

A Software Platform Design Approach for Knowledge Sharing and Self-Organization among Crowd Workers

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Business Informatics

eingereicht von

Julia Filler, BSc

Matrikelnummer 1225408

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Ao.Univ.Prof. Dipl.-Ing Dr. Stefan Biffl

Mitwirkung: Jürgen Musil, Angelika Musil

Wien, 30. August 2018

Julia Filler

Stefan Biffl

A Software Platform Design Approach for Knowledge Sharing and Self-Organization among Crowd Workers

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Business Informatics

by

Julia Filler, BSc

Registration Number 1225408

to the Faculty of Informatics

at the TU Wien

Advisor: Ao.Univ.Prof. Dipl.-Ing Dr. Stefan Biffl

Assistance: Jürgen Musil, Angelika Musil

Vienna, 30th August, 2018

Julia Filler

Stefan Biffl

Erklärung zur Verfassung der Arbeit

Julia Filler, BSc
Treustraße 92/3/9, 1200 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 30. August 2018

Julia Filler

Acknowledgements

At this point I would like to thank all those who supported and motivated me during the preparation of this master thesis.

My first thanks are due to Prof. Biffl, who supervised and examined my master thesis and for his helpful feedback during this time. A huge thank goes to Jürgen Musil and Angelika Musil who give me always helpful suggestions and constructive feedback. Thanks to them I was able to do research in this interesting topic. Generally, I would also like to thank for the great support and the time, which was invested.

Special thanks go to all participants of my survey, without whom this work could not have come about. My thanks go to their willingness to provide information and interesting contributions and answers to my questions. Especially, I want to thank Silvia Kuba from the chamber of labour in Austria and Karin Zimmermann from the ÖGB who supported this thesis.

Especially, I want to thank my dearest friend Romana Jakob who was always there for me and has provided me the most beautiful and joyful time at the TU Wien. Also, I want to thank Alexander Schörghuber, Julian Lehner and many other friends which were by my side during my studies, for their support over the duration of my entire studies. Furthermore, I want to thank Max Landauer who proofread this thesis.

Last but not least: My special thanks goes to my family, especially to my parents Josef and Gerlinde Filler, who give me the opportunity to study and supported me in all my decisions unconditionally.

Kurzfassung

Crowd Work ist eine beliebte Form der digitalen Arbeit. Jeder kann auf Crowdsourcing-Plattformen arbeiten und für Unternehmen ist es eine günstige Möglichkeit, Aufgaben extern auszulagern. Diese neue Form der Arbeit hat sich im Laufe der letzten Jahre stark entwickelt, so dass die Crowd Worker sowie die Politik mit gravierenden Missständen jetzt konfrontiert sind.

Daraus ergeben sich drei Probleme: 1. Es gibt keine Möglichkeit sich untereinander auf einer Crowdfunding Plattform auszutauschen. 2. Es gibt zu viele Foren und Chats auf denen Crowd Worker untereinander kommunizieren können. Das führt zu einer Verteilung des Wissen. 3. Mitglieder von Interessensgemeinschaften (AK Wien, ÖGB, IG Metall) haben keine Möglichkeit in direkten Kontakt mit den Crowd Workern zu kommen. Außerdem fehlt die Möglichkeit regelmäßig die Arbeitskonditionen der Crowd Worker zu sammeln. Daraus, ergibt sich unser Ziel: Eine zentrale Lösung zu entwickeln, die all diese Probleme löst. Über unser Forschungsfragen haben wir zuerst mal die Bedürfnisse von Crowd Worker, Plattform Betreibern, Arbeitgeber und anderen Interessengruppen (wie AK Wien, ÖGB, IG Metall) ermittelt.

Danach haben wir sieben Designprinzipien und Architekturmodelle ausgearbeitet. Diese sollen die wichtigsten Merkmale zur Erfüllung der Bedürfnisse von Crowd Workern und anderen Stakeholdern aufzeigen und identifizierten Geschäftsprozesse abbilden. Der vorgeschlagene Ansatz verfolgt ein Architekturdesign eines Collective Intelligence System. Diese Art von System ermöglicht eine indirekte Kommunikation unter den User. Diese Art von Kommunikation belebt das System und führt zu einem regen Austausch unter den Usern.

Zur Ermittlung der Richtigkeit und Nützlichkeit der Designprinzipien und Architekturmodelle, wurde ein Prototyp implementiert. Danach wurden Stakeholder eingeladen zu einer Case Study, um den Prototypen zu evaluieren. Die Ergebnisse zeigten eine sehr gute Übereinstimmung der Designprinzipien und der Collective Intelligence Modelle für die Crowd Work Domain. Daher wird geschlussfolgert, dass das Ziel der Arbeit erreicht wurde durch die ermittelten Designprinzipien und Modelle. Die Crowd Worker und Interessensgruppen werden optimal durch eine zentrale Plattform unterstützt werden.

Abstract

Crowd Work is a popular form of digital work. Anyone can work on crowdsourcing platforms and it is a great way for companies to outsource tasks externally. This new form of work has developed strongly over the last few years, so that the crowd worker and politics are now faced with grave grievances.

This results in three problems: 1. There is no way to communicate with each other on a crowdworking platform. 2. There are too many forums and chats where crowd workers can communicate with each other. This leads to a dissemination of knowledge. 3. Members of interest groups (Chamber of Labor in Vienna, ÖGB, IG Metall) have no opportunity to get in direct contact with the crowd workers. In addition, there is a lack of opportunity to regularly collect the working conditions of the crowd worker. Our goal is to develop a central solution that solves all these problems. In our research questions we first identified the needs of crowd workers, platform operators, employers and other stakeholders (like chamber of labor in Vienna, ÖGB, IG Metall).

Then we evaluated seven design principles and architectural models. These are intended to highlight key features to meet the needs of crowd workers and other stakeholders and to map identified business processes. The proposed approach follows an architectural design of a collective intelligence system (CIS). A CIS allows an indirect communication among the users. This type of communication animates the system and leads to a lively exchange among users.

To determine the correctness and usefulness of the design principles and architectural models, a prototype has been implemented. Afterwards, stakeholders were invited to a case study to evaluate the prototype. The results showed a very good match between the design principles and the collective intelligence models for the crowd work domain. Therefore, it is concluded that the goal of the work has been achieved through the determined design principles and models which optimally support crowd workers and stakeholders through a central platform.

Contents

Kurzfassung	ix
Abstract	xi
Contents	xiii
1 Introduction	1
1.1 Motivation	2
1.2 Problem Statement	7
1.3 Structure of the Work	9
2 Background & Related Work	11
2.1 Crowd Work	11
2.2 Socio-technical Systems for Crowd Workers	20
2.3 Collective Intelligence Systems	23
2.4 Summary of the papers and their limitations	28
3 Research Approach	31
3.1 Research Challenges	31
3.2 Research Questions	32
3.3 Research Methodology	34
4 Stakeholders, their Concerns and Needs	37
4.1 Crowd Workers	37
4.2 Platform Operators and Requesters	42
4.3 Regulators	45
4.4 Interest Groups	47
5 Design Principles	51
5.1 Independence and Centralization	52
5.2 Privacy	53
5.3 Trust and Reputation	55
5.4 Fairness and Objective Exchange	56
5.5 Distribution of Information and Awareness	57
	xiii

5.6	Analyzability and Transparency	58
5.7	Work Type-Specific Software Clients	60
5.8	Analysis of Platforms Used for Crowd Workers' Self-Organization	61
5.9	Summary	67
6	Software Platform Architecture Design for the Crowd Work Domain	69
6.1	CI Context Viewpoint	70
6.2	CI Technical Realization Viewpoint	80
6.3	CI Operation Viewpoint	93
7	WorkerHub Prototype	97
7.1	Register & Terms of Use	97
7.2	Users, Role Management & Reporting System	99
7.3	Preference System & Dashboard	102
7.4	Groups & Notifications	103
7.5	Alert System	107
7.6	Advocacy	108
7.7	Survey & Statistic	108
7.8	User Profile	110
7.9	Implementation Details	110
8	Evaluation & Discussion	115
8.1	Case Study Design	115
8.2	Results & Discussion	119
9	Conclusions & Future Work	133
	List of Figures	136
	List of Tables	139
	Bibliography	141
.1	Appendix A	149
.2	Appendix B	152

Introduction

For more than a decade, crowdsourcing has established itself as a popular form of digital work. Initially mentioned by Howe [1], the term is a *portmanteau* of the terms crowd and outsourcing as we can see in figure 1.1. In detail, the crowdsourcing mechanism works as follows [2]: A requester (a company) offers (many little) tasks (*microtasks*) via a digital platform (*crowdsourcing platform*) to anonymous workers (*the crowd*), who perform these tasks for a compensation. These tasks can either be location-independent (*cloudwork*) or location-dependent (*gigwork*) [3]. In recent years, the term crowd work has been used more often to refer to crowdsourcing as more work-related aspects have become particularly dominant in the societal discourse. On the positive side crowd work is beneficial for workers and requesters: Workers can easily search for jobs online and take on work when and how much they like. Requesters can offer jobs on the platform in an automated way and can offload risks and costs to the workers. The crowd workers can decide which job they want to do and companies can choose who should get paid for the task.

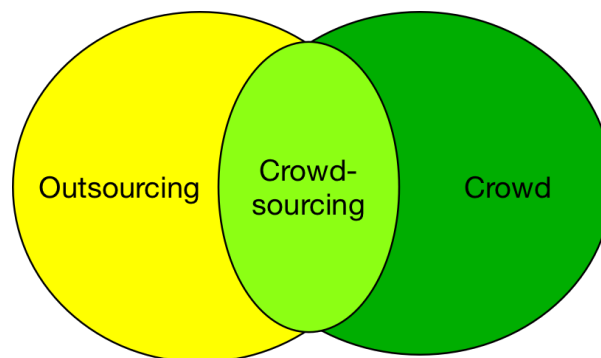


Figure 1.1: Origin of Crowd Work [1]

These tasks are typically data handling jobs, which need human cognition for implementation and the results can be evaluated mechanically [3]. Each microtask is a piece of a puzzle, which belongs to a larger task [3]. In Figure 1.2 typical microtasks are presented as mentioned by Schmidt [3]: verifying data sets, digitalization of credit cards, transcribing of audio files, creating product descriptions, sentiment analysis and content moderation.

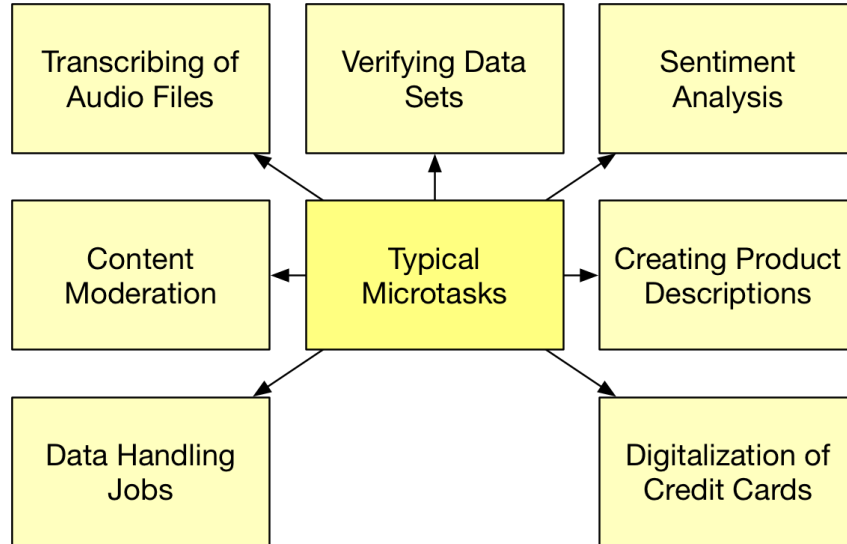


Figure 1.2: Examples of typical microtasks [3].

1.1 Motivation

The processing of these tasks is automated, anonymized and assembled line-like as far as possible [3]. Mary L. Gray from Microsoft Research talks about the „last mile of automation“ as the last piece, which cannot be automated [3]. The design of a microtask is essential. It should not involve the variability of a worker, otherwise it could have negative effects on the whole work [4] [5] [6]. This decentralization should support the accuracy of microtasks through independent judgment [7]. The advantage of the decomposition of the tasks into microtasks is robustness [8]. The drawback is the loss of context for the worker with every smaller decomposition [9]. For example, workers interpret tasks in different ways [10], edit adjacent paragraphs inconsistently [11] [12], or exhibit lower motivation [13] without context [9].

There are a lot of existing inequalities and grievances, which represent the motivation for this work to help crowd workers:

- **Few Cents and Uneven Payment**

The unit of times and remunerations of microtasks are simply estimated by seconds

and cents [3]. Workers are not rewarded for their creativity but for predictability and the payment is determined by an algorithm which does not understand humans [14] [15] [16]. Furthermore, workers of Amazon Mechanical Turk (AMT), a large crowdsourcing platform in the USA, are usually paid from the platform in form of Amazon vouchers except in the USA or India [3]. The wage level is different in every country. In Austria, a few cents for a task is nothing, but in India this is different [3].

Tasks can be refused which are not that easy to parse hence the workers do not get any compensation [17]. Furthermore, unsecured payment is something that happens again and again and it is not possible for workers to appeal because of low payment [3]. Most of the workers complain about uneven or unfair payments [16].

- **80 Hours a Week for Full-Time Compensation**

The working hours differ from worker to worker. For some workers it is only a secondary employment. They work in average 7,39 hours a week and a maximum of 25 hours [18]. A full-time crowd worker has to work 80 hours a week to get the same compensation as a normal employee with 40 hours a week [18]. And it is shown that workers who do crowdworking as a main activity earn in average more than the others [3]. Of course, these values depend on the kind of work. Simple "clickwork" has a lower compensation than design tasks [18].

- **Average Performance for Almost Certain Payment**

The prepared microtasks are compared by an algorithm and those who are deviant will be sorted out [3]. The computer is comparing all assignments with each other and is sorting out tasks, which are not fitting the standard - therefore the worker will not get paid [3]. Figure 1.3 illustrates this procedure. A requester splits a task into several microtasks. These microtasks are offered on a crowd work platform and the crowd workers can pick a microtask. After fulfillment of the microtask, the result is visible on the platform and the algorithm checks all results from all crowd workers against each other. The completed results are then sent to the requester who will only pay for microtasks, which are in their opinion done the right way. It could happen that despite a worker tries hard to finish a task and her result is much better than the average tasks, her task will not be submitted to the requester [3].

- **No Social, Pension and Health Insurance**

Crowd workers have no social, pension and health insurance. This fact makes it so cheap for the requesters to offer tasks. In the case of AMT, it is possible because they legally define the workers as contractor subject to law designed for freelancers and consultants [16]. So per terms of condition, they are self-employed [3].

- **Nearly 50% has a University Degree**

It might appear that crowd workers are under educated. Most recent studies show another picture. The study from Leimeister et al. [18] shows that more than three-quarters of those questioned had a high school graduation and 48% of those

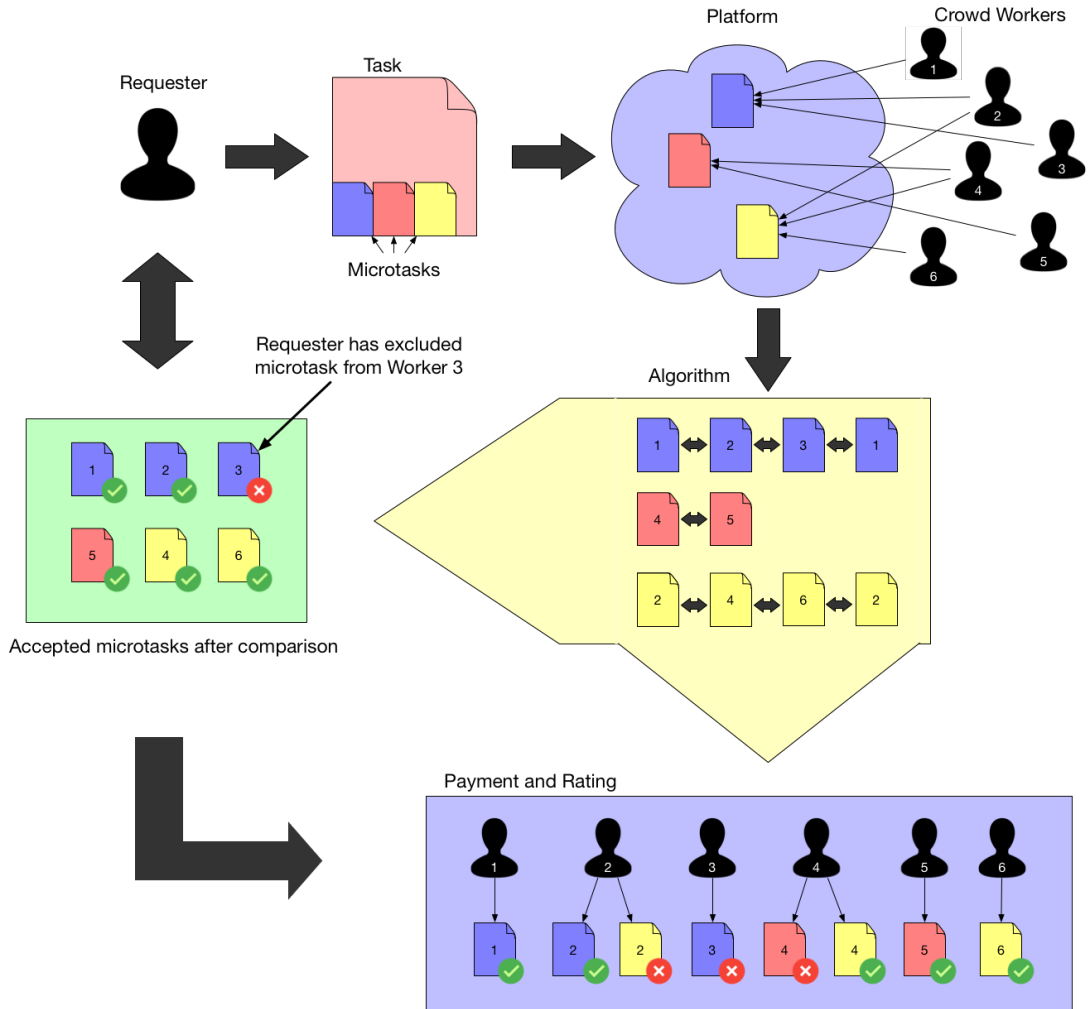


Figure 1.3: General process of the crowd work procedure with the involvement of the requester and the sorting of the results of the crowd workers by the algorithm of the respective crowd work platform.

surveyed had a university degree. In microtasking, 69% of the workers have finished their education [18]. Also, other researchers such as Irani [19] and Kittur et al. [20] describe a large number of highly educated workers who work full-time on crowd work platforms.

- **Mistrust and Irresponsibility as Result of Economic Efficiency and Technical Practicability**

As mentioned before, the approval of microtasks is automated through an algorithm [3]. If someone has been unjustly treated, it is not possible to lodge a complaint [3]. On these platforms there is certain irresponsibility since both parties (workers and

requesters) are anonymous [3]. The consequence is a lack of confidence on both sides and actually there are some actors who confirm this mistrust. For example, workers assess qualifications wrong or the requester provides a wrong description for the tasks and so it is possible that mistakes happen unintentionally [3]. In addition, fraud attempts are not unusual on both sides [3]. The economic efficiency and technical practicability of microtasking depend on workers who are not hand-picked [3]. If someone is choosing a certain worker, then it will not be crowd work anymore [3].

AMT does not bother about the problems of the workers but simply takes the money from requesters [16]. In addition, the workers believe that they do not get any attention from AMT and that AMT has forgotten about the fact that their platform does not work without workers [16]. Researchers have argued that this human computation in the current form will encourage invisibility of the workers, which in turn will eliminate the moral incentive for the necessities and working conditions of the workers [16]. Therefore, the crowd workers do not feel any appreciation for their work nor receive a fair payment [16].

- **The Vicious Cycle**

At the beginning of their career, crowd workers have to take every little and simple task, which is rewarded with a few cents. It also depends on the branch, but crowd workers with more experiences have it easier to find new jobs [18]. So, crowd workers will get better-paid tasks if they deliver consistently good work. Therefore, workers will dive in even more in this system. The recommendation and reputation system is to blame for that.

- **Noisy Rating and Two-Class-Society on Platforms**

Ratings and recommendation systems are an important factor on crowdsourcing platforms, because they strongly depend on reviews of requesters, e.g., the task acceptance rate states a lot about a worker [16] [21]. There is a deluge of jobs and crowd workers [3]. This deluge could only be handled by an algorithm on the platform [3]. The algorithm shows the tasks to the workers and also compares their results [3]. This phenomenon is called „Algocracy“ [22] [23]. The worker chooses a microtask on her own and, if the result is incorrect, it will be sorted out and the worker will get a bad rating [3]. This sorting process can affect the worker negatively in several ways, e.g., getting a bad rating or getting blocked [3].

The rating is very important for each worker [16] [18]. A new worker starts with little and low-paid tasks and with every good job the rating is rising [16] [18]. A worker with low rating will highly unlikely see the same tasks as a worker with higher rating [16]. The rating decides which job a worker could make or not [3]. At AMT there exists a certain category for workers with a high rating called „Master“ [24]. There are tasks on AMT, which are only for turkers in this special category [24]. On some crowd work platforms the rating is based on the requesters' feedback and these reviews are notoriously inflated and noisy [21] [25]. Furthermore, the

ratings are important for the platforms, because, if the platform has many workers with a high reputation, the work they do is also of high quality [26].

All these points show the predicament of the workers and lead to the questions: "why do people crowd work?" and "why do they not quit?" As many software systems, crowd work platforms are an invention with the goal to help people.

In the first place crowd work was invented for people who cannot do an ordinary job, e.g., disabled people or people who have to stay at home because of children or nursing care. It should give people the opportunity to earn easily some money as a sideline. Besides the financial aspect, there are other important factors: fun and learning effects, pastime and altruistic reasons [27]. Over the years, the systems and workers are more and more exploited, and now full-time jobs are outsourced to the crowd [18].

The second question is also easy to answer. Many people who do crowd working, especially those who see this activity as a secondary employment, perceive the work only as a job, which they do for a short period [3]. With every good rating they get better jobs, and then they stay on the platform - the vicious cycle - and soon weeks or months become years [3]. For crowd workers it is not easy to find a job in the real world, because they have a gap in their resume and with "clickwork" a worker has not learned some special skills.

There are a huge number of workers who complain about this system. Therefore, workers have started to interact with each other to be not anymore this anonymous collection of people [28]. They exchange information in forums, chats and social media or even in person [28]. This is the answer from the crowd workers regarding the inequalities on these platforms. They do not form other actions, because they do not have the time or as mentioned before, it is only a secondary employment and not worth to spend time to create a system to help each other [3].

