predicting quality have been shown. In this context, using crowd sourcing test leads to reduce cost, speedup testing, possibility of larger tests (more reliable statistics), and improved diversity. In this session different points have been discussed: First of all considering outlier detection it has been discussed if consistency of responses should be taken in account rather than accuracy for detecting outlier. When a worker is jugging all the time same thing it shows his/her different point of view. Second discussion point was what type of features should be provided by crowdsourcing platforms for facilitating conducting such a studies? Different points have been suggested like delivery mechanisms should not affect the quality of content (high bandwidth for serving videos), and while sessions of quality judgments are long (i.e. about 20 minutes), therefore platform should take care about consequences (e.g. worker becomes tired). Another comment regarding to this perspective was exploiting the crowdsourcing idea for doing science (here improve encoder using mathematical analysis) is a potential use case for crowdsourcing. Other suggesting like using real-time feedback of crowd regarding to encoding mechanism, effect of intrinsic motivation in subjective quality tests, and current state of research in this area has been discussed.

7.6 Late Night Session: Social Computing and Ethics of Crowdsourcing

Tobias Hoßfeld (Universität Würzburg, DE)

For the late night session, the participants from the parallel seminar on "Cloud-based Software Crowdsouring" (Dagstuhl Seminar 13362) organized by Michael N. Huhns, Wei Li, Martin Schader and Wei-Tek Tsal were invited to join the discussions. On one hand ethical issues in the area of crowdsourcing were raised in a stimulus talk by Martha Larson (TU Delft). On the other hand, Munindar P. Singh (North Carolina State University) intended to provoke with his talk on the critique of current research in the area of social computing and crowdsourcing.

Munindar P. Singh: *Honest-to-Goodness Social Computing: Critique of Current Research and a Call to Arms* In social computing, the word "social" is introduced. Thereby, the social relationships are the basis for the motivation and potential power of social computing. The social state is a snapshot of a set of social relationships, while social computing means to compute the social state, that is, operate a social machine. Doing so involves conceptually modeling and specifying a social machine. The principals are autonomous parties i.e. people or organizations who are accountable. There may be a hidden principal behind the machine. On top of communication, there is (a) communication transport i.e. the protocol, (b) machine which means API and communication, and (c) interactions combining protocol and social middleware. The social computing architectural schematic is composed of several layers that are (i) social application: employing social relationships, (ii) social middleware: maintaining social relationships, (iii) distributed substrate: computing, communications, information. This architecture is grounded in social abstractions and a decentralized conception to support autonomous operation. Agents enact a specified social protocol on behalf of their principals, while the social middleware maintains the social state.

The major research challenges are the following. (1) Modeling protocols and ensuring their enactability and viability. (2) Developing middleware to compute and progress the social state. (3) Specifying a social protocol (collectively by principals). Generalizing social computing means to have models of problems vis a vis architectures of solutions. Concrete questions are the following. *Who may initiate a computation? Who selects the participants? Are the parties interested in the outcome? Do the parties interact repeatedly? Do the parties learn and might useful outcomes emerge? Is it a majority or a minority game? How do participants interact?*

The conclusions on achieving the promise of social computing are the following. Computer science is a game of abstractions and for social computing we need new abstractions. In particular, human-level abstractions are to be incorporated, problems are to be elicited more precisely, more flexibility to participants is to be granted, a clearer accountability of actions is to be obtained, and work is to be held to higher standards of norms and ethics. The discussion then was about the following questions. *What is the relation to software development? May the agent be abstracted by API? What is the social middleware in case of crowdsourcing? Is it Enterprise Crowdsourcing?*

Martha A. Larson: *Crowdsourcing Ethics: Towards operationalizing ethics*

Ethical issues in the area of crowdsourcing were then raised in the second talk of the late night discussion session. First the concept of ethics was elaborated. It involves systematizing, defending, recommending concepts of right and wrong behavior. The normative principles in applied ethics considers (i) personal benefit, i.e. acknowledge the extent to which an action produces beneficial consequences for the invidiual in question, (ii) social benefit, (iii) principle of honesty, lawfulness, autonomy (acknowledge a person's freedom over his/her actions or physical body), justice, (iv) rights. The question arises why we should care about ethics in the context of crowdsourcing. We are actors in crowdsourcing systems to directly benefit from an environment that respects ethics. In general, long run altruism beats greed. An ethical system is more sustainable (everyone is empowered to reach full capacity).