Therefore, other researchers took on the problem and designed a few solutions from which most of them are not very successful, except for Turkopticon¹. Figure 1.4 shows that crowd workers currently use many systems, forums or social media. They are divided into these different forms, which causes three main key challenges. The first key challenge is that information is distributed, which leads to the second key challenge that workers do not know which channel they should choose to get relevant information. The third key challenge is the locked information. Especially in social media groups information is not easily accessible and crowd workers need to register on many platforms to get all important information they need and are looking for.

Another motivation for this work is that worker organizations know very well about these problems and had also formed a few actions, but they need more information to take better action against the current problems and limitations on a political level [29]. Also, the European Parliament has recently conducted a study and gained more perception for this problem.

¹<https://turkopticon.ucsd.edu/>. Accessed 26.3.2018

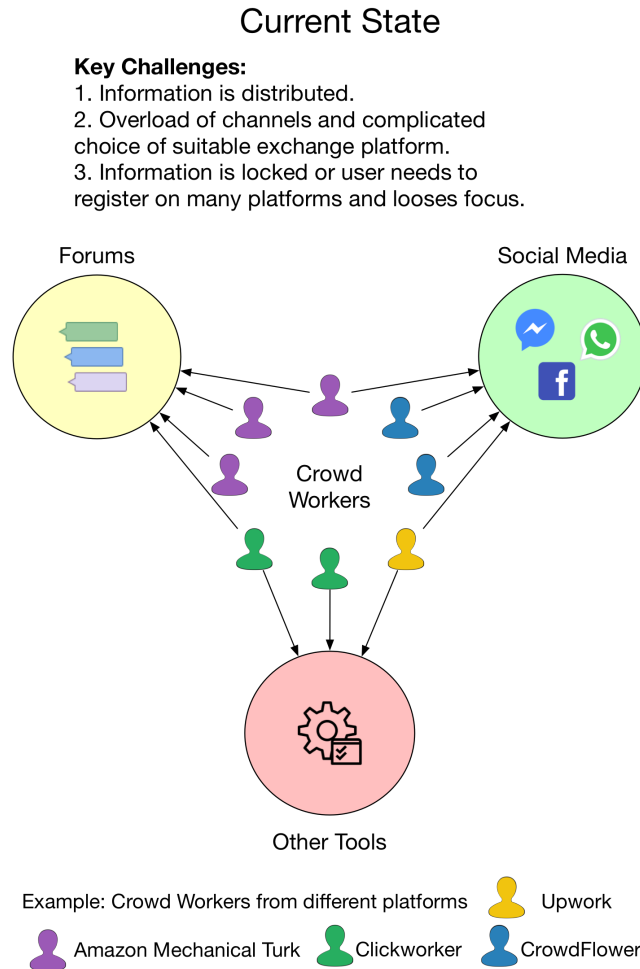


Figure 1.4: Current state of knowledge sharing and self-organization of crowd workers

Work 4.0 is nowadays a hot topic. Exploitation and undermining of human and labor laws is not something, which should go hand in hand with Work 4.0 and democracy. This work should be a step in the right direction and give guidelines how to help crowd workers in knowledge sharing and self-organization.

1.2 Problem Statement

Crowd work is a socio-technical application domain that can be investigated from numerous scientific fields and perspectives, whereby two of those perspectives are particularly interesting. Firstly, from a societal perspective: A recent study [30] on the crowd work landscape in Austria funded by the *Austrian Chamber of Labor* identified the following negative aspects: Workers only get paid for the output, regardless of the actual time invested in finishing the task. It may happen that the workers do not get paid even

though they completed the work, because the requester does not accept the work product. Also, crowd workers predominantly work as independent contractors, thus they need to pay social security themselves. In another recent work, Risak and Lutz [31] investigated the legal challenges of crowd work. Among many aspects, they identified a lack of transparency of internal platform mechanisms (like reputation systems) as particular issues.

Secondly, from a technical perspective: In the past, platforms have been proposed to support crowd workers to organize [15] and share knowledge [16]. However, with respect to software architecture, these platforms are typically designed in an ad hoc manner and are poorly documented. This leads to error-prone designs and implementations, but most importantly, it is difficult for other software architects and designers to systematically assess, replicate and adapt the designs of such platforms for their system-of-interest and their crowd work community, its needs and concerns.

Our overall goal is to systematically explore a central solution approach to help crowd workers in knowledge sharing and self-organization as we can see in Figure 1.5. Therefore, we address these technical and societal limitations in the form of research question. We study the needs and system requirements of crowd workers and design an environment for them. Also, the research questions aim to identify the needs and business processes of crowd workers in order to formulate design principles and create architectural models which cover the identified needs, and finally to check if these established principles and models are correct and useful for crowd workers and relevant stakeholders.

To address the problems and the formulated research questions, this thesis presents a collection of design principles and criteria that help software architects and systems designers of collaborative platforms to support crowd work communities on the societal level with transparency and knowledge sharing capabilities. The presented principles address conceptual and architectural aspects that are specific to the crowd work domain. Based on the identified seven design principles, architecture models for a central software system solution approach were created that address the needs and requirements of the stakeholders. Furthermore, we apply and evaluate the proposed architecture design and models in the context of a prototypical system implementation together with stakeholders from Austrian worker organizations and subject matter experts from the crowd work industry. An approach of a central system is shown in Figure 1.5 where we can see the current state, which we saw before in Figure 1.4 and the target state which is providing a central platform. With a central system the existing key challenges are addressed by providing a centralized information hub. One channel is suitable for all problems and discussions, and the information is accessible for every user. To verify if these principles and models fit the needs of the crowd workers and stakeholders, we conduct a survey and ask about the usefulness and practicability of the prototype and the principles and models itself. With these results, we can adapt our principles and derive future work.

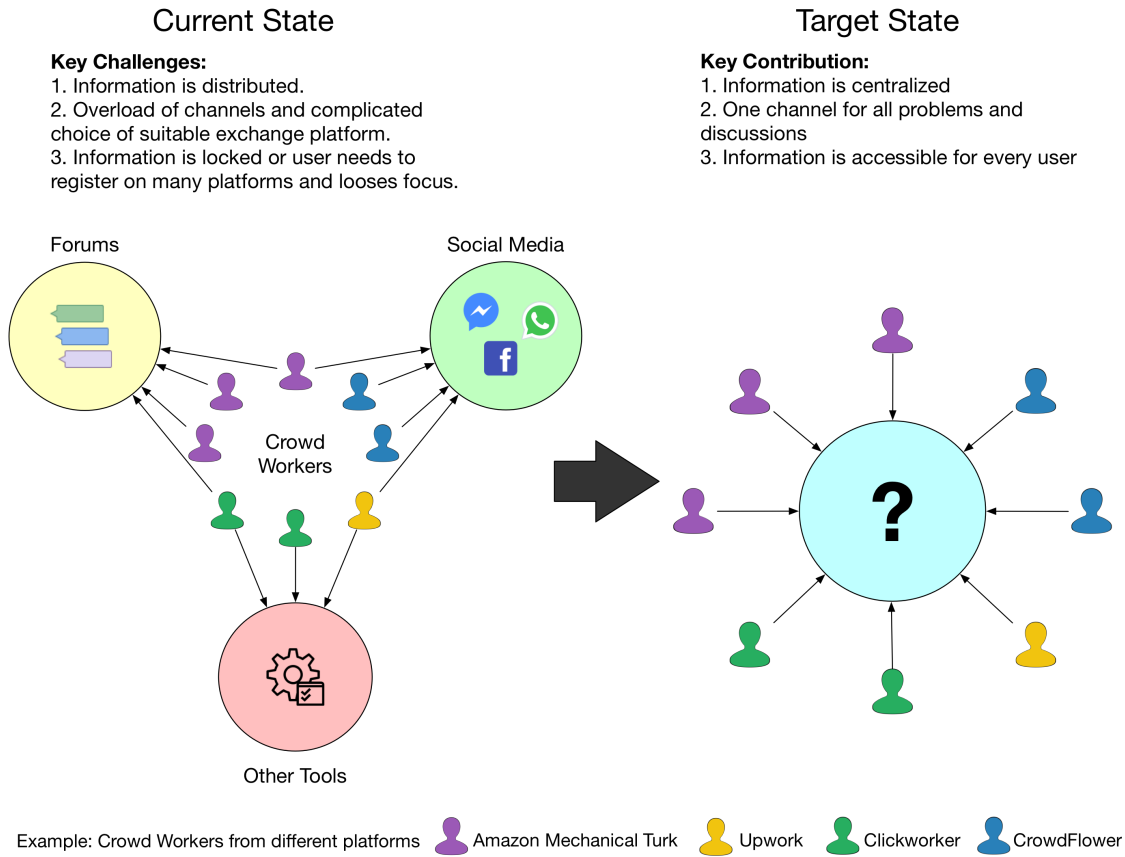


Figure 1.5: Target state of knowledge sharing and self-organization of crowd workers

1.3 Structure of the Work

The remainder of this work is structured as follows: Chapter 2 presents the related work and background information. Chapter 3 discusses our expected results, research questions and methodological approach. Chapter 4 describes the key stakeholders, their concerns and requirements. In chapter 5 the design principles are presented according to the findings and stakeholder needs. In chapter 6 the architectural models of a platform for crowd workers which covers all their needs is presented. In chapter 7 the prototype WorkerHub is described which is evaluated and discussed in chapter 8. Finally, in chapter 9 conclusions and future work is discussed.

Background & Related Work

The related literature presented here covers the scientific fields relevant in the context of this thesis. Therefore, the first part of this chapter deals with crowd work in general, the advantages and drawbacks, the network within the crowd and some studies which analyzed the current situation of the workers. The second part introduces relevant approaches which aim to help workers to improve their working life. In the third part we will take a look at collective intelligent systems and how they can support crowd workers.

2.1 Crowd Work

In this section we discuss crowd work in general to provide a better background of this topic. The presented studies describe the current situation of crowd work, the working conditions of the workers, their challenges and how they ease their work.

In her work, Simperl [2] describes different forms and types of crowdsourcing today where she differentiated between microtasks (crowdsourcing for routine work broken down into smaller, independent units), macrotasks (closer to classical outsourcing), crowdfunding (funding a venture with the help of monetary contributions from the crowd) and contests (competitions targeting grand scientific, technology, business, or social questions). In her paper she distinguish between four dimensions, which helps us to understand the complexity to design a microtask:

1. To define what the crowdsourcing is.
2. To evaluate who the crowd is.
3. To determine how to crowdsource.
4. To identify the right incentive.

The definition of each dimension is crucial to the quality of the results. So, Simperl [2] illuminated the different shades of crowd working. Therefore, this work is foundation to our work because of the definition of microtasks and the different dimensions which gives us a taste of the complexity to design a task. Also, it helps to understand that creating the ideal microtask is quite complex and highlights the importance of the involvement of requesters for explanations of tasks.

Leimeister et al. [32] examined crowdsourcing and crowd work in general. Thereby, they took a closer look on the phases how a task can be outsourced successfully and stated that this depends on the form of work. The form of work (e.g., competitive approach or teamwork approach) gives a better insight into how workers can be attracted to a task, how much the compensation is, which legal framework has to be considered, etc. Big advantages for workers are a high flexibility and a degree of autonomy. Low compensation and missing legal frameworks are major disadvantages for crowd workers. For a requester advantages are a fast task completion, access to more knowledge and low costs. But also there are some disadvantages like losing of intern knowledge and extra costs, which are hard to calculate.

Another work from Leimeister et al. [33] gives more insights of the phases of a crowdsourcing process from the point of view of the requester. These phases include first a specification of the task, then a selection of crowd sources, after that the execution of tasks, then the aggregation and selection of a solution and finally the remuneration of the crowd workers. Furthermore, they evaluated in which steps in the software development process crowd work can be used. For example: crowd funding can be used for budget allocation, or a coding contest can be used for implementation. This work shows that even for the requester, the effort of generating such a task is high. Therefore, the importance of a good quality of work increases with the complexity of the task. These contributions help us to understand crowd work from the view of a requester and getting better insights into crowd work and aspects, which go along with crowdsourcing.

Alkhatib et al. [9] examined parallels between crowd work and gig work and piecework in the 20th century. This work draws attention to the bad working conditions and tries to compare it with piecework. Therefore, Alkhatib et al. emphasized the danger of crowd work and the need of change of the working conditions. This work builds also on this paper. In addition, they took a closer look on the history of piecework because nowadays in the context of on-demand work there is a similar trend of work decomposition, distribution and payment. Therefore, they addressed three questions:

1. What are the complexity limits of on-demand work?
2. How far can work be decomposed into smaller microtasks?
3. What will work and the place of work look like for workers?

The paper presents a theoretical foundation of the most persistent questions in crowd work by doing research investigating the mechanism that enables and limits piecework historically. Also, the authors identified pitfalls, limits and differences of on-demand work. The differences between historical piecework and modern on-demand work limit the analysis. Also, the foreseen and unforeseen is a problem. Alkhatib et al. suspected that on-demand labor will follow the same trajectory of worker empowerment that piecework saw and that Utopian and dystopian outcomes will both occur, but in different parts of the world and to different groups of people. They also see a potential in the management of people. It will make a big difference if the crowd worker will not be managed by an algorithm but in a more personal way. Their final finding was that there is a reciprocal relationship between modern on-demand work and piecework because on-demand work can teach us about the broader phenomenon of piecework as well as piecework from on-demand work.

Schmidt [34] examined crowd work from the perspective of crowd workers in his study. He defined ethical principles especially with a view on payment. This paper shows the bad working conditions of the workers particularly when it comes to fairness, respect and economic sustainability. This contribution is motivation and foundation for this work to help crowd workers.

Some papers have a special focus on crowd work in the legal sense and give a basis for further improvements on the political side, which covers the motivation for this thesis. In a recent study, the authors evaluated crowd work platforms and they came across legal contradictions [35]. For example, the right to organize is a human right and some platforms decline this right. Another example is that the workers had a right of social insurance. Also, Felstinerf [36] took a closer look on the legal rights of crowd workers. It is a problem that nobody knows if a crowd worker is an employee or a freelancer. It is also difficult for a labor union, because they do not know if a crowd worker is part of their union.

2.1.1 Network of Workers

Some studies and papers state that crowd workers want to connect with each other. Therefore, crowd workers join forums and social media and thereby form a network. How do they do that and what do they expect from their "colleagues"?

Gray et al. [28] defined the crowd as a collaborative network. They showed that crowd workers collaborate to fulfill technical and social needs left by the platform they work on. They identified that workers are not independent, autonomous workers, but work with other crowd workers. There are three different reasons why workers work together. First, they have to manage the administrative overhead associated with crowd work. This overhead includes creating accounts on crowdsourcing platforms, avoiding employment cams and collecting checks. Second, they need to find lucrative tasks and reputable requesters. They found out that workers notify each other when a high quality task is

online. Third, they would like to recreate the social connections like working together on a task. The work combines ethnography, interviews, survey data and large scale data analysis from four different crowdsourcing platforms: Amazon's publicly available *Amazon Mechanical Turk (AMT)*, Microsoft's proprietary *Universal Human Relevance System (UHRS)*, the social entrepreneurial commercial start-up *LeadGenius*, and the not-for-profit site dedicated to translating content for transnational audiences *Amara.org*.

Ying et al. [24] investigated AMT deeper. They tried to map the entire communication network of workers on AMT, which is shown in Figure 2.1. To do so, they designed the task, also named HIT on AMT, in such a way so that over 10,000 workers from across the global self-report their communication links to other workers. This task should show that there is a substantial communication network within the crowd. However, their main contribution is a map of the entire communication network of workers on AMT (Figure 2.1). In this HIT, the workers get a few questions and can decide if they want to share this information with the other workers who also did this task. It is also possible that they can record at anytime with whom they communicate. 5,268 edges are added among all connected workers and the largest connected component consisted of 994 workers.

83 % of the connected workers reported that they are using at least one forum. Another interesting outcome is that in the U.S. the workers prefer communication via forums and in other countries one-to-one. Their results showed that workers share lucrative tasks and information about reputable requesters with each other. This connection enables workers to get more high quality tasks before other workers hear from them. In some cases, it could be that a worker who is not connected to other workers will be isolated and will not get lucrative tasks. Also, this paper underlines the connections between workers. Hence, it supports the goal of this thesis to build a platform for crowd workers and is a contribution where this work builds on.

Huang and Fu [37] conducted an experiment where they wanted to find out if the outcome of a task will be better if the transparency of workers increases. The generation of high-quality outcomes is outlined in Figure 2.2. This assumption is proven in their study and so it is a relevant contribution, which supports the goal of this work. The results from crowd working are significant better when social transparency of the workers is increasing and this enables the workers to work together. They also mentioned that this kind of cooperation needs a new reward system which can additionally motivate the workers.

Oppositely, there is a study that experiments with the introduction of a guild. The guild is a good example for an efficient system where quality of work and centralization of workers counts. Hence, it is a contribution that this work builds on, but it also shows some limitations of a social platform at a crowdsourcing platform. For example: To motivate crowd worker to rate the work of their colleagues, the platform had to pay them some money. Also, the system is quite complex and the platform cannot guarantee a fair rating. Whiting et al. [26] focused on crowd guilds and how they could help crowd workers. They stated that crowd work happens in a distributed and decentralized way all over the world. This decentralization undercuts behaviors and institutions that are

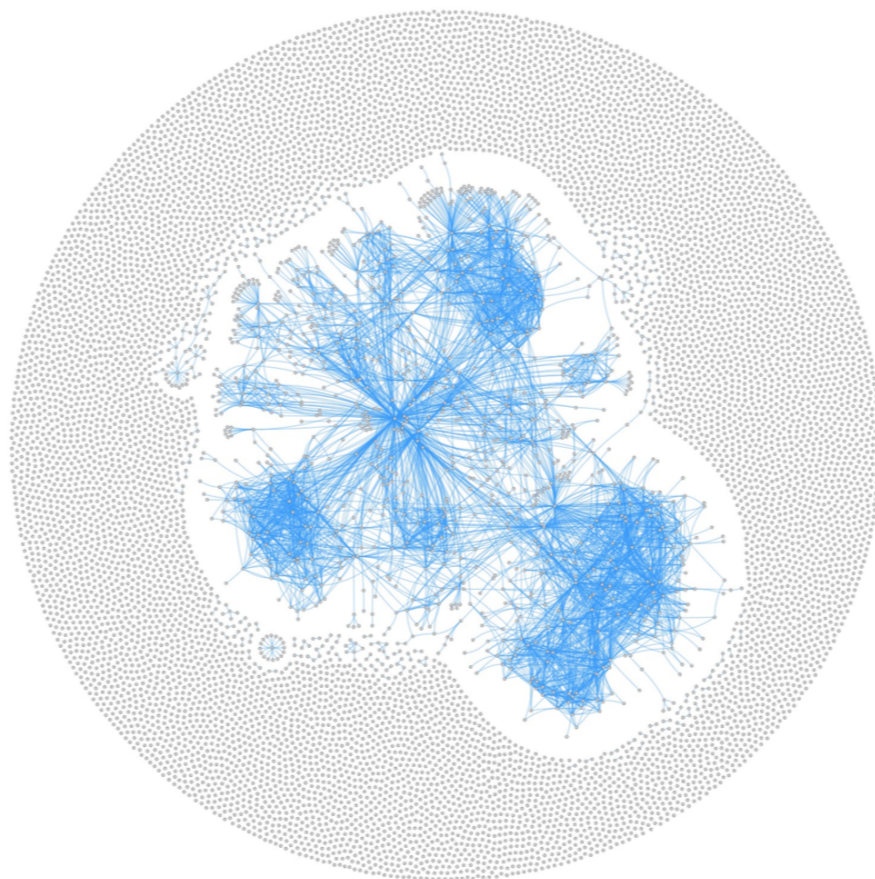


Figure 2.1: The communication network among Amazon Mechanical Turk Workers [24].

critical to high-quality work. Therefore, they drew inspiration from historical worker guilds to design and implement their prototype *daemo*, which supports the building of crowd guilds. This guild consists of centralized groups of crowd workers who collectively certify each other's quality. The focus is on reputation. It will be realized by applying a double-blind peer assessment. The process is shown in Figure 2.3. These assessments will be performed from other crowd workers. Whiting et al. introduced levels within the worker group. Ideally a crowd worker with a higher level assesses the work from a worker with a lower level. The reason for this kind of review is that the quality of work can increase and therefore, also the wage for the task will increase, because the requester can be sure that the work is done in a certain quality. The review process will also be promoted with a little compensation. So every reviewer gets paid for reviewing.

The result from this experiment was that almost all crowd workers rather tend to centralization than to decentralization. Only a few crowd workers did not like the idea, because of destroying the spirit of crowd work. Another result was that workers in the guilds provide one another with more practical and actionable feedback and advice. The

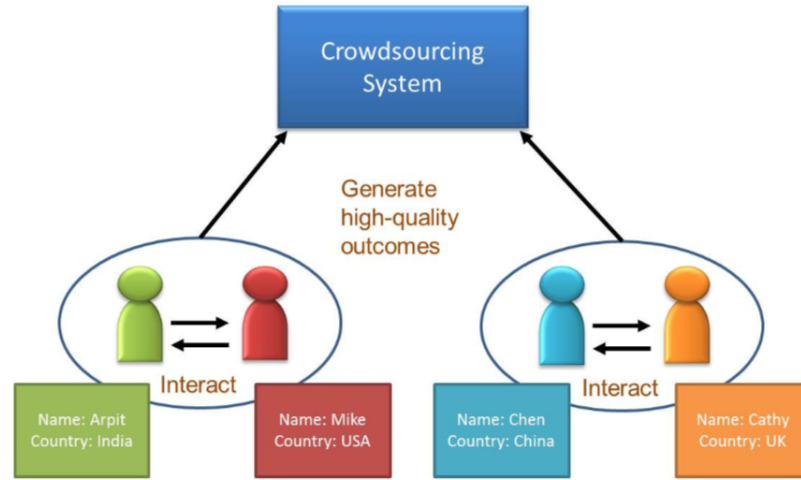


Figure 2.2: Interaction between workers to generate high-quality outcomes [37].

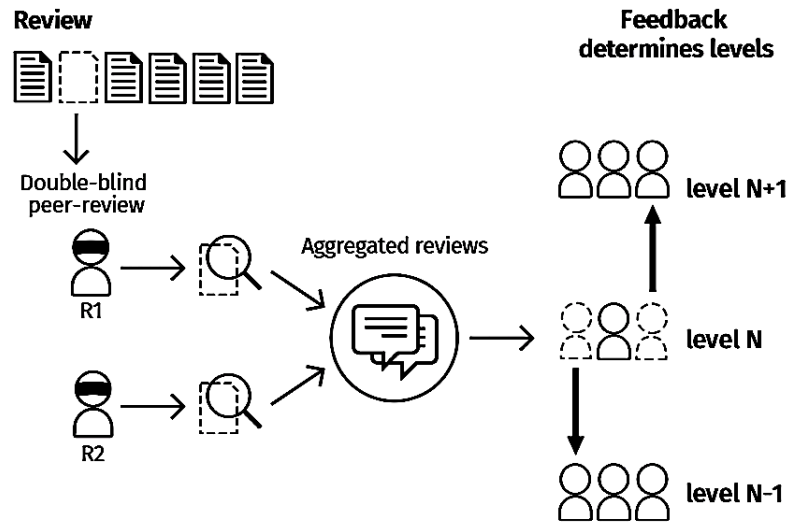


Figure 2.3: The review process of the guild [26]

ratings from the guild’s peer assessment were a significantly better predictor of workers’ actual accuracy than the workers’ acceptance rates on Amazon Mechanical Turk. Also, these results showed that crowd guilds produce more accurate reputation information than no guilds.

These contributions are important for this work because they show that crowd workers are social and want to engage with each other to overcome problems, which come along with crowd work.

2.1.2 Crowd Work-related Studies

In the following we summarize the results of some general studies about the working life of a crowd worker. These studies are a motivation factor of this work to help crowd workers in knowledge sharing and self-organization. Hence, the results of these studies represent an important basis of the thesis.

The Austrian Chamber of Labor [30] conducted a study about the crowd work scene in Austria. This study provides insights into the working life of a worker. They asked the workers about the regularity of their work, if it is a main or secondary employment, their personal income, the kind of work they do, ages, gender and about their requesters.

Schmidt [3] analyzed the different aspects of crowd work. Thereby, he showed the grievances on crowd work platforms and also the potential for improvements, which is important for this work. Figure 2.4 shows the main difference between cloud work and gig work, and some example platforms.

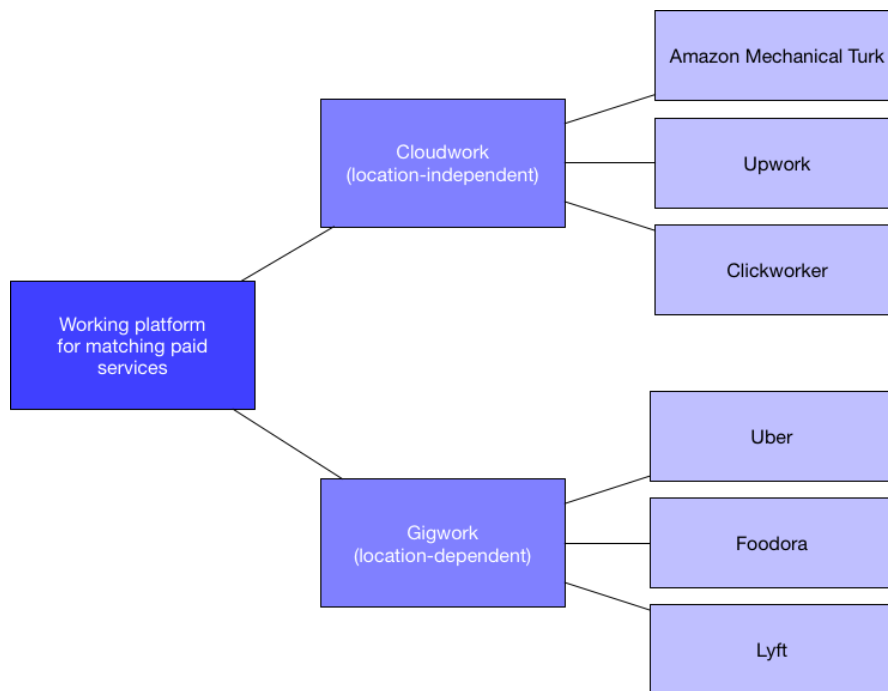


Figure 2.4: Overview of working platforms [3]

If the execution of the activity is transportable, meaning that everything is executed over the Internet, Schmidt talks about cloud work. Additionally, if it is also irrelevant who is going to do a certain task and therefore this task will be presented to an open and unspecific group, then he talks about crowd work. If this certain task is divided into minor tasks within this group to a fixed unit price, then it is called microtasking. A creativity contest occurs when the group solves one task and only one result will be chosen and paid. If a task is stationary and will be assigned to a certain time and specific

person, then he talks about gig work. In summary, Schmidt defined six fundamental types of paid services:

1. Freelancer marketplace (Cloud work)
2. Microtasking (Crowd work)
3. Creativity competition (Crowd work)
4. Catering, restaurants (Gig work)
5. Passenger transportation and delivery service (Gig work)
6. Household and personal service (Gig work)

He explained that this new business model depends on an army of millions of employees who are available or can be fired depending on the order situation. The operator of the platform relies on private individuals and autonomous persons who want to have an additional income. This business model turns out to be an attractive alternative to a regular job, because the individuals can decide when, where, for whom and how much they like to work. This flexibilization is accompanied by infiltration of social and labor law-related standards. Schmidt [3] analyzed some working platforms and made the following statements:

- It is a matter of online marketplaces with at least three parts (requester, platform operator and crowd worker).
- The operators of the platforms are intermediates and mediate between supply and demand.
- This kind of business model enables the operators of the platform to shift all costs, social and legal regulations to one of the other two parties. So it is possible that the platform can grow exponentially but without increasing the overheads.
- Only the operators of the platform have full access to all data and rules.
- This software architecture results in information and power asymmetry.

The socially explosive nature of these working platforms lies in the fact that the operators can avoid regulations of work and consumer protection, minimum wage and social security contributions according to Schmidt. Also, the operators can control how the two parties interact with each other and under which conditions or what an individual see upon the interface with the appliance of design (e.g., via interface) and legal conditions (e.g., via terms of service).

Also, Leimeister et al. [18] defined crowdsourcing and crowd work and their different forms in their study. They dealt with the world of the crowd workers and their environment.

Thus, they provided important insights on which this thesis can build on. In the first step they made an analysis of the crowd working market. In the course of this analysis they identified five different clusters: *Microtask platforms*, *Marketplace platforms*, *Design platforms*, *Testing platform* and *Innovation platforms*. These clusters are mapped in Figure 2.5 with examples of market-leading crowdsourcing platforms in Germany.

Analysierte Crowdsourcing-Plattformen⁴



Figure 2.5: Five clusters and the analyzed platforms [18].

In the second step Leimeister et al. [18] examined the working conditions of the crowd workers in Germany. 434 crowd workers who live in Germany or are Germans were interviewed regarding their social data, their payment, their working time and quantity, their precaution and protection, and their experiences and cognition of crowd working. This analysis of the work of crowd workers was the first analysis in this field, which was based on empirical data. The results of the study delivered important insights into the structure and composition of the crowd as well as the working conditions. In the following some of the insights of this analysis:

- Crowd workers are working on average for two different platforms.
- Most of the workers have a good to very good education.
- The distribution of gender is over all platform types balanced, but on marketplace platforms the amount of women is up to 38% which is low compared to the other platforms.

- Most of the workers, who are doing crowd working full-time, prefer the same job as an employee in a company.
- A full-time crowd worker earns in average 1.500 Euros per month and work 80 hours a week.
- 66 % insure themselves against illness and unemployment. Only 53% are taking care about their pensions scheme.
- On average the crowd workers are not feeling exploited, but they are also not satisfied with their working environment.

The starting point of the study by Al-Ani and Stumpp [38] was that crowd workers are difficult to classify according to labor law and labor unions are unsure how to deal with crowd workers. The study evaluated the living situation and motivation of the workers as well as their expectations on labor unions. An interface between labor union and workers will also be crucial according to their study. Therefore, this work encourages us to involve unions in our platform.

There are lots of forums, which workers use to connect. One forum was analyzed and gave insights into the viewpoint of a Turker (crowd workers of Amazon Mechanical Turk). Martin et al. [39] conducted an ethnomethodological analysis of the forum [TurkerNation](http://turkernation.com/)<http://turkernation.com/>. Accessed 26.3.2018. This analysis gave insights into the methods of a turker, turker-requester-relationships and the perspective of a turker. The whole analysis is considered under following aspects: practical, emotional and ethical. The results showed that for turkers the most important factor to choose a task is the payment. Also, one of their biggest concerns is how to make good decisions on selection jobs and having a good relationship to the requesters. A suggestion from Martin et al. [39] is to implement a forum like [TurkerNation](http://turkernation.com/), which helps to reduce information deficits and promotes better collective actions.

2.2 Socio-technical Systems for Crowd Workers

In recent years, researchers identified the inequality of crowd workers on crowd work platforms and thus they developed software systems to help them. The described contributions in this section represent an important foundation on which this thesis builds on, but also reveals limitations that this thesis aims to address.

Salehi et al. [15] studied challenges of collective action efforts in relation to online labor. Based on an ethnographic fieldwork, they understood the workers' barriers to collective action. In the study they talked to workers about their relationships to other workers and the collective action. They created a platform to support the Amazon Mechanical Turk community in forming publicly around issues and then mobilizing. This platform is called [Dynamo](http://www.wearedynamo.org/)¹. The idea is to gather ideas, energy, and support directed towards collective

¹<http://www.wearedynamo.org/>. Accessed 26.3.2018

action. The design of Dynamo focused on three principles: trust and privacy, assembling a public, and mobilizing. These three design choices are based on the interaction with workers. Trust and privacy are very important aspects, because some workers are afraid of Amazon and that if Amazon finds out what they do, they might lose their account and so their income. In assembling a public, Dynamo focuses on idea pitches, which are illustrated in Figure 2.6. These idea pitches act as polls that enable public to form around them. So workers can vote on such idea pitches.

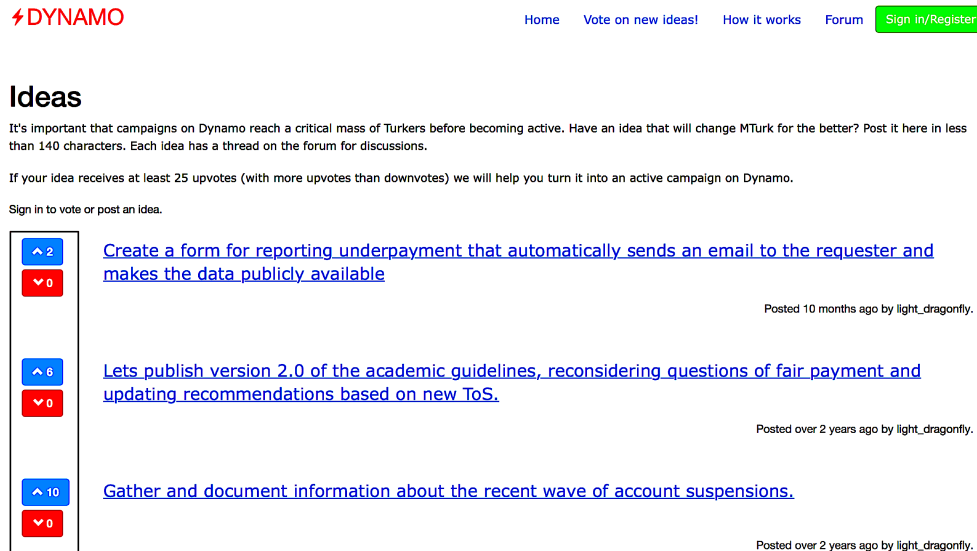


Figure 2.6: Voting system in Dynamo.

The idea will become a campaign if the idea gets at least 25 up votes and has more up votes than down votes. In the lifetime of Dynamo, only two ideas made it to campaigns: Guidelines for ethical research on Amazon Mechanical Turk and a „Letter Writing Campaign“ to create a positive image of Turkers in the public eye. This paper provides a very good insight into collective actions of crowd workers and their limitations.

Irani and Silberman [16] analyzed the technical relationship between the workers and the requesters on Amazon Mechanical Turk. The outcome raised a number of issues:

- More than half of the workers felt that their work was regularly rejected unfairly or arbitrarily.
- 40 % demanded a faster payment. The requester has 30 days to evaluate and pay for the work.
- 20 % mentioned “fair” compensation generally.
- 12 % expressed dissatisfaction with the requesters and Amazon’s lack of response to their concerns.

- 10% want a “minimum wage” or “minimum payment” per HIT.

Turkopticon² was introduced as a response to the hazards of crowd work. Turkopticon is an activist system that allows workers to publicize and evaluate their relationships with requesters. It is an ethically motivated response to workers’ invisibility in the design of Amazon Mechanical Turk. Amazon Mechanical Turk workers can rate requesters in four different categories and also have the possibility to write a comment. Figure 2.7 illustrates a rating of a requester and the interface of Turkopticon. The four categories are:

- Communicativity: How responsive has this employer been to communications or concerns you have raised?
- Generosity: How well has this employer paid for the amount of time their HIT’s take?
- Fairness: How fair has this employer been in approving or rejecting your work?
- Promptness: How promptly has this employer approved your work and paid?

Crowd workers can install Turkopticon in their browser as a browser extension that augments workers’ view of their AMT HIT lists with information other workers have provided about requesters. The findings of this contribution and the success of Turkopticon is a basis for our research.

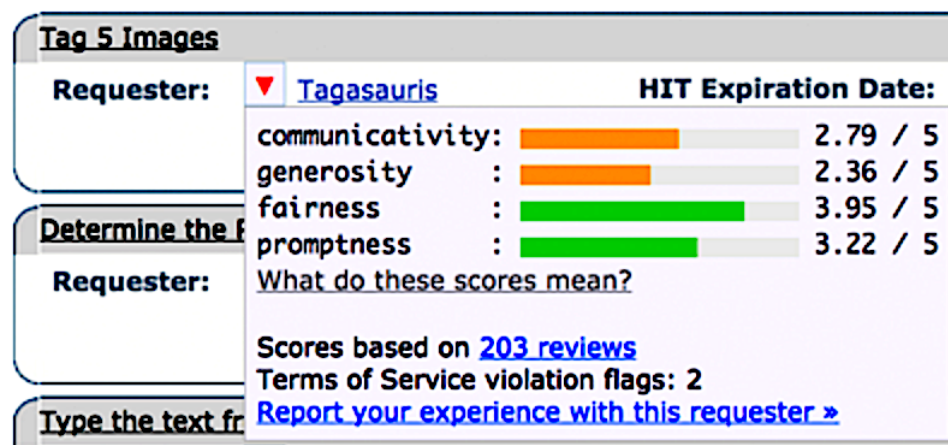


Figure 2.7: A screenshot of Turkopticon [16].

Totally contrary are the two papers from Vakharia and Lease where they examined crowd work platforms to find platforms with better working conditions of the workers than on

²<https://turkopticon.ucsd.edu/>. Accessed 26.3.2018

AMT. In the first paper, Vakharia and Lease [40] evaluated seven crowd work platforms in contrast to AMT. They formulated key criteria on the base of the analysis of AMT to compare different platforms. With regard to the criteria "ethic" and "sustainability" they found out that there are some crowd work platforms, which have humane working conditions or medical benefits. In their subsequent work, Vakharia and Lease [41] defined more detailed superior criteria and compared more platforms with AMT. They identified some issues across all examined platforms like insufficient support for crowd workers. In their opinion, future research should more focus on alternative platforms than comparing them with AMT. These two publications show another way to help crowd workers by finding crowd work platforms with better conditions. Therefore, the two papers are contributions where this thesis builds on because we want to motivate for an exchange of information among all platforms to get a better overview of the current circumstances.

Another approach to improve the working conditions for the workers is to build a crowd work platform with better working conditions addressing workers from all platforms. This approach originates from Kittur et al. [20]. In their paper they asked themselves how should the future working place look like where they want their children to work. The framework presents research addressing 12 areas including quality control, reputation and motivation. The goal of this work was to generate a better system, work and experiences for the next generation. This contribution depends on a new crowd work platform. This new approach motivates us to explore an independent platform for crowd workers because it offers no advantages to the platform operator and it will be hard to establish such a platform referring to the giant AMT.

2.3 Collective Intelligence Systems

Collective intelligence refers to the value that is created by collective contributions of all people who share their knowledge, e.g., in forums [42]. Having all this knowledge aggregated in a computing system with the cognitive capabilities of human groups then the system is called collective intelligence system [43]. Wikis, social networks and content sharing platforms are examples of collective intelligence systems [43]. Therefore, a collective intelligence system is suitable for our approach of a central system. Before getting deeper into the topic of collective intelligence systems, the following paragraphs introduce the areas of human computer interaction (HCI) and computer-supported cooperative work (CSCW) systems as introduction to collective intelligence systems (CIS). Both aspects are important parts for modeling a CIS.

2.3.1 Human Computer Interaction & Computer-supported Cooperative Work

Bigham [44] mentioned that HCI is tightly coupled with crowdsourcing. The human interface has to be simple and the tasks should be clearly visualized so it is easy to understand them. He defined three types of crowdsourcing and their challenges for the HCI in the context of crowdsourcing and collective intelligence. The first type is *directed*

crowdsourcing, which means that an algorithm or person directs the workers to do a specific goal. The second type is *collaborative crowdsourcing*, where the motivation of the workers is intrinsic which means that most of the work is done as a volunteerism. The last type is called *passive crowdsourcing*, which means that the results are coming as a side-effect of other crowd work, e.g., the web search and click behavior of a crowd worker. His contributions are relevant for this thesis with regard to the system architecture design as well as to the needs of crowd workers.

Hollan et al. [45] dealt with the new way we perceive things, because the world is getting more and more complex and computer-mediated content increases every day. They examined how HCI changed in that way and how distributed perception affects our activities. They observed that a focus on the distribution of cognitive processes through members of social groups and coordination between inner and outer structure is important. Hence, it influenced this work to have a clear structure and keep an eye on key processes, which are important for crowd workers to give them a higher priority.

Ackerman [46] followed up with the gap between social and technical requirements and mechanism of a CSCW system. He found out that it is important to understand how humans really work and live in a group, organization and community to design a good CSCW system. One important result that there is a fundamental gap between what we can do technically and what is required socially. So, it is important to understand the boundaries of a CSCW system. This contribution reminds us about the limitations of a CSCW system and where we should put our attention to in the process of design such a system.

2.3.2 Collective Intelligence

Malone et al. [47] evaluated the division of collective intelligence in genomes. The goal of their work was to understand collective intelligence better and how to combine different genomes to get the desired capabilities. Also, they examined under which conditions a genome is important or not. Figure 2.8 shows the questions used to classify the genomes. Therefore, the paper gives us good guidance to define important elements of collective intelligence.

Bonabeau [48] analyzed whether collective intelligence leads to better decisions and discusses key issues. Collective intelligence shifts the decision of a company in a new way. There are two high-level tasks in the field of operation research: 1) Generating solution: This includes framing the problem and establishing a set of working assumptions about it. 2) An evaluation about the different alternatives in the first step. Collective intelligence can help to fulfill these two tasks. For example, humans tend to see patterns where no pattern exists and so collective intelligence can provide a diversity of viewpoints. At the first glance, an application that taps into collective intelligence for improved decision-making seemed to be a simple concept, but it can be very difficult to implement. So Bonabeau had defined a few key issues:

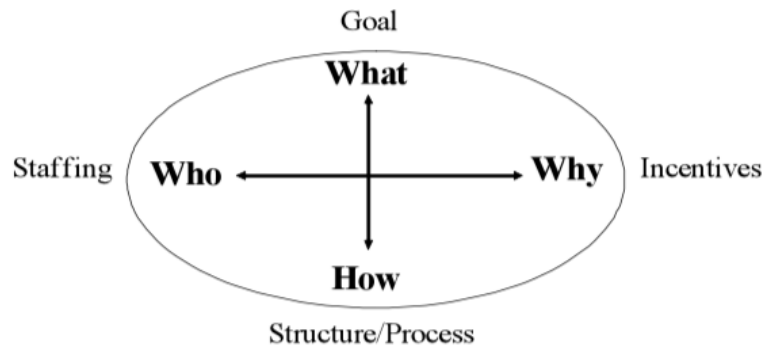


Figure 2.8: Elements of collective intelligence building "genes" [47].

- **Control.** A key concern is a loss of control. It can be that the outcome is unwanted and undesirable or it can be unpredictable, so the organization is not prepared to deal with it.
- **Diversity versus expertise.** It is important to have the right balance between diversity and expertise when a company makes use of a collective. Also, the composition of diversity is another factor. A large population can be skewed and may lead to distorted decisions. So the company had to decide which people should be involved based on the ability of those individuals to understand the problem at hand and collectively make positive contributions to solving it.
- **Engagement.** The motivation of the people can vary widely. Some will do it for cash rewards, prizes or other promotions, but it can also be a desire to transfer knowledge or share experiences. This combination with civic duty can be powerful motivators. The best example for this is Wikipedia.
- **Policing.** People who misbehave will raise with an increasing group size. Mutual policing can control such transgressions for which an implicit code of conduct helps govern people's behavior.

Collective decision-making is largely empirically driven. For every success story like Wikipedia ³ there are many projects that failed because of faulty mechanism design [48]. Any company that is developing a *Decision 2.0* application needs to understand some fundamental issues like the distinction between decentralized and distributed decision-making or the balance between diversity and expertise. The contribution of Bonabeau shows the power of collective intelligence and its application and pitfalls. Two key issues (engagement and policing) are also very important to this work, and the paper of Bonabeau [48] was the impetus.

³<https://www.wikipedia.org>. Accessed 12.8.2018

2.3.3 Collective Intelligence System Architecture

A collective intelligence system is a special variant of a IT-mediated collective intelligence [49]. With regard to the design and the architecture of a collective intelligence system (CIS), Musil et al. [50] evaluated the concept of stigmergy and proposed the model of a stigmergic information system that describes a CIS architecture. This model should support the development of such services in an effective and efficient way. The authors also identified that in the early stage design of such a system self-organization and feedback mechanisms play an important role [51]. In their subsequent work, Musil et al. [43] dealt with the systematic design of collective intelligence system architectures. They proposed an architecture framework for CIS (CIS-AF). First of all, this framework defines key principles of a CIS design and provides guidance to architects to describe and to design a CIS, which fits in the context and goals of an organization. The CIS-AF is grounded in a very detailed analysis of existing CIS. They investigated well-established CIS to derive key elements like stakeholders, processes, components, behavior, etc., and to identify the common foundational principles of CIS. The CIS-AF comprises three complementary viewpoints together with rules that express relations across the viewpoints:

- The context viewpoint describes the conventions to derive an architecture view that frames usefulness and the stigmergic mechanism that is central to the enduring aggregation and dissemination of knowledge in the CIS. Figure 2.9 shows us the meta-model how to design a As-Is Workflow/To-Be Workflow of the current business processes. Also, the meta-model of the stigmergic coordination model is defined in Figure 2.9. In the stigmergic coordination model is mapped the indirect communication among users.