The next question is when we should care about ethics. The answer is simple – always. In cases where crowdwork income is not essential: *What are people not doing when they are crowdsourcing? Does crowdsourcing have an addictive side?* In cases where people are subsisting on crowdwork income: *Is the crowdwork paying only money, without building a longer term future? Is the money enough? Is the timing of the payment ok?*

There are easy ways for taskmakers to take ethics into consideration. (1) Try some crowdwork. (2) Do not reveal identity of the crowdworkers. (3) Address crowdworkers politely (and try to apologize if you goof up). (4) Don't engage in the "race of bottom". (5) Review work promptly. (6) Realize there is an investment of time and not only money. (7) Respond to inquiries (even the ones with horrible grammar). (8) Use contracts to clearly define your relationships. (9) Be open and honest.

Nevertheless, there are challenging ways towards ethics. There is more than one system involved in determining what is ethical and non-ethical behavior on the platform. Further it should be ensured to give people the information that they need to make informed decisions about the use of their work and their information. Understanding the past will also be an important issue. Privacy is crucial and health state and personality information need to be protected like it was priceless art. More challenging ways towards ethics address the following. (A) Think beyond instant gratification. (B) Spend lots of time following guidelines for human subject research. (C) Spend lots of time doing non-science activities like joint task forces to improve and innovate crowdsourcing practices. (D) Be close to workers. (E) Mirror the ethics mechanisms of the real world on the crowdsourcing platform. (F) Engage in ethics

by design. A crowdwork unit should be associated with certification information (e.g. fair trade). Crowdworkers can feedback information and self organize if they can communicate among themselves. Every party on a crowdsourcing platform should reach all other parties.

Some interesting notes taken from the discussion: "CS workers are not anonymous, they have skills and motivations." "Crowdsourcing reduces humans to mechanical devices." "In cases where crowdwork income is not essential: What are people not doing when they are crowdsourcing? Does crowdsourcing have an addictive side? In cases where people are subsisting on crowdwork income: Is the crowdwork paying only money, without building a longer term future? Is the money enough?" "There are easy ways for task makers to take ethics into consideration: do not reveal identity of the crowdworkers; address crowdworkers politely; review work promptly; realize there is an investment of time and not only money; respond the inquiries; use contracts to clearly define your relationships; mirror the ethics mechanisms of the real world on the crowdsourcing platform."

Participants of the Seminar

In total, 22 researchers (18 male, 4 female) from 9 different countries and 16 different institutions or companies participated in the seminar. The majority of participants is working in academia (86 %) in Europe (77 %). But it has to be highlighted that especially the participants from industry were enriching the seminar and brought in important and complementary viewpoints. Similarly, the young researchers (36 %) actively brought in their opinions and enriched the discussions especially in the working groups. The complete list of participants can be found below.

Participants

- Abraham Bernstein
Universität Zürich, CH
- Kathrin Borchert
Universität Würzburg, DE
- Alessandro Bozzon
TU Delft, NL
- Cristina Cabanillas
Wirtschaftsuniversität Wien, AT
- Joseph Davis
The University of Sydney, AU
- Gianluca Demartini
University of Fribourg, CH
- Klaus Diepold
TU München, DE
- Matthias Hirth
Universität Würzburg, DE
- Tobias Hoßfeld
Universität Würzburg, DE
- Andreas Hotho
Universität Würzburg, DE
- Deniz Iren
Middle East Technical University – Ankara, TR
- Christian Keimel
TU München, DE
- Shinichi Konomi
University of Tokyo, JP
- Vassilis Kostakos
University of Oulu, FI
- Markus Krause
Universität Hannover, DE
- Martha A. Larson
TU Delft, NL
- Babak Naderi
TU Berlin, DE
- Nhatvi Nguyen
Weblabcenter, Inc. – Texas, US
- Munindar P. Singh
North Carolina State Univ., US
- Phuoc Tran-Gia
Universität Würzburg, DE
- Maja Vukovic
IBM TJ Watson Res. Center, US
- Florian Zeiger
AGT International – Darmstadt, DE