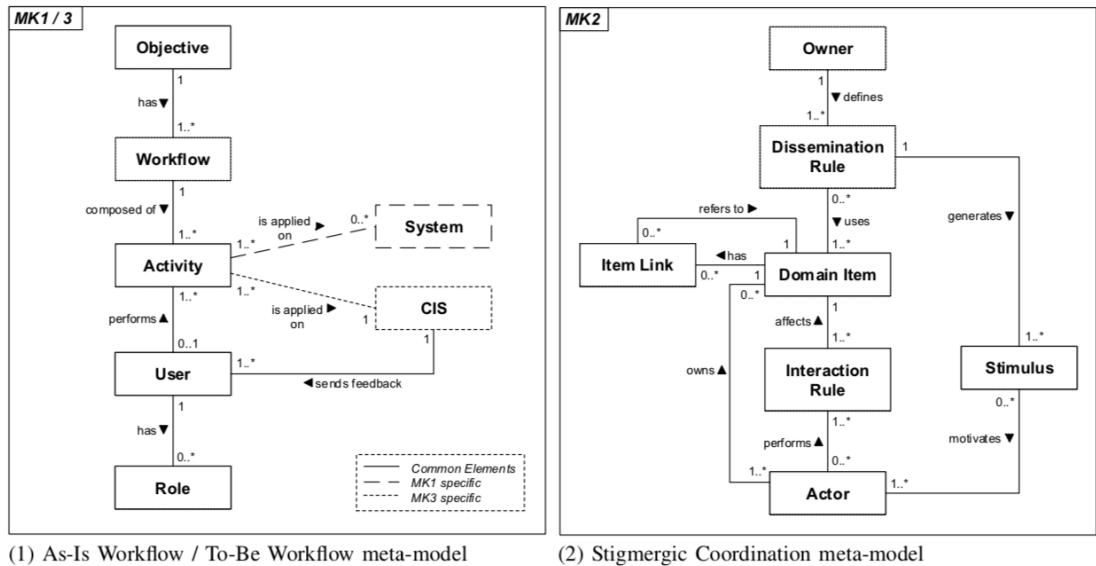


Figure 2.9: Meta-Models for the Context Viewpoint [43].

- The technical realization viewpoint describes the conventions to derive an architecture view that frames the data aggregation, knowledge dissemination, and interactivity concerns. Figure 2.10 presents three meta-models. The first one shows us the artifact definition meta-model where the main artifact is defined. The second meta-model is about the aggregation which gives insights into the agents, actor roles and their activities, and the dissemination meta-model is the last model where the content is disseminated back to user.

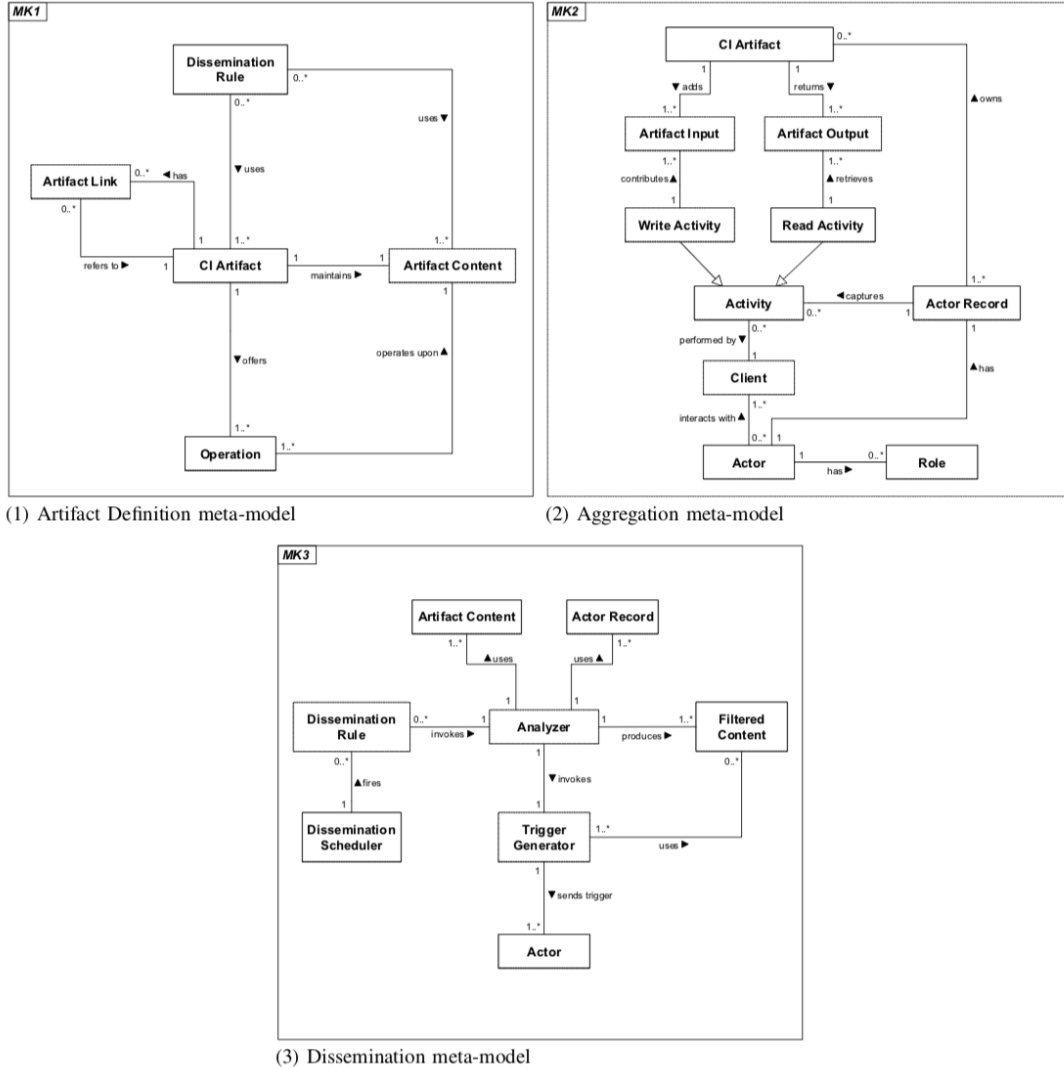


Figure 2.10: Meta-Models for the Technical Realization Viewpoint [43].

- The operation viewpoint describes the conventions to derive an architecture view that frames the kick start and monitoring concerns. Therefore, Figure 2.11 presents

the initial content acquisition meta-model where it is defined how to attract new members and the CI analytics meta-model where key metrics are defined.

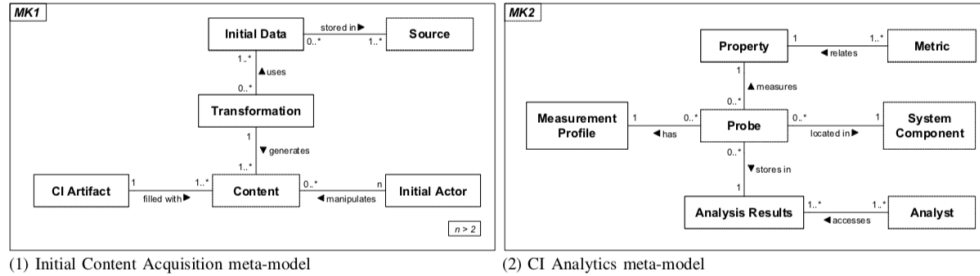


Figure 2.11: Meta-Models for the Operation Viewpoint [43].

2.4 Summary of the papers and their limitations

The presented papers make clear that the situation of crowd workers need to be improved and should gain more attention to societal stakeholders (politicians, unions). In the general part papers were described who underline the bad working conditions for crowd workers and also identified a responsibility to help them with negative examples like the non-compliance of human rights. Furthermore, some papers (e.g., Felstinerf [36], Schmidt [34]) call for actions on the politic sides and underline an understanding of the scope of the problem. These papers serve as motivation and underlines the importance of involving policy makers.

Other introduced papers explicitly proofed that workers are communicating among each other and that this communication is increasing the quality of their work, like the paper of Huang and Fu [37] showed. Different actions take place to increase the quality of their work and therefore the compensation for the work. For example Huang and Fu [37] aimed to improve the current situation with more transparency and Whithing et al. [26] tried it with the introduction of a guild. The limitation of these experiments is that these approaches are installed within the crowd work platform. Also, the limitations of these papers point out that it is important to define actions to bypass traditional crowd work platforms and therefore to develop a system that is independent of the crowd work platform.

The studies of the crowd workers give us a good insight into their needs and to form principles to support them. Also, they highlight the need to involve worker organizations like labor unions, which participate in the interaction with crowd worker. [38]. The key functions of forums are to reduce information deficits and promote collective action [39]. But a drawback of forums is usually their limitation to one big platform like Amazon Mechanical Turk and the non-existing involvement of worker organizations.

From other socio-technical systems we take some principles as important to be considered in a central solution design like trust and privacy from Dynamo and transparency from

Turkopticon. The problem with Dynamo was the huge effort of the campaigns and maintenance of the collective action. Turkopticon is a great addition to the work on AMT, but it is limited to the rating system. Turkopticon was a first step into the right direction, but the focus is only on support for AMT workers.

In the last section an introduction to collective intelligence and collective intelligence systems was provided. We showed how important the perception is and how humans perceive relationships. Collective intelligence enables bottom-up information sharing and knowledge aggregation [43]. A collective intelligence system lives from engagement and interaction between workers.

Research Approach

This chapter introduces the research challenges and research goal, which are derived from the motivation and problem statement. To solve the problems and achieving our goal, we formulate research questions. Furthermore, we give an overview over our approach, which results in section 3.3.

3.1 Research Challenges

The challenges lie in the different needs and concerns of the stakeholders. Crowd Workers has no time to take care about their rights because they are very busy in doing the work or it is not important for them [18] [3]. Nevertheless, they experience the disadvantages of crowd work like uneven payment [3], vague task description which results in average results [3], no health insurance [16], huge effort in finding and doing low-paid tasks [18], no possibility to exchange information and unfair treatments [3]. To overcome some of these grievances, they started to interact with each other over forums [28], which on the hand do not sufficiently fulfill their needs but on the other hand they have no time for developing a solution.

We assume that platform operators didn't want to change anything because for them the system works nearly perfect because they have the full power. On the side of the requesters arise a wish to get in contact with crowd workers [39]. It can be helpful to explain the task in more details to get a good result [39].

Worker organizations know from the problems of the crowd workers. They want to help them and need more transparency in the whole crowd working process [29]. Also, they want to get in contact with the crowd workers and collect data about their working life to do something against these grievances on the political side [29].

Figure 3.1 shows the research challenges:

1. Crowd workers, requesters and platform operators can get in contact with each other over the crowd working platform only in form of exchanging tasks [3]. Platform operators only monitor the platform and have full access [3]. Requester offers their tasks via an API over the platform [16]. Still there is no possibility where they can exchange detailed information and get in contact with each other e.g. to clarify tasks description or to communicate grievances [16].
2. Crowd workers are busy with their tasks and need to work 80 hours to earn as much as full-time employee [18]. Also, they see their job as a sideline [3] and blunder into the crowd work trap. The inequalities on crowd sourcing platforms increase and so crowd workers began to exchange information over different forums [28]. The result is an uncontrolled growth of different forums where the crowd worker don not know which channel is appropriate for a certain information and also the trustworthiness of the information on such platforms is questionable [26].
3. Members from worker organizations, which are also named in this thesis as advocacy want to help crowd workers which is proven from different studies like the one from the European Parliament [29]. The lack of transparency and a missing link to crowd workers make it hard to get in contact with them [29]. To overcome this issue a few worker organizations decided to build some information platform like the website faircrowd.work¹. On these sites are information about crowd working and crowd working platforms, but there is no chance to get in contact with crowd workers. They want more information about the working life of crowd workers to provide the needed support [29].

3.2 Research Questions

In Figure 3.2 we see the research goal. The goal is to build a central solution, which overcomes all identified challenges. The platform operator and requesters get a central platform to get in contact with each other. Also, the crowd workers have to choose one appropriate channel for their problems and they also get trustworthy information with the involvement of the advocacy. Furthermore, the advocacy has also access to central solution. So, the members of an advocacy can exchange information with the crowd workers and collect the necessary information to help them.

To fulfill this goal, it is to identify a set of design principles and investigate an appropriate software architecture design approach that supports software architects in the design and technical realization of socio-technical software platforms that support crowd workers with efficient knowledge sharing and self-organization capabilities. Such an approach should support software architects with design and decision support of assistive digital platforms that improve the transparency and collaboration among various crowd worker communities and domains. Based on this goal, this work address the following research questions:

¹<http://faircrowd.work/de/>. Accessed 28.3.2018

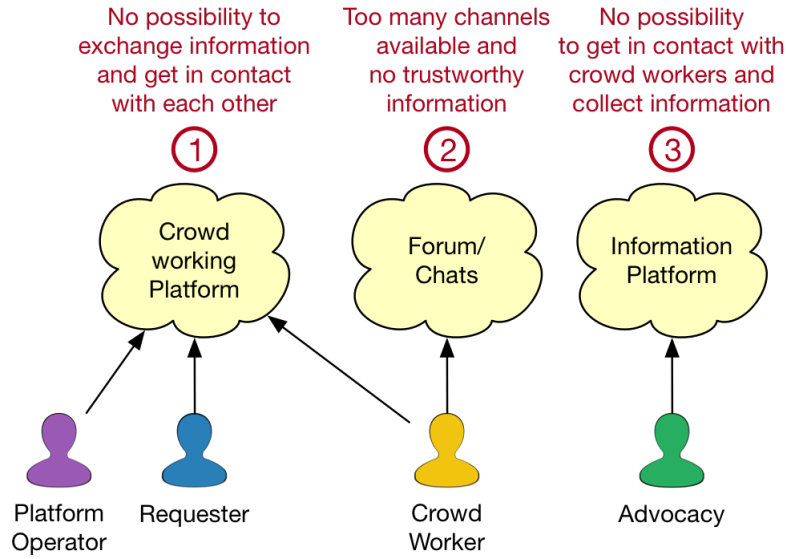


Figure 3.1: Overview over the research challenges

RQ 1: What are relevant requirements, capabilities and business processes of software systems that support crowd workers and worker organizations in knowledge sharing and self-organization?

The first research question aims to understand the needs and grievances of the crowd workers as well as to determine the business processes and system requirements that support a platform for crowd workers best. This understanding is important to design criteria and a system according to the needs of the stakeholders. The answer of this research question will be the foundation for our solution for the stakeholder.

RQ 2: What are major conceptual software architecture design principles of social-collaborative platforms for crowd workers?

Crowd workers share knowledge and self-organize themselves over different forums, platforms or social media. Therefore, we need to identify what elements and features fit the needs and business processes of crowd workers and worker organizations. Based on these results, we aim to formulate design principles and architectural models for software architects to plan and implement a central system. In addition, we plan to provide a design approach of a system where crowd workers can share their knowledge and self-organize themselves. The results elicited for RQ 1 will provide the foundation for the definition of a catalog of design principles, success criteria and quality attributes.

RQ 3: How do the identified principles support the design of collaborative crowd worker platforms?

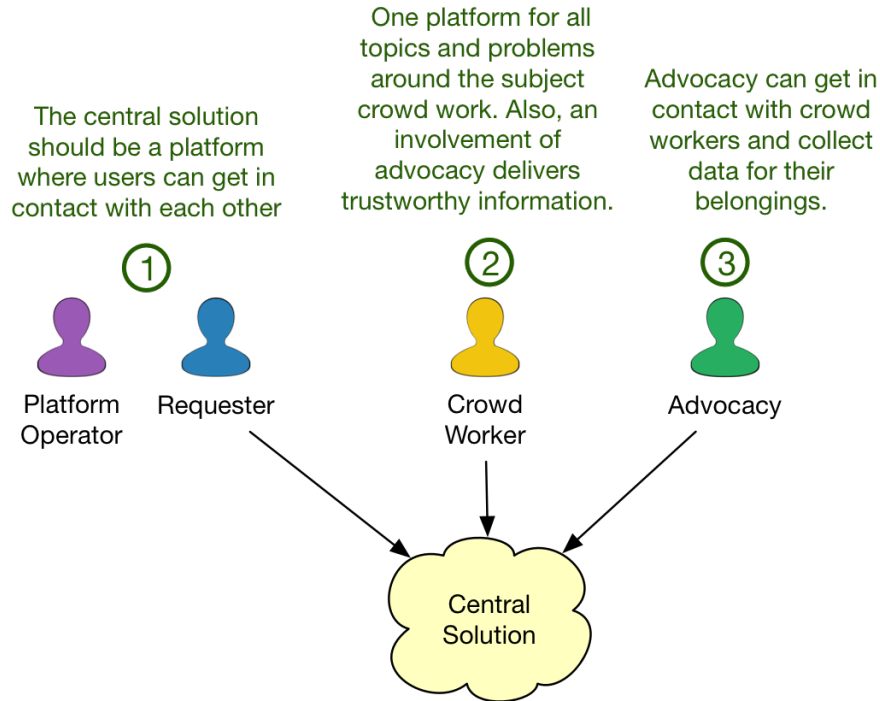


Figure 3.2: Overview over the research goal

The expected design principles and architecture models should fit the needs of crowd workers and worker organizations. Therefore, we need to investigate the validity of the provided contributions of this work. From the software architect's perspective we will evaluate how the identified principles support the architecture design process of a platform with regard to utility, applicability, and completeness. Secondly, from a user perspective we will evaluate how well the created system, and in particular the components that are related to the design principles, address the quality attributes and key performance indicators associated with the needs and business processes of key stakeholder groups.

3.3 Research Methodology

This thesis aims to contribute (1) a collection of system requirements and business processes in the domain of crowd work, (2) design principles and architecture models for the crowd work domain and (3) a prototypical implementation of the solution approach as well as an evaluation of the prototype. The goal is that this platform counteracts with the negative effects caused by the crowdsourcing platforms. The methodological approach is mapped in Figure 3.3 and designed according to Wieringa [52], which provides us guidance to design a methodology. This Figure illustrates the steps we plan to take to answer the research questions.

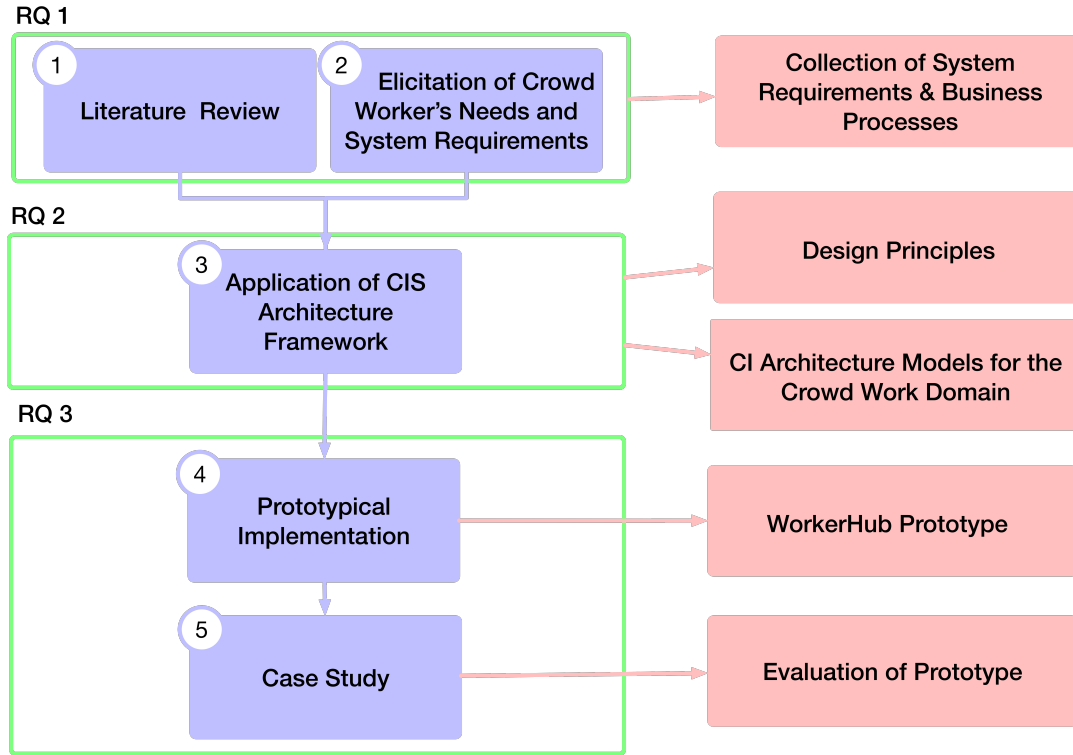


Figure 3.3: Steps of the planned methodological approach

1. Literature Review

To answer RQ 1 we do a literature review and then elicit the needs of crowd workers and system requirements. The result will be a collection of system requirements and business processes, which reflect, crowd workers' needs. The first step is to perform a review of existing literature to get an overview about the topic of collective intelligence system architectures which, we assume, provides the ideal system requirements to support crowd workers, and on the other hand to get deeper into the world of crowd work. That implies research about the use of different types of computer-supported cooperative work (CSCW)-like systems. The goal is to get an insight view of all relevant fields like working situation of crowd workers or collective intelligence which is the foundation of the following steps.

2. Elicitation of Crowd Workers' Needs and System Requirements

In this step, the needs of crowd workers will be identified through a literature review. In addition, input will be collected from worker unions and subject matter experts of crowd work from their websites and published studies. A core aspect of this step is to collect information about what system capabilities are required for certain stakeholder groups regarding the identified needs. The focus is on the identification of a comprehensive set of relevant stakeholders to support, their concerns and required processes based on the literature review. The consolidated

results from step 1 and 2 aim to answer RQ 1. On the basis of these findings, a set of system requirements will be determined.

3. Application of the CIS Architecture Framework

For the second research question, we do an application of CIS architecture Framework. Based on step 1 and 2, key design principles for developing a system which fulfills the needs of our stakeholders will be derived. Based on these design principles for the domain of crowd work, we will define architecture models of a central system solution for crowd workers by applying the CIS architecture framework [43]. The description of these architecture models is compliant to the well-established ISO/IEC/IEEE 42010:2011 standard [53] on software architecture descriptions. Being standard compliant is an important criteria as this enables integration alongside other more rigorous architecture approaches (like architecture frameworks, domain-specific languages or model-driven engineering methods) in the future. A comparative analysis of other existing platforms will also be delivered in this step. The platforms will be compared with our design principles to see limitations and potentials of them.

4. Prototypical Implementation

We will answer RQ 3 by applying our design principles and CI architecture models in the form of a prototypical system that realizes the main principles, criteria and quality attributes identified in step 2. The prototypical software platform will be implemented using an agile software development process and state-of-the-art web development frameworks (e.g., Ruby on Rails) and cloud-based SaaS infrastructure (e.g., Heroku).

5. Case Study

The prototype will be empirically evaluated in a quantitative case study in order to identify completeness, utility and usefulness in addition to specific quality attributes in order to answer RQ 3. Therefore, we do a case study where we invite our stakeholder to test our prototype and do a survey which designed according Runson et al.[54]. From this quantitative user survey we conclude whether our principles and model meet the needs of our stakeholders. In addition, we take up improvement suggestion and refine the principle catalog. The planned target groups consist of domain-relevant stakeholders like crowd workers and worker organizations, as well as subject matter experts from the field of crowd work. The study procedure, quality attributes, measures and analysis methods will be reported. Also, descriptive statistics will be used and provide information about the fit of our design principles and the architecture design. Finally, we will evaluate improvements of our design principles and models.

Stakeholders, their Concerns and Needs

The stakeholders are crucial in this thesis, because they can evaluate the appropriateness, usefulness and completeness of our study. To develop a software platform, which fits the goal of establishing design principles and architectural models, we need to take a closer look on the key stakeholders in the domain of crowd work, their concerns and needs. Crowd workers, platform operators and requesters, regulators, interest groups like work unions are the relevant target groups. In figure 4.1 we give an overview of the stakeholders and their relationships. The requester offers a microtask, which a crowd worker can accomplish and delivers the fulfilled task back. The platform operator has full access to all information on the platform and monitors every action. The regulators care about the crowd workers and are interested in more information about their working life and needs to support them. Interest groups focus on more transparency during the whole process and of the platform operators but also care about the working life of a crowd worker.

In the following we discuss each key stakeholder, her concerns and needs in detail.

4.1 Crowd Workers

First we take a closer look on the main stakeholder - *the crowd worker*. Crowd worker are people who work on "Share Economy" platform (e.g.: Upwork or Amazon Mechanical Turk) where digital short time jobs are offered [30]. Anyone can work on such platform, therefore most of the workers do it as a sideline [18]. The concerns and requirements of this stakeholder group are our main points of decision for our design principles and CI architecture models. Their needs in order to fulfill their tasks and providing them support in knowledge sharing and self-organization have highest priority.

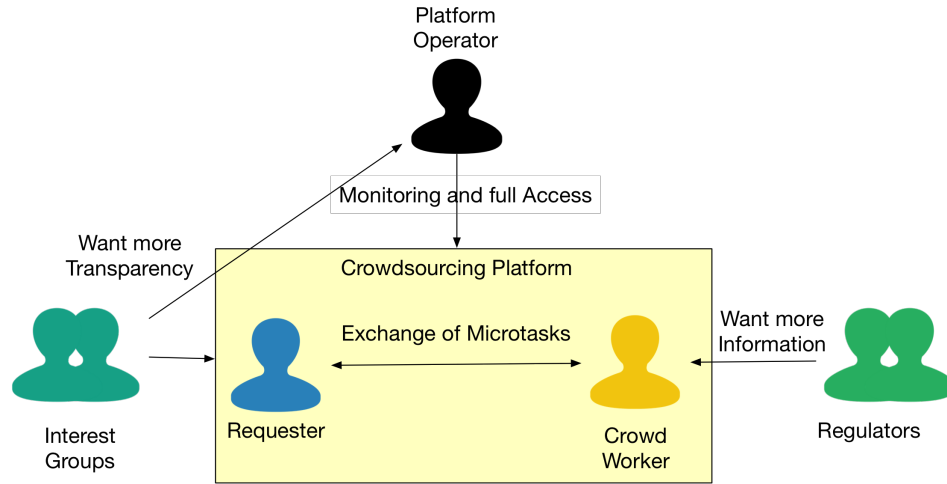


Figure 4.1: Overview of the key stakeholders and their relationships in the domain of crowd work

The review of existing literature and studies showed that the workers currently have bad working conditions and the crowd work platforms and requesters do not treat them very well. A few research studies took a closer look on the interaction between the workers, because by definition the crowd is an independent collection of people [1]. There are workers who participate in active realizations for improvement of their working conditions [16]. Some crowd workers only are looking for a forum where they can share information with each other without censorship or condescension and worker visibility and dignity in general [16]. Others are interested in a long-term work relationship with productive requesters and also in a general improvement of their relationship to the requesters [16]. In addition, some of them asked about work unions while others want the opposite [16].

Workers are collaborating with each other regularly and social interaction is a basic human need [28]. There are various reasons why they are collaborating to fulfill their social and technical needs [28]:

- One reason is to manage the administrative effort which occurs by doing crowd work [28] [15]. Such effort could be signing up for an account, avoiding scams and receiving checks [28]. If someone decides to do crowd work than she has to give the platform sensitive financial details [28]. In this case it is very helpful to have a friend who is familiar with the platform. So she can be sure that the platform is not a scam or something else [28]. Social endorsement through word of mouth and presence of a friend help them to feel safe [28].
- Another reason is that workers are communicating by phone, forums, chats, social media and sometimes even in person to share information about new tasks and requesters [28] which is shown by figure 4.2 where the different communication

channels of the crowd workers are illustrated. They even call each other when there is a good task. Also, recommendations from a friend are trustworthy [28] [15]. This exchange with friends causes a certain level of confidence by the worker, which cannot be made by a technical solution [28].

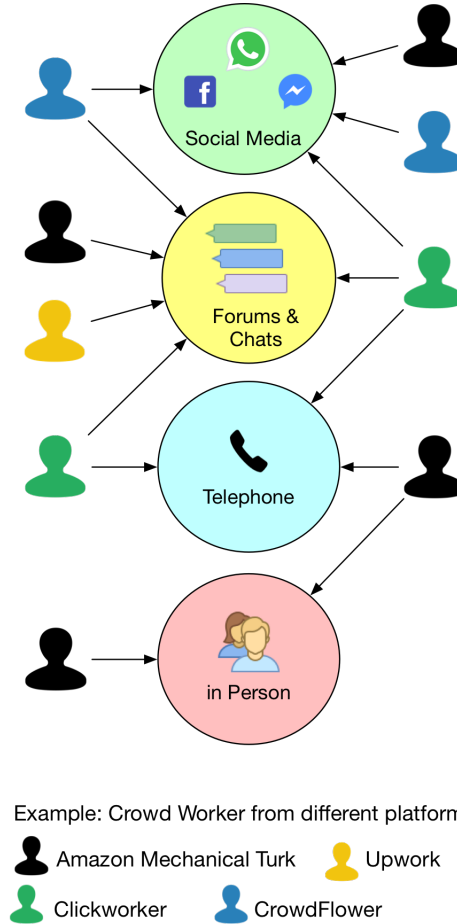


Figure 4.2: Communication channels between crowd workers

- Workers are interested in comparison of each other because they want to gain more knowledge and information [39]. Furthermore, they want to know which earning potential the others have reached and how they can also gain a better potential [39].
- Finally, the crowd workers are collaborating to do the job and to support each other [28] [15] [26]. Also, workers try to motivate each other through night shifts [28]. It is easy to ask a friend to understand the description of a task than to ask the requester, because answering is not cost-effective for microtasks [28] [15]. The community is very important for new workers to get informed about all rules,

norms and averages or how to get good tasks and which plug-ins and add-ons are needed and what should be done in case of rejection [15].

The worker is not an independent and alone human who only performs the task and wants to be paid and does not want any social interaction [28]. The crowd is a network of dynamic, self-organized humans [28]. The workers long for a fair and objective communication not only with each other but also with the requesters [39]. All this information is exchanged by using forums and social environments [28] like *TurkerNation* <http://turkernation.com/>. Accessed 26.3.2018 or Facebook Groups, and there is already a trend from traditional forums to a working environment for crowd workers [26]. These forums and environments are working like office break rooms, where workers emphasize, communicate and confide with other workers [28]. Figure 4.3 shows the social area of *TurkerNation* <http://turkernation.com/>. Accessed 26.3.2018 where users can talk about personal topics of interest like TV shows or games. Crowd working systems cannot

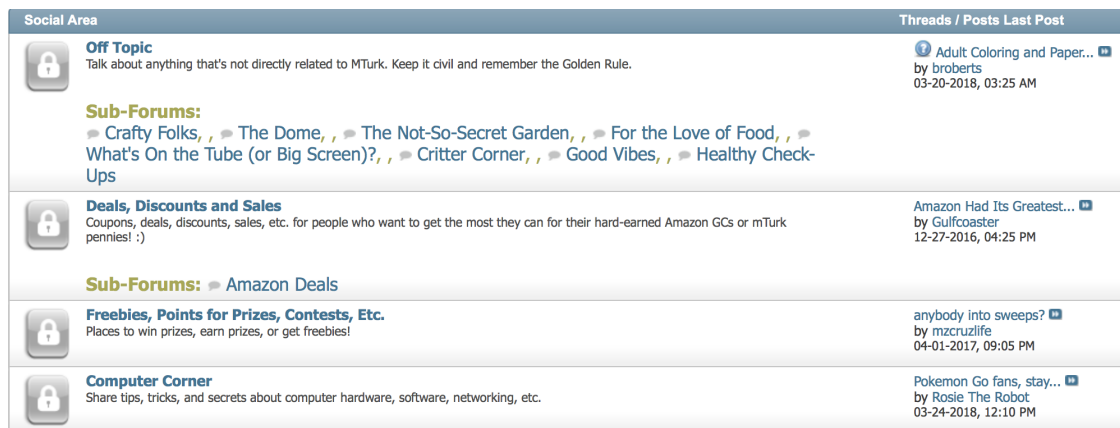


Figure 4.3: Snippet of the social area on *TurkerNation*

eliminate certain needs and feelings of workers like the simple human need for social interaction, endorsement, recognition and feedback of their work [28].

4.1.1 Stakeholder Concerns

AMT describes crowd work as „Human-as-a-Service“ and that is how the workers are treated [3]. It is like the cloud and brain performance, a service, which can be rented [3]. Humans are treated as machines [3] and this is the main critique - the dehumanization [16]. In 2015, a group of turkers wrote a letter to Amazon’s CEO Jeff Bezos [15]. They reminded him that humans are made out of flesh and blood and wanted to be treated with respect and fairness [15].

There is a trend in the community that workers will rather communicate with others when they originate from the same country or work on the same kind of task [24]. For example, Amazon Mechanical Turk is not providing a forum where their workers could talk about their tasks, and so the workers evade to other platforms [24] [39]. One of these platforms is Reddit HWTF¹ and it is used to share links and information about good tasks [24]. More different topics will be discussed in three other forums: MTurkGrind², TurkerNation³, and MTurkForum⁴ [24]. Each of them has moderated areas and tens of thousand discussion threads, which touch wide ranges of themes [24] (e.g.: requesters, good tasks, etc). Each forum member can contribute input to any thread. By using the chat function or a private message system it is possible to directly communicate with other members [24]. CloudMeBaby⁵ is another forum helping with navigation and improvement of crowd work platforms. In addition, there are also numerous public and private Facebook groups [24].

The most popular platforms to be used by crowd workers from Amazon Mechanical Turk are TurkerNation and Facebook [24]. Though, a third of those surveyed said that they are active at least on more than one platform [24]. A drawback of this approach is that this decentralization makes it even harder to localize and obtain information of reputation [55]. But forums or other social environments are important for the workers, because individual workers have not many options to build solidarity and they offer them a chance of creating sufficiently coordinated actions to make pressure on AMT [16]. The reasons why it is not possible to form any action against these systems are divided loyalties, time pressures to earn enough money, and the risk that agitation poses to their reputations and to availability of crowd work more generally [15].

In summary, we can say that

- crowd workers would like to share information with other workers.
- they are interested in connecting with each other to lead them through the platform and their idiosyncrasies.
- they long for others to ease their work.
- workers are looking for a social environment where they can exchange information without censorship and with more visibility and higher status.
- most of them like the idea of worker unions.
- reputation is an important value for them.

¹<https://www.reddit.com/r/HITsWorthTurkingFor/>. Accessed 28.3.2018

²<http://www.mturkgrind.com>. Accessed 28.3.2018

³<http://turkernation.com>. Accessed 28.3.2018

⁴<http://mturkforum.com/index.php>. Accessed 28.3.2018

⁵<http://www.cloudmebaby.com/forums/portal.php>. Accessed 28.3.2018

Based on a study from Al-Ani and Stumpp [38] with a focus on the support of crowd workers, this list is extended with the following suggestions:

- An assessment system for evaluation of crowdsourcing platforms,
- a platform-independent reputation system,
- the certification of work algorithm of platforms through objective instances and
- possibility of development of a platform-independent community.

4.1.2 System Requirements

According to the described concerns and needs, crowd workers are looking for a safe and social-focused platform where they can exchange information towards various topics related to crowd work. They would like to have a reliable source and a reputation system available on one location where all workers can interact and which is an organized environment addressing all topics. Most of them wish a corporation with worker organizations to get trustworthy information [15]. This will increase the working conditions of the crowd workers because they get more knowledge about their rights.

4.2 Platform Operators and Requesters

To define the concerns of them according to this work is a challenge, because they are the only ones who have full-access to data and the control over the rules of the platform [3]. Therefore, we assume that they didn't want a change because they have the most power. The platform operators are responsible for the 24/7 operation and running of the crowd work platform. The purpose of the platform is to mediate between offer and demand [3], which is illustrated in figure 4.4. The platform operator has access to all data that are collected by the platform. The requesters and crowd workers communicate over usually different interfaces with the platform. This business model of three laterality makes it possible that the operators can outsource the business, legal and social risks and the costs for labor and means of production to the other two parties [3].

Requesters offer a task over a platform and pay the crowd workers for their results [1]. Some advantages for the requesters are

- the access to extensive knowledge and expertise,
- acquisition of more innovative concepts,
- higher processing time of tasks due to split into microtasks,
- a potential to save costs because of low rewards,
- high flexibility,

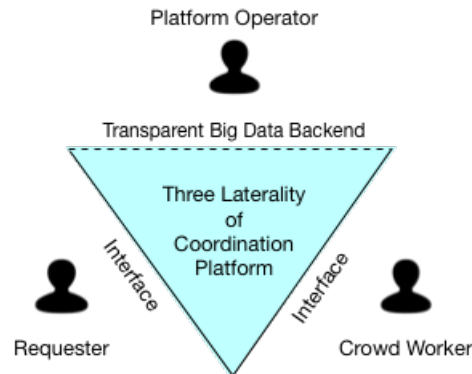


Figure 4.4: Model of the three laterality of a crowd work platform [3]

- focus lies on the core competences
- and it could increase market competence because the worker could be potential customers and thus they are involved in the innovative development process [33].

The drawbacks on the other side for the requester could be that

- the tasks need a precise and detailed description,
- it is difficult to calculate costs for crowdsourcing initiatives,
- danger of resistance of the worker,
- risk of losing control over the activities of the crowd worker
- and risk of outflow of internal know-how [33].

4.2.1 Stakeholder Concerns

There is a great problem with regard to the relation of the requester and the crowd workers. Leimeister et al. [18] evaluated the relation to the requester and asked workers about three categories: specification, payment and reviews. The ratings in category microtasking were worse but overall all workers rate it appropriate. Furthermore, they asked about the support and mentoring of the platform, because this is the „central

workplace“ for the workers and therefore it has crucial significance. The evaluation criteria were the terms of use, further training, qualification and support when there is a problem, dimension of care and well-being. Surprisingly, microtasking got quite good ratings in all categories except for qualification, because workers do not need any specific qualification. Overall, the reviews of a requester given by workers are subjective and noisy [39].

Concerns from the side of requesters arise through the flexibilization: The casualization and infiltration of social and labor standards [3]. It is a step back into early industrial period regarding protection of workers and social protection systems [3]. The decentralization should support high-quality work, but it undercuts behavior and institutions, which are critical for high-quality work [26]. In many traditional organizations, central worker co-operations are a key-element that guarantees quality of work inclusively skill development [56], knowledge management [57] and performance rating [58]. The relief of the worker is not happening nor higher self-determination, because the reward is very low like a „digital sweatshop“, the competition is very high and the work process is unilateral [33]. So, the working quality decreases because of the flexibilization which is crucial concern on requesters side.

4.2.2 System Requirements

If there is a good relation to the requesters, the workers would more like to work for them and this would also have an impact on the work they make [39]. A good example is the forum *TurkerNation* <http://turkernation.com/>. Accessed 26.3.2018 where the requesters can contribute to a sub-forum where they can interact with workers. Requesters can usefully discuss problems of their tasks or work together with the workers to understand problems better [39]. So this is beneficial for both sides to invest in a good relationship, because then the quality of the outcomes will increase. [39].

If the working conditions would get better on the platforms or the workers can use a platform where they can share their knowledge, then the quality of the work will increase and this will result in better outcomes [37]. Furthermore, if there is a platform where all workers can exchange their experiences about crowd work platforms and requesters then more qualified workers are available to perform the requesters' tasks. Maybe it could also facilitate collaborations between operators and requesters to collect ideas how to be more attractive to qualified workers.

There are also some other platforms that improved the working conditions [26]. These platforms are *Duolingo*⁶ *LevelUp*⁷, which integrated interactive tutorials to enhance the quality of work [59].

For our central solution it will be helpful to have a place where requesters can interact with the crowd workers and discuss about tasks and problems. Also, a section where

⁶<https://en.duolingo.com>. Accessed 28.3.2018

⁷<https://www.thelevelup.com>. Accessed 28.3.2018

the crowd workers will get more information about other platforms with better working conditions could be provided. This will not only be helpful the crowd worker but also to for the requester to share their tasks on more social crowd working platforms. Maybe it will push the platform operator to change bad working conditions on his platform.

4.3 Regulators

One of the most important regulators in the german-speaking area are the Austrian Chamber of Labor and the IG Metall. These regulators gained the importance of better working conditions for workers and have also triggered various actions to help them, e.g., the website faircrowd.work⁸. The development of the platform faircrowd.work reflects their goals and current actions. The platform serves to inform crowd workers about certain crowd working platforms, rights of the crowd worker, unions for crowd workers and basics about crowd work. So, their goal is to inform and help as many crowd workers as possible to improve the working conditions of the crowd workers.

4.3.1 Stakeholder Concerns

The Austrian Chamber of Labour realized a study [30] about the crowd working scene in Austria. This study is a collaboration with the University Hertfordshire, Ipsos MORI, Foundation for European Progressive Studies (FEPS) and UNI-Europa. They performed an online survey and asked 2003 adults to their working situation. The data from the sample are weighted towards age, sex, region and working status for reflection towards the adult population of Austria. In their survey they asked the sample about their income, kind of employment and their work, how often do they work on crowdsourcing platforms and their customers. A few insights of this study include:

- 36% of the interviewed persons tried to get a job in the crowd working scene, but only the half found one.
- For 59% crowd working is an additional income.
- Only 3% earned more than 60.000 Euros per year and 48% earned less than 18.000 Euros per year.
- Office work, little tasks and “clickwork” is the most popular activity with 74% followed by creative tasks and different IT tasks with 62%.
- The age distribution is more or less evenly distributed.

The IG Metall developed with eight crowd work platforms a code of conduct [27]. Figure 4.5 shows an overview of the principles of this code of conduct. From this overview four groups of the principles of the code of conduct can be derived. The blue elements aim

⁸<http://faircrowd.work/de/>. Accessed 28.3.2018

an open and fair communication. The green elements are about the task itself. The elements in turquoise stand for concessions of and for the crowd worker. Finally, the red elements cover the legal situation. The goal of the code of conduct is to have universal guidelines additionally to the law to establish a basis for trustful and fair togetherness between platform operators and crowd workers. In this code of conduct all eight platforms accepted to comply and promote these principles. The code of conduct shows clearly the concerns of the regulators, which are:

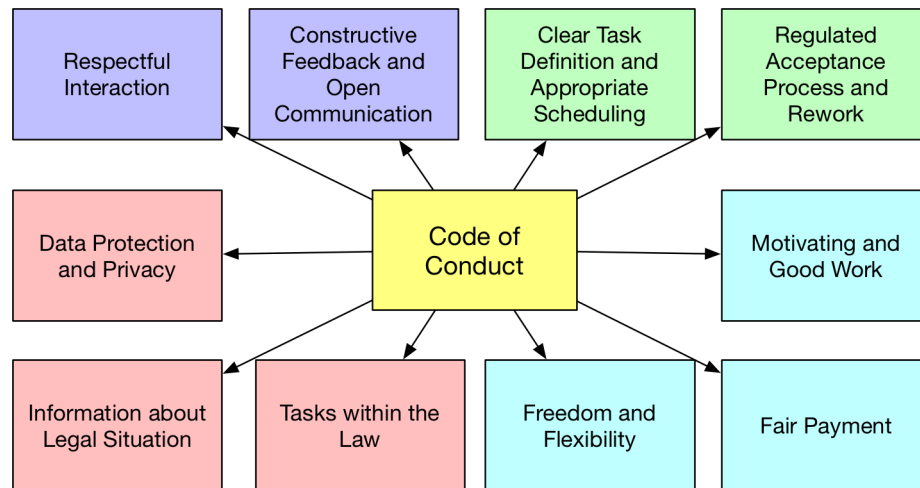


Figure 4.5: Overview of the principles of code of conduct from IG Metall

1. **Tasks within the law.** The platform operators do not offer projects, which are against the law and they have the obligation to review the projects and tasks according to legal compliance.
2. **Information about legal situation.** Crowd workers will be informed about their activity in context of tax regulation and guidelines.
3. **Fair payment.** The calculation of the payment depends on complexity of the task, necessary qualifications, the local nature of its business, local wage standards and expenditure of time. The terms of payment, especially time limits and cycle payments, has to be revealed.
4. **Motivating and good work.** The signatories help to ensure a user-friendly and intuitive interface for the workers. Also, there should be a contact opportunity if someone needs help.
5. **Respectful interaction.** The relationships between the platforms, requesters and crowd workers are based on reliability, trust, honesty, openness and mutual respect. The platform operators have the responsibility to esteem these values.

6. **Clear task definition and appropriate scheduling.** Tasks have a clear and precise definition. The crowd workers get a detailed description about the textual and temporal criteria to fulfill the project.
7. **Freedom and flexibility.** The crowd worker decides by herself if she wants to take a job and for non-acceptance, without being afraid of any negative consequences. The worker is not tied to the platform.
8. **Constructive feedback and open communication.** The operator and requester commit to give the worker support and prompt feedback.
9. **Regulated acceptance process and rework.** The acceptance process for the finished tasks have to be in written form and transparent for the worker. A non-acceptance has to be explained and based on facts of the project description. A possibility for rework should be given unless the project cannot allow it.
10. **Data protection and privacy.** Data protection and privacy have the highest priority, especially personal data or data which is strictly confidential.

4.3.2 System Requirements

The regulators already deal with the drawbacks of crowd work and would like to improve the situation of the workers. These samples show that there are concerns on the side of regulators and they are striving for improvement. According to the concerns, the regulators need a platform where all information over the situation of the crowd workers is summarized. These information should be analyzable and trustfully. Furthermore, for this stakeholder it is important to get in contact with the crowd workers to fulfill their goal in helping and informing crowd workers. A place for the regulators in a central solution to exchange information will be required.

4.4 Interest Groups

Relevant interest groups include worker unions and governments. The unions do not know how to deal best with the crowd workers [38] [36]. Most of the crowd workers want unions and only a few reject them [16]. Also, the government noticed the poor working conditions of crowd workers and so the European parliament has conducted a study about the social protection of workers in the platform economy [29]. They interviewed over 50 expert stakeholders in the platform economy across eight EU countries and this report includes an original survey of 1.200 crowd workers. The results were that they identified a need to help crowd workers to increase their working conditions. Also, they understand the importance of this topic as well as need for action.

4.4.1 Stakeholder Concerns

The European parliament noticed the grievances and in their study formulated actions against these circumstances. The established recommendations give a good insight of their concerns [29]:

- Legal classifications:
 - Address the framework of legal classifications to redress exclusions from social patterns. The goal is to establish a framework which meets all particular legal definitions.
 - Working on framework of actions to ensure the inclusion of all those who need social protection.
 - Ensuring that enforcement of legal employment status is conducted by an independent authority.
 - Having a legal binding way to make the needs and desires of the crowd worker heard by platform operators.
- Transparency:
 - Proving by the employer the reasons why a person is not an employee by carrying out a paid work.
 - Providing full and easy accessible information to all workers.
 - Getting more information from the platforms to support policy development around social protection. This should be ensured with measures.
 - Access for the unions to inform all workers about their policies and to potentially organize them.
 - Developing clear and transparent review mechanisms, which not should be used for rating the work quality.
- Establishing instruments to protect crowd workers:
 - Extending the right to written statement of particulars.
 - Introducing independent mediation panels for settlement of platform-worker grievances.
 - Knowing the customers and the purpose of work.
 - Able to contest non-payment work evaluations and qualification test outcomes.
 - Including unions but only if the workers wish it, and they also should get the opportunity to choose them by themselves.

4.4.2 System Requirements

Based on the concerns of the interest groups and their goals and their possibilities to support crowd workers, they would like to have a tool where they can easily access information from the workers and where the working conditions and policy environments in the respective country are traced. That is also one main problem as defined in the study of the European Parliament [29]:

"The first problem is definitional. Different definitions lead to different conclusions about size, and about the most important issues for policy-makers to address. The second research problem is who to access suitable data for making an accurate assessment of the size and growth of the platform economy." Also, they wish more transparency in the whole crowd working process, which includes more information from the platform operator and a solution where platform operators and crowd workers can communicate with each other.

Furthermore, the introduction of a worker organization for the crowd workers should be first point of contact.

CHAPTER 5

Design Principles

In this chapter we will answer one part of RQ 2. Based on the results of the literature research and some studies, we identified needs of crowd workers and key stakeholders and how we could support crowd workers in knowledge sharing and self-organization. So, we derived seven design principles relevant for a technical solution approach in the crowd work domain. In the following each principle is introduced by providing an overview of all principle characteristics followed by an explanation of the principle and suggestions for technical realizations. The technical realization part involves general implementation suggestions and references to implementation details at our prototype. We will take a closer look at other platforms that crowd workers use to share information and evaluate them according to our defined design principles to identify the potentials of the proposed principles of the crowd work domain. All seven principles together support the workers and the worker organizations where each principle is important to implement in order to achieve knowledge sharing and self-organization. The goal of the principles is to deliver a basis for a crowd work platform-independent community.

5.1 Independence and Centralization

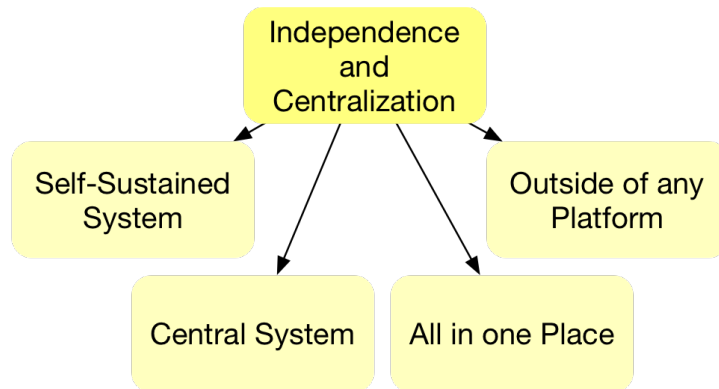


Figure 5.1: Overview of all characteristics of the design principle *Independence and Centralization*

Figure 5.1 shows the four characteristics, which are central to this principle. A central solution should be self-sustained and external of any crowd work platform. This is important, since the terms of use of each crowd work platform are contrary to law [35] [3]. The terms of use are an issue on crowd work platforms not only because of their volume of text, which is in the case of AirBnB about 550.000 words (approximately the length of a novel), but also regarding the unilateral elaboration in favor of the interests of the platform operator and their area of life [3]. In general, the terms of use are opaque and from a legal point of view it is questionable [18]. The platforms abused software license agreements as employment contracts and in their terms of use it is governed that a worker is a self-employed person or „independent contractor“. Therefore the platform operator is not in charge of social security of the workers [3]. Also, in the terms of use it is not mentioned what is needed to be a crowd worker on their platform [28].

As mentioned in chapter 4, workers fear about their job if they say something, which is not allowed according to the terms of use [3]. Some terms of use prohibit the right to organize, which is a universal human right [35]. This reason is not only an argument for *Independence* but also for the *Privacy* principle, which is described later on. So, terms of use rule the personal and business relationship between the workers and the requesters [3] [35]. Therefore, crowd workers need a system, which is separated from their working platforms in order to have a place where they can talk about their work or organize themselves without any restrictions. Also, Al-Ani and Stumpp [38] defined in their paper suggestions like a possibility of development of a platform-independent community on how to support crowd workers and the building of a platform independent community.

Centralization is upon *Privacy* probably the most important principle, because you cannot gain collective power if information and people are spread all over different forums and private Facebook groups. But what is the problem with decentralization and private

groups? One problem is the double content in forums. There are many forums and also many private groups where the content is overlapping and locked. This is especially the case in forums of big platforms like Amazon Mechanical Turk. The number of forums is increasing with the popularity of a crowd work platform. Crowd workers who work on platforms with smaller reputation have to face the same issues as workers on better-known platforms, but they do not get in touch with the content provided in the separated forums.

It is important to have as many workers as possible centralized at one place to get as much information about their working life and to gain huge collective power to change their working conditions. Another advantage would be that with all of this information, workers get a good overview about other platforms and could shift to platforms with better working conditions. In chapter 4 platforms with better working conditions like Duolingo or the code of conduct, which is signed by eight crowd work platforms were mentioned. Amazon Mechanical Turk is a giant and most of the other platforms and especially new platforms with good conditions are unknown or not famous enough. Another problem with the new platforms is that people are always afraid of scam [28]. A system with collected knowledge about vast amount of platforms would solve the problem.

Technical Realization

The technical requirements for a new independent system have to deal with huge amounts of information. The architecture should support a good structure of information and should be expandable.

5.2 Privacy

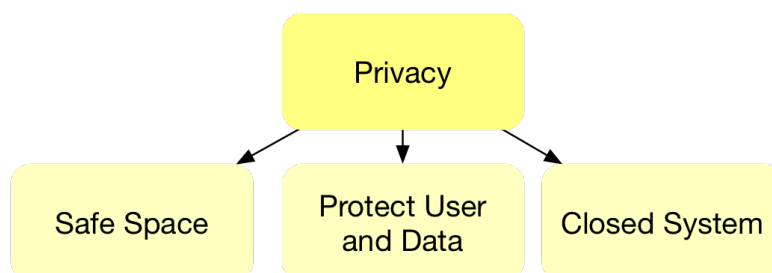


Figure 5.2: Overview of all characteristics of the design principle *Privacy*

In Figure 5.2 we see three main key characteristics for the principle *Privacy*: Safe Space, Protect User and Data and Closed System. These characteristics are now explained as follows: A central solution should provide a safe space for sharing, aggregating, and disseminating of information. Crowd workers' identities need to be protected to prevent their identification. As mentioned in the first principle, workers weigh collective action against personal risk to their account, work environment, or reputation as online workers

[15]. Especially for turkers, AMT is a crucial source of income and therefore, they actively prevent any action that may put it at risk [15]. They worry about losing their source of employment, about taking actions that attracted legal attention to crowd work or worrying about being singled out and getting blacklisted by Amazon or other employers [15]. So, they are afraid of be banned from the platform if they organize themselves or do something against the terms of use which are in the most cases legally questionable [15] [35]. It is very easy for AMT to ban them, because they have a very high turnover on hiring and firing [15]. Workers cannot do anything against it. There are no face-to-face interactions, no way to communicate with others, no organizational body to make or enforce decisions or complain about decisions [15].

On one side it is part of the system to protect the data and on the other side it is part of the politics to help crowd workers and protect them. To do so, the second part needs more information from the workers in order to take action. To obtain information about working conditions and working life of crowd workers, it is important to have a system where they can share their experiences without fear. This will be further discussed in principle *Analyzability and Transparency*.

Technical Realization

For the technical realization it is important to decide if it is an open or closed system. Is it necessary to have a registration and login or not? Because of the privacy and the third principle *Trust and Reputation*, the answer to this question is yes. The system should be a safe space and thus it should be closed and there should only be like-minded people.

The next question deals with the kind of data that are important in such a system. The first action to protect the privacy is that every user gets a unique username independently from their real names [60]. Workers are more likely to interact with other workers from the same country or do the same tasks [24]. So it is important to collect information about the city where workers live and the country they are from. This information gives the system the ability to suggest workers the right forums or topics. Email address for registration and password for login is essential in the system. More required information is not needed from the worker. The advantage of less private data is that it is hard to find out the worker who is behind the username and, in case of hacking attacks against the system, they will not be able to get much information [60]. Sensitive data is of course the email address and password [60]. A recommendation to protect these sensitive data is to encrypt them [60] and to use an email address, which is not used on the platform.

5.3 Trust and Reputation

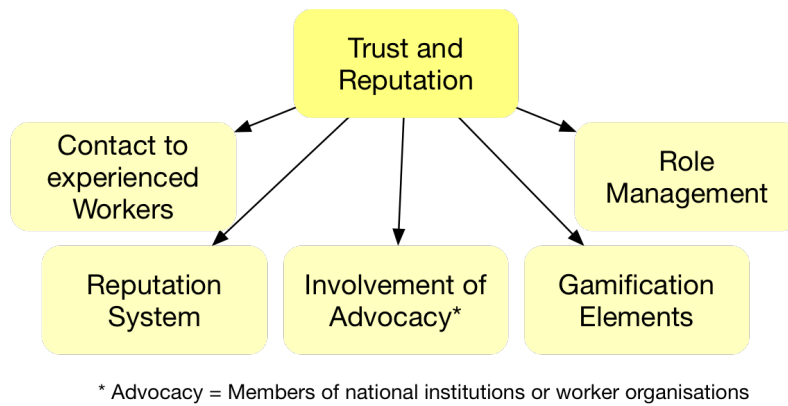


Figure 5.3: Overview of all characteristics of the design principle *Trust and Reputation*

The five main characteristics for the principle *Trust and Reputation* are shown in this Figure 5.3. Each characteristic will now specified. Trust and reputation enable workers to engage in the worker community. This is the key to a thriving and active community. Salehi et al. [15] recognized this relevance on their platform Dynamo. They worked a long time to gain the trust of the crowd workers. When they achieved the trust of the community, it was possible to make pursuing changes [15]. Crowd workers have to trust that admins or other instances act neutrally, fairly and predictable [15]. To meet these requirements we collaborate with stakeholders who are part of national institutions or worker organizations and have an interest to connect with crowd workers and help them.

Trust is also important when workers get information from experienced workers. Gray et al. [28] found out that workers consistently described to rely on finding someone who is willing and able to support new workers in the disorienting world of survey questions or culturally specific knowledge like "twerking" [28]. For Indians, it has more importance because of the dizzying range of appliances and other consumer goods, which are relatively unknown or uncommon to them [28]. But it is necessary to understand them to complete the most mundane tasks posted at the platforms [28].

Reputation is a key for success. Many studies and papers defined reputation as key for a good and sustainable communication with crowd workers [26] [15] [9] [38]. Reputation strengthens the feeling to be in good hands [28]. It is important to have people in the community with a good reputation like experienced crowd workers or members from national institution or worker organizations. Schmidt [3] recommended in his study to establish a platform where institutions can warn workers from platforms who are especially problematic in labor law. Building up a community to support them in knowledge sharing and self-organization, we see it as important to have an official institution as part of the community.

Technical Realization

To have a role management in the system is a necessary feature, which also supports trust and reputation [61]. Especially, if there is a certain role for official institutions and experienced workers. Giving them a special part in the system makes the system itself reliable and trustworthy. Thus, these parts can take care of the quality of the content.

In addition, a reputation system is recommended. It can be realized with gamification tools like levels, points, role of a member, badges or special signs. These awards would strengthen the trust within the community and be an intrinsic motivation for the members to take part and contribute. The success of a reputation system is proved by the study of Hamari et al. [62] where they found out that gamification provides a positive effect but depends on the context.

Another part of such a reputation system could be the involvement of crowd work platform and requester ratings. A rating system for platforms is implemented on the faircrowd.work¹ website from IG Metall and a rating system of requesters is realized on Turkopticon.

5.4 Fairness and Objective Exchange

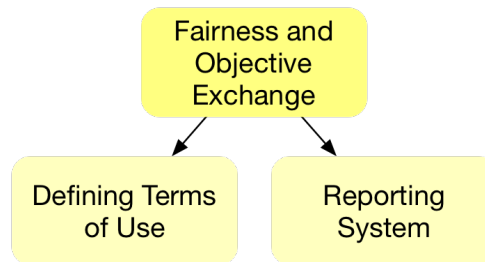


Figure 5.4: Overview of all characteristics of the design principle *Fairness and Objective Exchange*

The characteristics of *Fairness and Objective Exchange* are Defining Terms of Use and Reporting System as shown in Figure 5.4. The identified needs and ways of workers to interact with each other and exchange information revealed the importance of a fair and objective exchange with the requesters and the crowd work platform [39] [28], but also among each other [39]. It is not only important that exchange takes place, but also how it takes place and in which manner. A respectful and kind social interaction where every opinion is taken seriously is fundamental for the system. The desire to act fairly is also an important characteristic of a forum [39]. To act fair on a platform means no hate speeches or inappropriate language. Also, it is important that every user is taken for serious and will be treated in a respectful manner.

¹<http://faircrowd.work/de/> Accessed 04.04.2018

Technical Realization

A system should provide mechanisms to guarantee a fair and objective exchange. Defining terms of use, like Bonabeau [48] recommended, that are virtually signed by registration on the system could be a first mechanism to address fairness. Every member of the system is bound to the terms and has to observe them and, in case of not fulfilling it, is allowed to report the violation. Such a reporting system would assist to enforce the compliance of the terms of use. It is necessary to think about the reasons of reporting, consequences of wrong reporting, consequences of many reports and administration of reporting.

5.5 Distribution of Information and Awareness

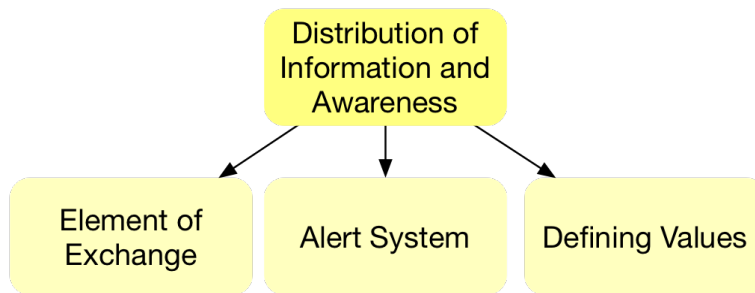


Figure 5.5: Overview of all characteristics of the design principle *Distribution of Information and Awareness*

Figure 5.5 shows three characteristics of the principle *Distribution of Information and Awareness*. These characteristics are specified as follows: As mentioned in the principle *Fairness and Objective Exchange*, workers collaborate and interact with each other

- to manage the administrative effort [28] [15].
- to share information about new task and requesters by using telephone, forums, chats, social media and even in personal [28] [15].
- to gain more knowledge and better quality of their work [39].
- to support and motivate each other through the working process [28] [15] [26].

To have social contact is a normal basic human need [28] and so it is plausible that a wide range of forums are sprout. It is quite common to talk about microtasks, because the requesters' replies do not come immediately and asking a friend is much easier [28]. It reduces the costs of spending time finding tasks and reliable requesters when workers collaborate [28]. The interviews in the paper of Gray et al. [28] showed that collaboration happened routinely among workers. The workers could only get in touch with each other

if a system offers a possibility to exchange information, like forums do. Gray et al. [28] found out that workers alert each other when a good task is online. In their study they also analyzed how workers find their tasks. So they created a microtask on AMT. 36% of the traffic came from search on the platform and 41.3% came from online forums [28]. Another interesting insight of their study was that most of the major spikes correspond to online posts by searching in the forums for mentions of their microtasks. This finding also refers to the *Trust and Reputation* principle, because workers convey trust to the information from forums [28]. Therefore, the system should inform other workers about new tasks.

A main point is to create awareness. Gaining more awareness of a worldwide problem among every worker, not only a few. Especially in industrial countries, where the wage level is very high. We think that establishing values on a system would raise attention to the importance of collaborating and supporting each other. To explicitly reference the goal of the system, we should motivate workers to cooperate and form action against the grievances of crowd work. So, we recommend to form values additionally to terms of use.

Technical Realization

The system should contain an element where workers could exchange and distribute their information. For example, it is possible to give them a forum, private chats, FAQ or to offer a mentoring system. Many new workers are looking for somebody they can turn to when they have problems and need some more support in the beginning [28]. Also, a key feature should be to regularly inform users of a new reliable tasks and relevant information. How should the users get informed? What data are needed for an alert? Therefore, it could be practicable to install a mechanism where a user will get notified if a new task is only. This information could contain all important information about the task like the origin platform, time and compensation.

5.6 Analyzability and Transparency

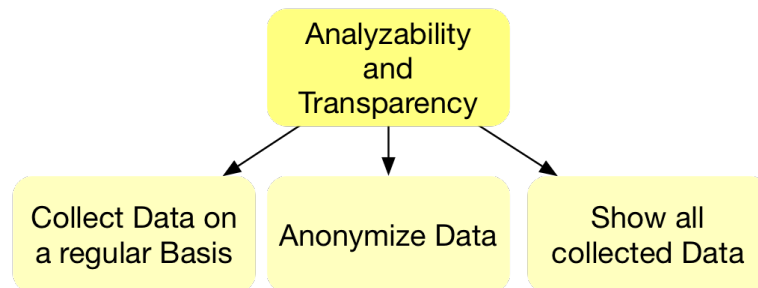


Figure 5.6: Overview of all characteristics of the design principle *Analyzability and Transparency*

Three main characteristics are crucial for the principle *Analyzability and Transparency*. These are Collect Data on a regular Basis, Anonymize Data and Show all collected Data. On the one hand this principle is based on suggestions about an independent institution who certifies algorithms of platforms [38] and the introduction of a rating organization for platforms [3]. On the other hand this principle is the result of the suggestion and recommendations published by different worker organizations and official institutions like European Parliament [29], IG Metall [27], Chamber of Labour in Austria [30], etc.

Transparency is crucial for a change in the political system. Therefore, we want to quote the Frankfurt Paper on platform-based work [35]:

"Municipal, regional, national, and international policy makers have very limited access to data describing the number and value of transactions conducted over online labor platforms, the geographical locations and demographics of clients and workers, or the importance of platform-based work to the business strategies of firms and the livelihood strategies of workers. In short, the knowledge base required to make sound policy is missing."

According to this fact they proposed that platform operators, researchers, workers, worker organizations, policy makers and other actors as appropriate work together to increase transparency in the world of platform-based work [35]. They also suggested to develop an infrastructure for ongoing data- and information-sharing to provide policy makers information for stable political decisions and to guarantee observance of legal regulations [35]. Because of the lack of concrete laws for the good of the crowd workers, IG Metall [27] defined a code of conduct to complement legislation as mentioned in chapter 4.

Analyzability arises from the limitation of data [35] and the problem of accessing suitable data [29] that means that worker organizations have no access to regular data over the working conditions of crowd worker. To help national institutions or worker organizations to generate data and content for political decision makers, it is important to collect enough reliable data and transparency and gather them.

Technical Realization

Anonymizing data is crucial in this principle, especially regarding to the privacy of a user as described before for the design principle *Privacy*. According to the Frankfurt Paper [35] the collection of demographic data are of particular importance. But how detailed should this data be collected regarding to *Privacy*? Data from the working life of a crowd worker are also interesting and should be collected. All these data should be collected on a regular basis and it should be presented transparently. Not only national institutions or worker organizations should have access to the data, but also the crowd workers.

5.7 Work Type-Specific Software Clients

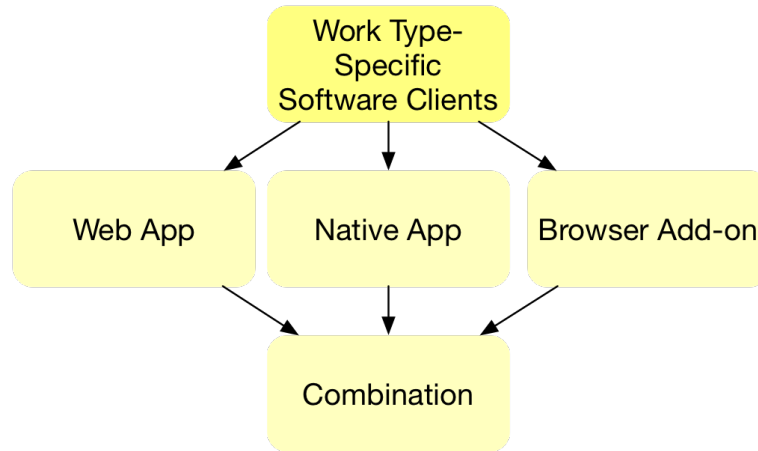


Figure 5.7: Overview of all characteristics of the design principle *Work Type-Specific Software Clients*

In Figure 5.7 we can see the different characteristics of this principles. These characteristics are explained as follows: When planning a platform for crowd workers, the first thing to consider is what group of workers should be the target group of the system. Crowd work is defined as every task, which is offered to crowd over the internet [2]. Crowd work is location independent in the narrow sense. Delivery service or leasing is location dependent are summarized as gig work [3].

After choosing the target group of workers, the second thing to consider is the preferred communication within the group of workers. As mentioned before regarding the principle *Distribution of Information and Awareness* and in chapter 4, crowd workers like to communicate in forums, chats and social media. Especially, gig workers tend to use chats like WhatsApp or Facebook to coordinate and communicate with each other, because navigation to the destination has to be fast. A web-app or website is therefore not sufficient. On the other hand, crowd workers work on their laptops at home. For them it is easier to use a website or browser add-on. The difference is that gig workers primarily use their phone when working and crowd workers their laptops.

Technical Realization

A web app is a good design choice to cover all needs. It is accessible by a mobile phone and a laptop. An additional native mobile app would fit better the needs of gig workers. Depending on the system, it could sometimes be more useful to have a browser add-on where a crowd worker can have fast access to the system. A good example for this option is Turkopticon where they have a website with all ratings of a requester. In addition, they offer an add-on where a user can hover over a requester's name on the Amazon Mechanical Turk platform and the ratings of the requester appear.

5.8 Analysis of Platforms Used for Crowd Workers' Self-Organization

In this section an analysis of three different platforms according to the defined design principles is presented that are regularly used by crowd workers to interact and organize themselves. These platforms include TurkerNation², Reddit HitsWorthTurkingFor³, and CloudMeBaby⁴.

TurkerNation

TurkerNation is an online forum for crowd workers who work on Amazon Mechanical Turk. On this platform they exchange general information about AMT, requesters and they also have a social area to discuss topics that have nothing to do with crowd work.

TurkerNation is a typical online forum. Workers have to register to get access to all functions and it is a self-sustained system that is not dependent on any crowd work platform. The platform is restricted to Amazon Mechanical Turk workers. Therefore, it fits the independence part of the principle *Independence and Centralization*, but not the centralization aspect, because it does not involve workers from other crowd work platforms.

The platform is a safe space for the users and a closed system. A user has to register to get access to all information on the website and has to choose a username and password. To unlock the account it is necessary to indicate an e-mail address. TurkerNation protects the users and their data, because it does not ask for sensitive data except the password, which is obligated when a platform is a closed system. So, the forum fulfills the *Privacy* principle.

On TurkerNation a certain role management system is implemented. It offers roles for crowd workers, moderators, researchers, requesters, and journalists. Each role has different permissions. There is no specific role for the advocacy, but they can participate with the role of a researcher. Every user on TurkerNation can view statistics on her profile. The statistics shows posts, messages, activities and thanks of a user. These statistics allow a better insight into the activities of other workers. The forum suits mostly to the principle *Trust and Reputation*. The only thing that is not integrated is gamification elements.

With regard to the *Fairness and Objective Exchange* principle it is important to have defined terms of use and a reporting system which is shown in Figure 5.8) that TurkerNation offers such system.

²<http://turkernation.com> Accessed 28.5.2018

³<https://www.reddit.com/r/HITsWorthTurkingFor/> Accessed 28.5.2018

⁴<http://www.cloudmebaby.com/forums/portal.php> Accessed 6.6.2018

5. DESIGN PRINCIPLES

Report Post
Forum: [Newbie Questions](#)

Your Message

Message:

Note: This is ONLY to be used to report spam, advertising messages, posts that break the rules, problematic (harassment, fighting, or rude) posts, and posts which need to be fixed (for example, changing an icon on a HoF/S post, a post in the wrong thread, or a thread in the wrong forum.) Please do not abuse the report function!

Figure 5.8: Report a user on TurkerNation

TurkerNation is a whole exchange element. A exchange element is an element where users can exchange information for example over a chat or forum. It is one big forum (see Figure 5.9) with different threads and topics (e.g.: threads over different requesters or new tasks, but also about discussions like tv-shows). They also take care of their users and so they defined a TurkerNation Mission Statement. There is no explicit mechanism where workers get notified when a lucrative task is online. So, an alert system is missing. Hence, *Distribution of Information and Awareness* is partly fulfilled.

General Turkling		Threads / Posts Last Post
	mTurk General General chit chat about mTurk topics. Sub-Forums: mTurk Site Problems , Suspension Discussion , Master Worker Discussion	Would you take a HIT that... by Turker297 03-24-2018, 10:18 AM
	mTurk Questions Questions about mTurk specifically, like "How do I do a search for a requester's HITs?" This is not for general mTurk chat. Sub-Forums: Etiquette, Rule Clarifications & Safety Questions , Make Money Fast/More Efficiently Questions , Payment Questions , Tax & Other Government-Related Questions , Qualification Questions , Newbie Questions	Any way to auto-catch the... by Flower Today, 04:25 AM
	Requester Introductions This is the place for requesters to introduce themselves to the Turker community! Please do not post announcements of new HITs here (unless it's your first time posting), further announcements should be posted in Everyone Else. If you don't have access to EE, please PM Spamgirl and she'll set you up.	Private
	Turker Introductions This is the place for Turkers to introduce themselves to the forum!	Private
	Crowd Work in the Media News articles, videos, interviews, etc. belong here! Academic articles can go here also.	Quick update: another... by spamgirl Yesterday, 08:28 PM
	Crowd Work Statistics and Other Data Papers on mTurk, statistics, and more.	"We Regret to Inform You"... by spamgirl 02-28-2018, 04:09 PM

Figure 5.9: Forum structure of TurkerNation

The principle of *Analyzability and Transparency* is hardly fulfilled on TurkerNation. It

is possible to do a poll (see Figure 5.10), but there is no menu item to get an overview about polls or surveys, and the collected data is not on a regular basis. They perform polls when a topic is urgent. The collected data are anonymized, but there is no option for advocacy to download the data.

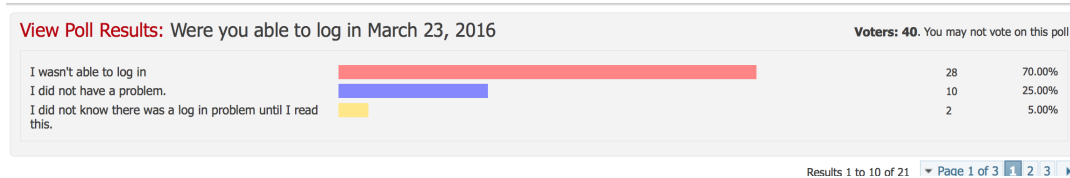


Figure 5.10: A poll on TurkerNation

TurkerNation is a forum system on a website. When we categorize it based on the different types of the principle *Work Type-Specific Software Clients*, we would assign it close to a web app, because you can access it from the computer or mobile phone and it is not static.

Reddit HitsWorthTurkingFor

Reddit HitsWorthTurkingFor (Reddit HWTF) is a subreddit from the platform Reddit. Reddit is a user-powered social news site [63]. A subreddit is a reddit for certain content [63]. This subreddit addresses only Amazon Mechanical Turk workers where they can post lucrative tasks. Regarding the *Independence and Centralization* principle, it is a self-sustained system and it is outside of any crowd work platform, but it is not a central system because it is only for turkers and doesn't offer any other feature than sharing and discussing about new tasks.

The *Privacy* of the users is protected because usually Reddit HWTF only asks for a username, password and e-mail address.

Trust and Reputation is valued on Reddits HWTF, because they offer different roles: user, moderator and admin. Trust arises with the "karma" of users. It is also possible to vote on a post. Involvement of advocacy is possible, but the member of an advocacy will not be certain highlighted. An experienced worker is identifiable by the points of karma, which is illustrated in Figure 5.11.

If you want to participate on Reddit HWTF, then you have to agree on three different agreements: user agreement, privacy policy and content policy. These agreements are formulated in such ways that a user is protected and should behave in a trustful and respectful way. Thus, the principle *Fairness and Objective Exchange* is fulfilled and it is also possible to report a user.

This platform does not offer a specific forum structure. There are only posts with content regarding to Amazon Mechanical Turk. The main focus on this site is to be informed about new, lucrative tasks. So, they offer an element of exchange, which is notification

system and this element has functions of an alert system. The part *Awareness* of the principle *Distribution of Information and Awareness* is not fulfilled here, because they do not explicitly mention values or a mission.

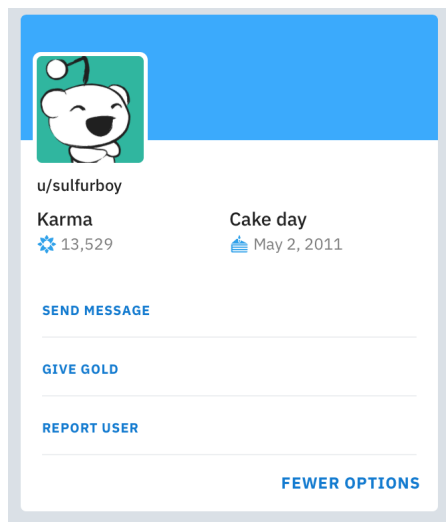


Figure 5.11: Profile of a user of Reddit HitsWorhtTurkingFor

Reddit HWTF does not offer any possibility to collect data or to conduct a survey.

This platform offers a web app and a native app. Therefore, it fits the needs of a crowd worker and a gig worker. A benefit is definitely the new interface and the usability of Reddit HWTF.

CloudMeBaby

CloudMeBaby is a platform with the intention to help crowd workers in navigating and improving the cloud based working space. This platform is one big forum and it is divided into different sections and their sub forums addressing different crowd work platforms. It is a central and self-sustained system, which is external of any crowd work platform. Therefore, the principle *Independence and Centralization* is fulfilled.

CloudMeBaby is a safe space because only registered users can view all content and the user data consist only of username, password, e-mail address, language, time zone and sex. Thereby, it is a closed system. So, CloudMeBaby matches the *Privacy* principle.

This platform offers a mix of a role management and reputation system. There are three main roles: user, moderator and administrator. The users have different levels, which are shown by a bar out of symbols under their name. Another element of reputation are the statistics under the user's name which is shown in Figure 5.12.

The only thing that does not match the *Trust and Reputation* principle is that there is no involvement of advocacy.

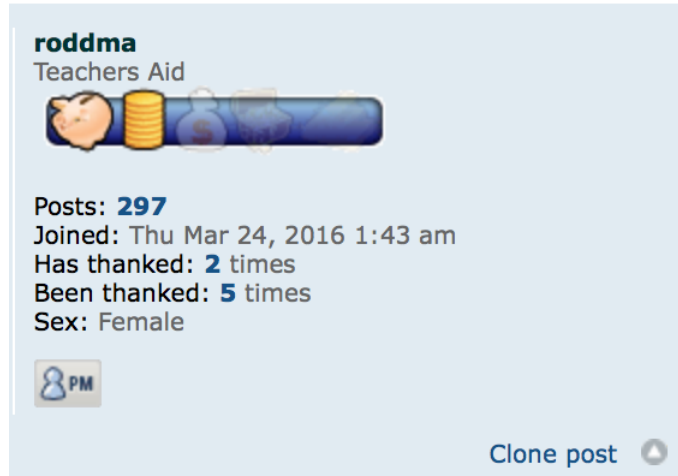


Figure 5.12: Profile of a user at CloudMeBaby

CloudMeBaby also has terms of use and privacy policy and it is possible to report a post and instance a reason for report. The terms of use and privacy policy ensure a fair and respectful exchange of information among the users. Also, there are some penalties formulated when a user misbehave. Therefore, the platform fulfills the principle *Fairness and Objective Exchange*.

This platform provides a structure of different forums. Therefore, it gives crowd workers the possibility to exchange and distribute information over different crowd working platforms. There are some forums where they talk about new and lucrative tasks, but the platform does not offer any system, which notifies a user. And the users do not get an overview of all tasks. For this reason, CloudMeBaby does not match 100% the *Distribution of Information and Awareness* principle. The platform also has defined rules, which could be put on the level with defined values.

On CloudMeBaby the principle *Analyzability and Transparency* is not addressed.

Regarding the *Work Type-Specific Software Clients* principle, CloudMeBaby is a web app, which is accessible from the mobile phone and the computer.

Comparison according to the design principles

		TurkerNation	Reddits HWTF	CloudMeBaby
Independence and Centralization	Self-Sustained System	✓	✓	✓
	Outside of any Platform	✓	✓	✓
	Central System	✗	✗	✓
	All in one Place	✗	✗	✓
Privacy	Safe Space	✓	✓	✓
	Protect User and Data	✓	✓	✓
	Closed System	✓	✓	✓
Trust and Reputation	Contact to experienced Workers	✓	✓	✓
	Involvement of Advocacy	✓	✗	✗
	Role Management	✓	✓	✓
	Reputation System	✓	✓	✓
	Gamification Elements	✗	✓	✓
Fairness and Objective Exchange	Defining Terms of Use	✓	✓	✓
	Reporting System	✓	✓	✓
Distribution of Information and Awareness	Element of Exchange	✓	✓	✓
	Alert System	✗	✓	✗
	Defining Values	✓	✗	✓
Analyzability and Transparency	Collect Data on a regular Basis	✗	✗	✗
	Anonymize Data	✓	✗	✗
	Show all collected Data	✗	✗	✗
Work Type-Specific Software Client	Web App	✓	✓	✓
	Native App	✗	✓	✗
	Browser Add-on	✗	✗	✗

Table 5.1: Comparison of the analyzed platforms according to the design principles

Table 5.1 provides a comparison of the three analyzed platforms according to the defined design principles. TurkerNation and Reddits HWTF do not offer a central system or an alert system. Also, Reddits HWTF and CloudMeBaby do not have a role where an involvement of advocacy is possible. None of the three platforms fully implements the *Analyzability and Transparency* principle and has a system where they collect data on a regular basis and give access to them, which is important for the political change. Furthermore, only Reddits HWTF offer a native app which is more likely used by gig workers.

On the other hand a forum like TurkerNation, Reddits HWTF or CloudMeBaby covers most of the defined design principles if we not consider the clarity and usability of the platforms. This finding may be the reason why workers are not willing or see a necessity of another platform to operate because the forums are sufficient for them. The need of a new platform is instigated from the advocacy side. Nevertheless, crowd workers can benefit from a new platform that is centralized and where they can actively do something to change the whole system, e.g., with their participation in surveys.

5.9 Summary

All defined principles arise from needs of crowd workers and grievances of the crowd work platforms. They should support crowd workers in knowledge sharing and self-organization, and provide transparency and insights into the working life of a crowd worker for worker organizations.

- The principle *Independence and Centralization* suggests a central system, which is external of any crowd work platform and provides all information at one place.
- The *Privacy* principle takes care about the data of the user collected by the system and is aware about sensitive data. It should guarantee a safe space for workers.
- The third principle highlights the importance of *Trust and Reputation* and suggests a few realizations how to promote an active community.
- To have guaranteed *Fairness and Objective Exchange* in a system it is important to define terms of use and a reporting system.
- The principle *Distribution of Information and Awareness* recommends places where workers get in contact with each other and creates a higher awareness of collective actions.
- *Analyzability and Transparency* is important to collect data of the current situation and problems for decisions makers to make a change for crowd workers in the political system.
- *Work Type-Specific Software Clients* helps to think about the ideal technical realization to support crowd workers in the best and most efficient way.

Software Platform Architecture Design for the Crowd Work Domain

In this chapter we will discuss the second part to answer RQ 2: "*What are major conceptual software architecture design principles of social-collaborative platforms for crowd workers?*". In the first part we identified and collected the concerns, needs and requirements of crowd workers, worker organizations and other stakeholders. Based on these requirements for a central technical solution, we formulated seven design principles.

In the next step, we will start to describe an architecture design of a central technical platform for the crowd work domain. Therefore, we will use and apply the *architecture framework for collective intelligence systems (CIS-AF)* proposed by Musil et al. [43]. The use of this framework provides the crowd worker and advocacy several advantages. For example: the framework guarantees an active and indirect communication among all users, it is location independent and a user gets also notifications when she is not online. Therefore, the user will get notified when a new message arises and the user can access the solution from every where. This framework is ISO/IEC/IEEE 42010 standard compliant which is important because this enables being integrated along side other more rigorous architecture approaches in the future. So this chapter presents CI architectural models based on the application of the CIS-AF, which form together with the design principles the core of the conceptual software architecture design of a proposed central technical solution for the crowd work domain.

For building a new CIS solution, we have to consider three viewpoints based on the CIS-AF [43]. First, we deal with the *CI context viewpoint*, which examines the main stakeholders, the as-is workflow regarding how crowd workers currently get their information, and the to-be workflow regarding how we could improve the knowledge sharing and self-

organization of crowd workers with a CIS. With the *CI technical realization viewpoint* we will identify our main artifact, the activities that can be performed upon this artifact, and its in/out data flows. The *CI operation viewpoint* then explores how we get an initial database for kick starting the feedback loop and metrics to continuously analyze the system's behavior.

In the following, the described architecture models are presented based on the CIS-AF.

6.1 CI Context Viewpoint

In this section we identify the main stakeholders relevant for the proposed technical solution in the crowd work domain. Then, we examine the current workflow used by the stakeholders to identify limitations. Based on these limitations, we design a new workflow applying the stigmergic mechanism of a CIS, which enables an indirect communication among the users. The to-be workflow shows then how the limitations can be addressed by a CIS solution.

6.1.1 Context

Individual crowd workers are performing different kind of tasks offered by crowd work platforms. Since these platforms do not provide any social communication and interaction features to their workers, crowd workers use external platforms to share and gain knowledge about new tasks or their working conditions or organize themselves.

6.1.2 Stakeholders

Crowd workers

Individual humans who are performing tasks offered by a crowd work platform and would like to exchange information and knowledge among each other.

Their concerns include sharing information efficiently with other workers, connecting with each other to overcome problems of the platforms or tasks, easing their work and exchanging information without censorship, and gaining more trust within each other.

Advocacy

A group of persons who are members of national institutions or worker organizations like the Austrian chamber of labor or IG Metall.

They are interested in efficiently supporting crowd workers and help them to understand their rights, but also explain their problems and needs to politicians. Therefore, they need to get insights into the current situation and more information that is analyzable and trustworthy.

6.1.3 As-Is Workflow Model

The as-is workflow of the crowd workers presents the business process how crowd workers nowadays share information with each other. This workflow is illustrated in 6.1 and its activities are described in Table 6.1 in more detail. In this process, different problems and limitations occur. For example, crowd workers search for task and collect information. Then, they have to choose the medium they would like to use to exchange information. In this way, the valuable information is spread all over the world wide web.

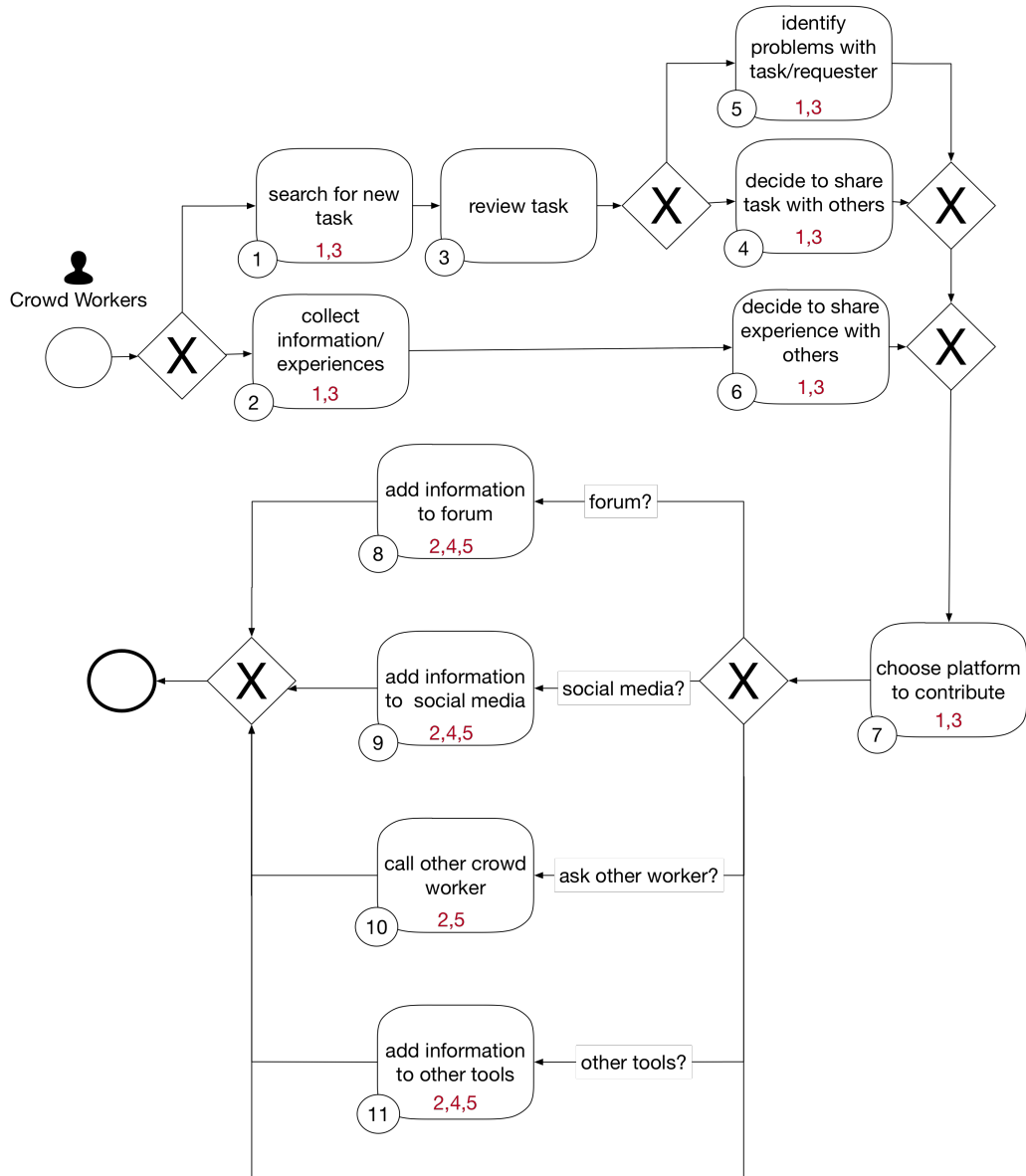


Figure 6.1: As-Is workflow of current communication between crowd workers.

The as-is workflow for the members of an advocacy is shown in Figure 6.2. Advocacy members want to inform crowd workers about their rights and collect data about their current working situation to help them in the long run. Also, they have the problem of the wide range of different platforms. This problem comes along with two challenges. The first challenge is to find an appropriate channel and the second challenge is to find one where they are highlighted as advocacy members. The last point is crucial for the mutual trust between crowd workers and advocacy.

An overview of the identified limitations and their impact on the stakeholders is given in Table 6.2. Every number shows one limitation and a limitation can occur in several activities. Also, an activity can have more than one limitation.

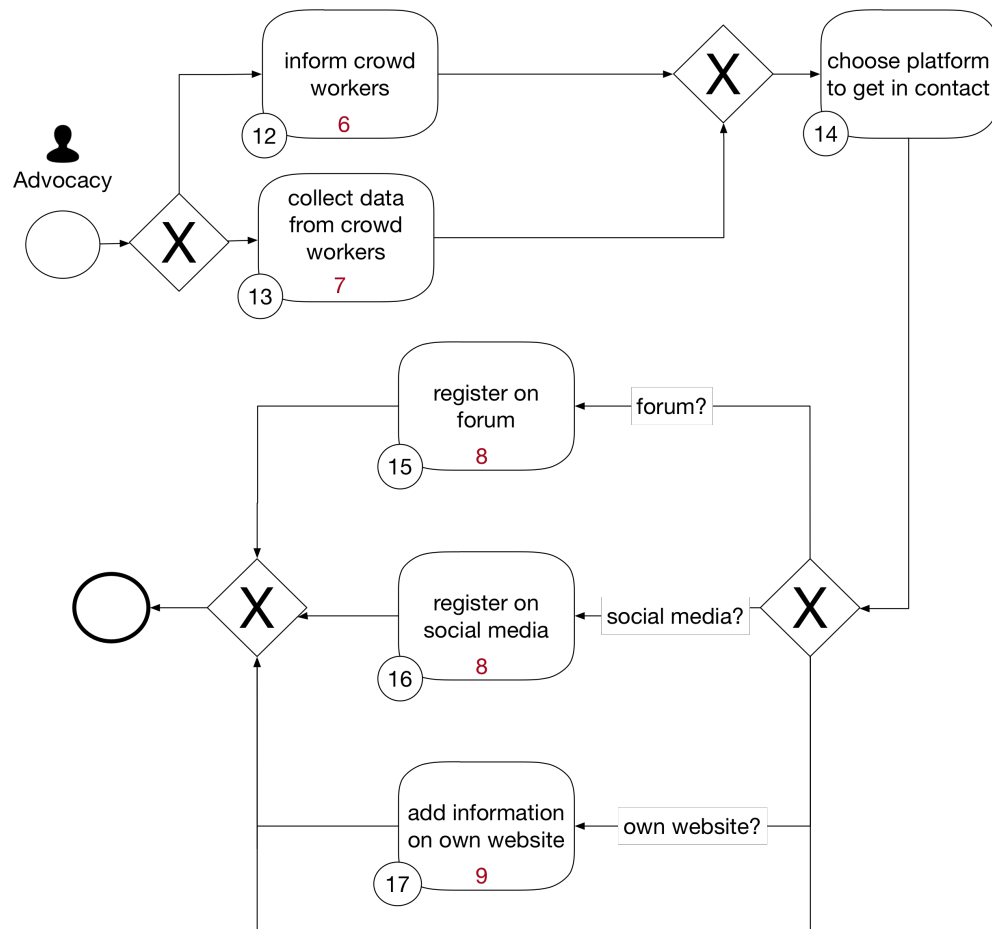


Figure 6.2: As-Is workflow of current communication between advocacy members.

List of Activities

A#	Activity Name	Activity Description	Stakeholders
1	Search for new task	To get lucrative tasks a crowd worker searches for new tasks on the crowd work platform.	crowd worker
2	Collect information/experiences	crowd worker collect different information (e.g.: about tasks, requesters, etc) and experiences over the time.	crowd worker
3	Review task	After choosing a new task, the task will be reviewed from the crowd worker	crowd worker
4	Decide to share the task with others	There is a need to share a task with other crowd workers	crowd worker
5	Identify problems with task/requester	The crowd worker has identified one or more problems with a task or requester	crowd worker
6	Decide to share experience with others	There is a need to share these information/experiences with other crowd workers	crowd worker
7	Choose platform to contribute	A crowd worker creates a new post with her concerns, questions, information, etc. on a platform	crowd worker
8	Add information to forum	A crowd worker post her concerns or information in a forum like TurkNation or CloudMeBaby.	crowd worker
9	Add information to social media	A crowd worker creates a new message consisting her concerns, questions, information, etc. on social media platforms (e.g., WhatsApp, Facebook).	crowd worker
10	Call other crowd workers	A worker calls another worker to get help with her concerns, questions, etc.	crowd worker
11	Add information to other tool	To get help or share information a crowd worker uses other tools, e.g., the plug-in of Turkopticon.	crowd worker
12	Inform crowd workers	Advocacy want to inform crowd workers about their rights	advocacy
13	Collect data from crowd workers	Members from an advocacy want to collect data about the current working life of crowd worker to help them.	advocacy

14	Choose platform to get in contact	Advocacy has to choose the right platform to get in contact with crowd workers	advocacy
15	Register on forum	On crowd working platforms, members from an advocacy has to register.	advocacy
16	Register on social media	Members from an advocacy have to register on a social media platform to get in contact with crowd workers.	advocacy
17	Add information on own website	Every advocacy has an own website where they share their information.	advocacy

Table 6.1: Overview of crowd worker activities in as-is workflow

List of Limitations

ID	A#	Description	Stakeholders	Impact (1-low, 5-high)
1	1,2,3,4,5,6	Since the provided information is scattered over several platforms and is organized in a decentralized way, crowd workers have a hard time to identify the communication channel that fit best their needs.	crowd worker	5
2	7,8,9,10	Knowledge is only provided to and stored in one certain system and is not available for all affected workers.	crowd worker	3
3	1,2,3,4,5,6	Different systems provide different kinds of activities to the crowd workers. For example, if a worker would like to share and discuss a lucrative task, then she has to deal with at least two different systems.	crowd worker	4
4	7,8,10	Crowd workers cannot be sure whether the source of information is reliable and trustworthy or not.	crowd worker	1
5	7,8,9,10	It is not possible to get in contact with national institutions or worker organizations to receive professional help or information necessary for do crowd work.	crowd worker, advocacy	1
6	12	Advocacy members don't know how to reach for crowd workers. They have all the information for them but don't know how to inform them.	advocacy	4

7	13	Getting data form the working situation of the crowd workers on a regular basis is not possible yet.	advocacy	5
8	15,16	To get in contact with crowd workers and to inform them is not easy, because advocacy members have to choose the appropriate channel and also has to be highlighted as trustworthy. Otherwise, it is quite long process to gain trust and inform them.	advocacy	4
9	17	Some organization has build their own website to share their information. The problem is that crowd workers do not search actively for their help and so they might not find their page or to go through all the content can be a huge effort. Also, on their platform they couldn't get in contact with them.	advocacy	3

Table 6.2: Limitations in the current as-is workflow

6.1.4 Stigmergic Coordination Model

The stigmergic coordination model shows how the indirect communication among the users as well as the perpetual feedback loop can be managed. figure 6.3 presents an overview of the stigmergic process of the new CIS solution including the domain item of interest, the possible interaction mechanisms with the domain item, and the feedback loop back to the users. The actors that interact with the platform comprise crowd workers as well as members of advocacy. Depending on the actor different kinds of interaction rules are possible. These interactions affect the domain items and create links between them resulting in a network of connected domain items. The domain item for the new CIS is *Groups*. The groups help to group same information together. For example: One group contains only general information about crowd work and another group consists information about certain platforms. Group links represent relationships between one or more groups, such as tags and categories. Tag represents a certain property. A group can have more than one property, but a group can only have one category. Categories help to group the same information together. The activities and data provided by the actors are collected and analyzed. Based on the resulting data analysis and the defined dissemination rules, triggers are created and disseminated back to the actor base to create stimuli for subsequent activities. Finally, the process starts again.

Figure 6.3 shows our stigmergic coordination model. Here, we see that crowd workers and advocacy has different interaction rules. While a member from an advocacy has access to all interaction rules a crowd worker can only perform a few. Some thing, which is important to explain is the difference between discussions and posts. A discussion

is part of a group. Everybody can create a discussion. If someone will participate at a discussion then the user will leave a post. So, a post is part of a discussion. After a interaction rule is performed, the domain item is effected. Then, the right dissemination rule grabs, which depends on the interaction rule and the stimulus is send back to the actors.

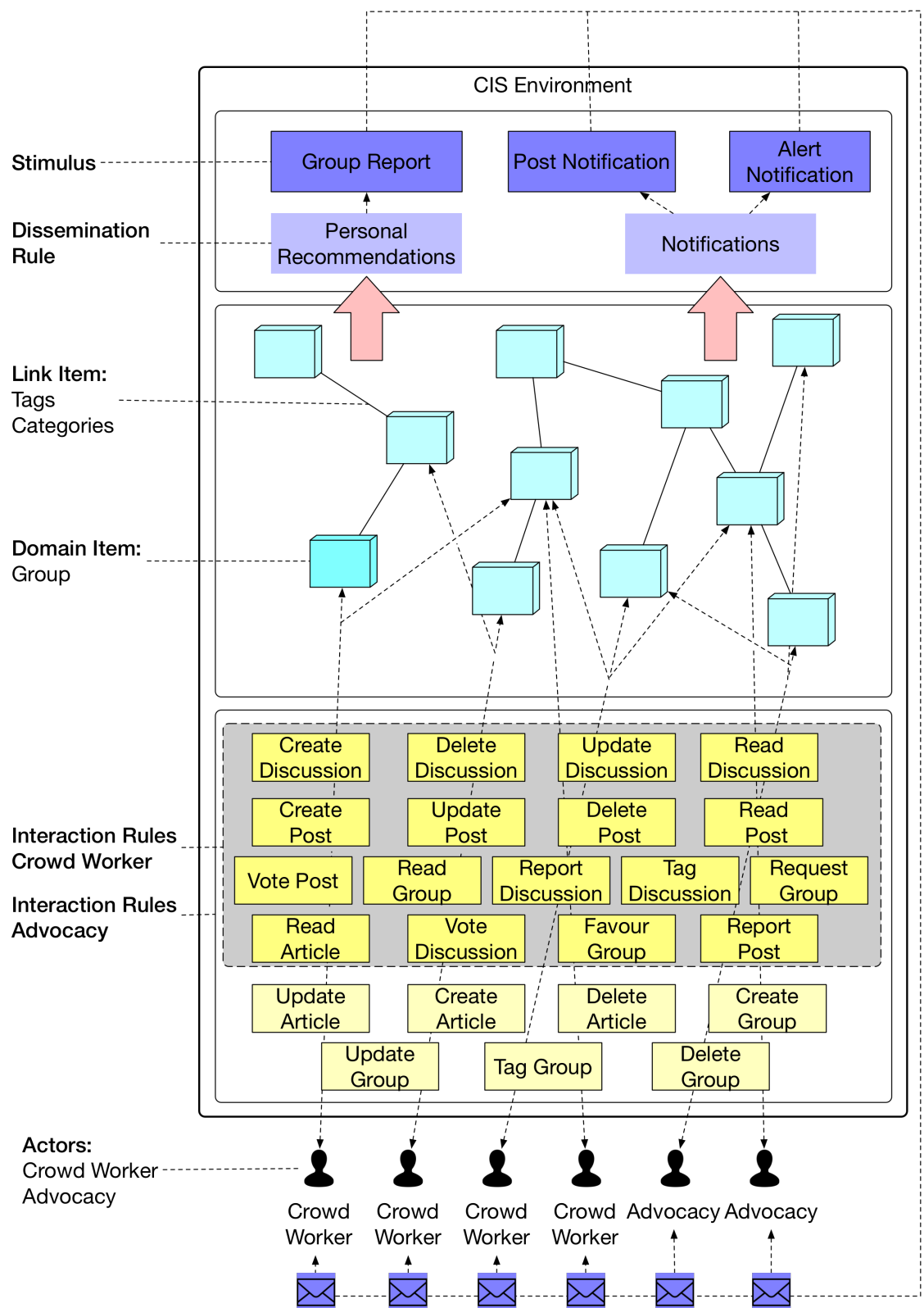


Figure 6.3: Stigmeric coordination model

6.1.5 To-Be Workflow Model

The to-be workflow model shows the planned business process involving a new CIS solution. In figure ?? we show the adapted business process for crowd workers. The big difference compared to the as-is workflow is that the users access a single central platform where they can share all their information and interact with each other. The information is centralized on one platform, which represents a central point of contact. The green numbers in the activities refer to the improvements, which are provided by the CIS and address the issues or limitations described in Table 6.2. Table 6.3 gives an overview of the improvements and addressed limitations according to activity number. Also, in table 6.3 we indicate concrete roles as agents, which will be in chapter 6.2.2 discussed in detail. The roles in our system are user, moderator, admin and external. The roles have different permissions. A member form an advocacy always get the role of an external. The other roles are for the crowd worker.

Figure 6.5 pictures the to-be workflow. A crowd worker search for a new task or collects information/experiences. If she has a found a new tasks, then she reviews the task. Afterwards, she found a problem with the task/requester or she wants to the share the new tasks. Also, when she has collected some information then she had to decide if she wants to share it or not. If she shares the task/information/problem then she has to choose the right platform and therefore, she has nothing to choose because there is only one central platform.

Figure 6.4 shows the to-be workflow for the advocacy. An advocacy member wants to inform crowd workers or collect data from them. Therefore, they have to choose the right platform to get in contact. Thus, there is only one central platform where they can get in contact.

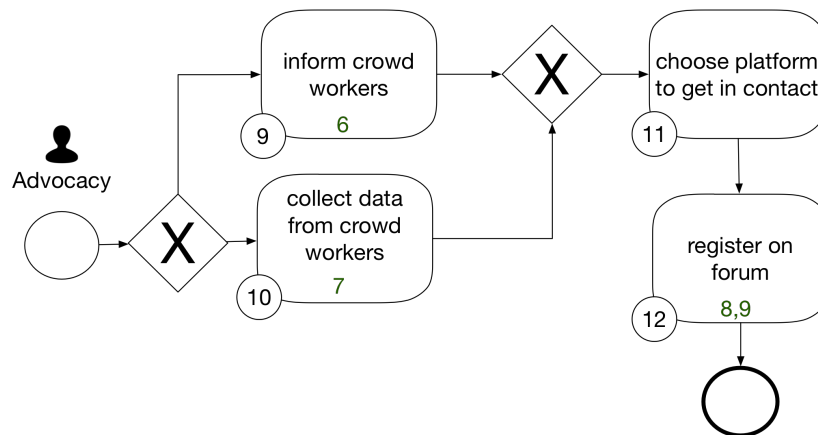


Figure 6.4: Activities of advocacy in to-be workflow

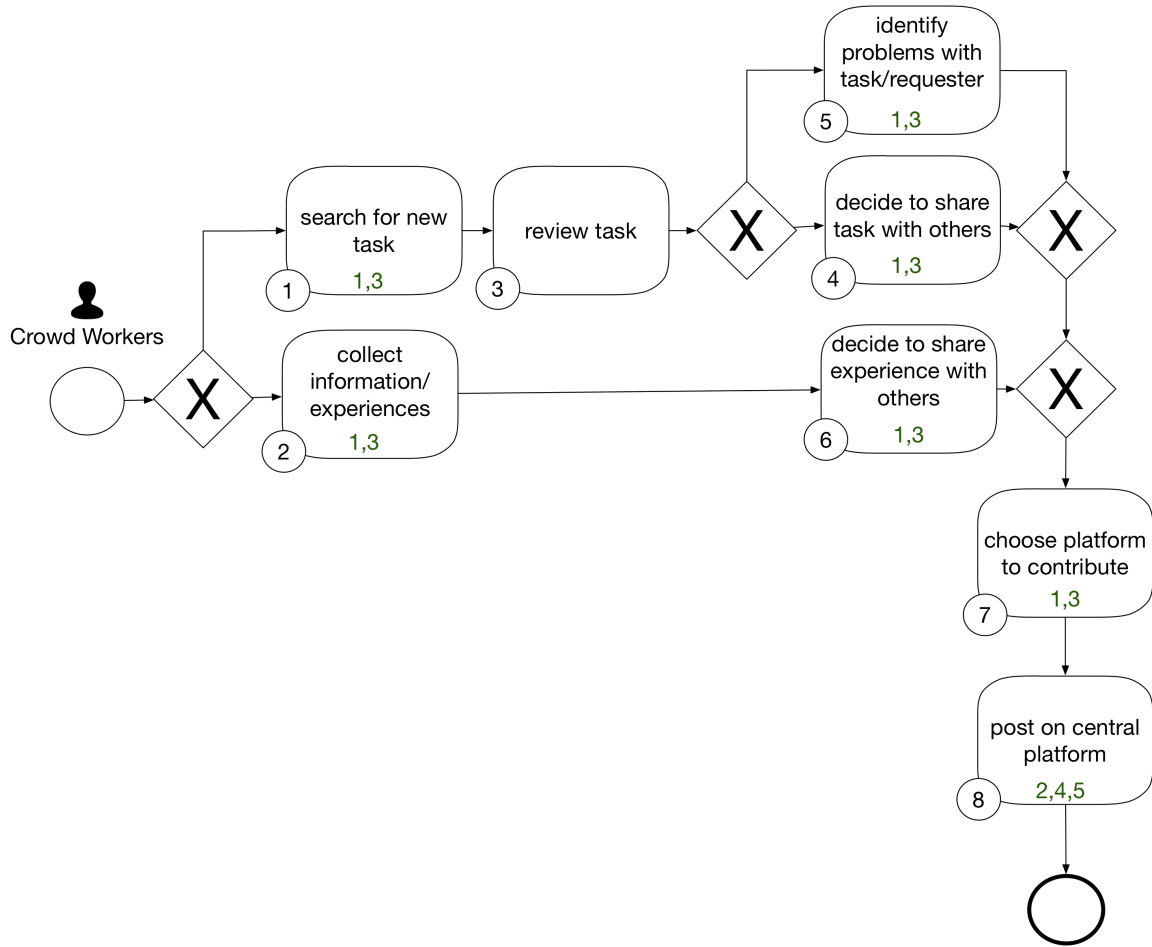


Figure 6.5: Activities of crowd workers in to-be workflow

Improvement-Limitation Mapping Table

A#	Improvement provided by the CIS	Limitation ID	Agent	System
1, 2, 4, 5, 6, 7	There is only one central system solution providing all activities a worker would like to perform.	1	user, moderator, admin	CIS - Worker-Hub
8	The content of workers is collected and stored in one single system and is provided to all registered crowd workers. There are no restricted groups or necessary registration on multiple platforms.	2	user, moderator, admin	CIS - Worker-Hub

1, 2, 4, 5, 6, 7	A user can join a discussion or share a new task without switching between two systems. All actions are possible on one central platform.	3	user, mod-erator, admin	CIS - Worker-Hub
8	The system offers different gamification-elements and badges to insure a certain level of confidence.	4	user, mod-erator, admin	CIS - Worker-Hub
8	On this CIS national institutions and worker organizations exchange their knowledge with crowd workers. So they have access to the system and deliver reliable information.	5	user, mod-erator, admin	CIS - Worker-Hub
9	It is easy to inform crowd worker because they know where they can get in contact with them	6	external	CIS - Worker-Hub
10	Members from an advocacy know where they can collect data from the crowd workers	7	external	CIS - Worker-Hub
12	There is one central system where advocacy members have to register. They can get in contact with the crowd workers and they are also highlighted as a trustful source. Also, it is possible to ask them about their working life and collect data from them.	8 & 9	external	CIS - Worker-Hub

Table 6.3: List of improvements provided by the to-be workflow

6.2 CI Technical Realization Viewpoint

This section describes the architectural view based on the technical realization viewpoint. We describe the CI artifact and its content in more detail as well as define possible links between them. The artifact content is described in Table 6.4 where we break down the artifact in different attributes, each with a name, type and description. Then, we define the operation specifications in context of the CI artifact. Furthermore, we describe the aggregation model which consists of an overview of all agents, their roles, which are shown in Table 6.5 and Table 6.7 describes the aggregation mechanisms applied in the CIS. In Table 6.8 all input and outputs of the artifact are summarized. The last part of this section describes the dissemination model, which defines how analyzed content is

distributed and stimulate subsequent actor activities and is shown in Table 6.9.

6.2.1 Artifact Definition Model

CI Artifact

Group

A group contains specific information from a certain discussion. For example: A group about the legal situation in crowd working in Austria. Also, a group is divided into two parts. One part is articles where members from an advocacy can post information to the certain discussion. The other part is the discussion board where every user in the system can start a discussion.

Artifact Link

Tags

A tag is property, which describes a group. Tags are important for the preference system. The preference system helps the user to find groups of her interest. So a user can be interested in the crowd working platform Amazon Mechanical Turk, then she could choose the tag and all groups who include this tag are shown to her.

Categories

A group has one category. Categories are generic term to categorize groups. For example: in the central system a category about all crowd working platforms is established. In this category are only groups, which deal with a certain crowd working platform (e.g.: Upwork)

Artifact Content

Table 6.4 shows the name, type and description of the artifact attributes in a technical context that form the artifact content.

Attribute Name	Attribute Type	Attribute Description
group ID	integer	Every group has a unique ID
group name	string	A group has a unique name
group description	string	description of the discussion of the group
group deleted	boolean	if a group is deleted or not
group unlocked	boolean	A crowd worker can suggest a group. An external or admin has to unlock a group to make it visible for all users.
group category	string	The category of a group. There are only two categories: "General" and "Specific". In the "General" category are only groups, which deals with general discussions about crowd working like the legal situation in Austria. The "Specific" category contains information about specific crowd working platforms (e.g. AMT)
group tag list	set<string>	A set of tags can be assigned to a group.

group articles	set<article>	A group contains a set of articles. Articles are official information from an external.
group discussions	set<discussion>	Every group has a set of different discussions. A discussion includes a concern or information from a worker or external which can be discussed by the users.
group userfavs	list<user>	A list of all users who defined this group as their favorite.
group created at	timestamp	Date and timestamp when a group was created
group updated at	timestamp	Date and timestamp when a group was last updated

Table 6.4: Artifact Content

Operation Specifications

- **CREATE.** A crowd worker can create a discussion or a post. Advocacy can additionally create articles and groups.
- **REQUEST.** Only a crowd worker can request a group. This should prevent uncontrolled growth of groups.
- **ADD.** A user in the system can add a group to her favorite groups.
- **READ.** A user in the system can read groups, articles, discussions or posts from other users.
- **UPDATE.** A crowd worker can update her discussions and posts. Additionally, A member from an advocacy can update groups and articles.
- **DELETE.** Within the delete operation a delete-flag is set. Advocacy can delete her discussions, groups, articles and her posts. A crowd worker can only delete her discussions and posts.
- **VOTE.** A user in the system can vote on discussions and posts.
- **POST.** Crowd workers and advocacy can post their comment under a discussion or article.
- **REPORT.** It is possible that user report each other for bad behavior.
- **ADD TAG.** Advocacy can add some tags to groups.

6.2.2 Aggregation Model

The aggregation model gives insights into the agents, actor roles and their activities.

Agents

In Table 6.5 we give an overview of all agents, the actor roles as well as a detailed description of each role defined for the CIS. The group of agents comprises actors and observers, whereby actors are users who actively contribute information to the system and thus need a registration, and observers are not registered users and thus are not allowed to contribute content. Observers can only passively consume information. In the CIS of interest both agent groups are defined. Furthermore, we define four roles for actors in the system: user, moderator, admin and external.

Agent	Actor Role	Actor Role Description
Observer		
Actor	User	A user can only be a crowd worker. The user contributes content, sets alerts, does surveys or communicates with other workers.
Actor	Moderator	A moderator is an experienced crowd worker. She has more permission like handling the reporting of a user and manage the requested groups.
Actor	Admin	An admin is a special user who has access to every information in the system. This role is limited to only few users.
Actor	External	An external is a member of the advocacy group contributes official content, interacts with the crowd workers and analyzes the surveys to improve strategies and give recommendations for politicians.

Table 6.5: List of agents

Actor Record

In Table 6.6 we break down the attributes of our actors and describe them. Therefore, we get a bigger picture what data we need from the user.

Attribute Name	Attribute Type	Attribute Description	Actor Role
ID	Integer	An unique ID to identify a user.	no specific actor role
username	String	Required. Unique name for the user.	no specific actor role
firstname	String	Optional. First name of the user	no specific actor role

lastname	String	Optional. The Last name of the user.	no specific actor role
password	String	Required. The password of the user with a minimum length of six characters.	no specific actor role
email	String	Required. The e-mail address of the user.	no specific actor role
city	String	Optional. Origin city of the user	no specific actor role
country	String	Origin country of the user	no specific actor role
deleted	Boolean	Sets a deleted flag if a user don't want to participate anymore.	no specific actor role
created at	Date	Registration date of the user	no specific actor role
updated at	Date	If the user changes her information	no specific actor role
password confirmation	String	Required. The user has to indicate the password.	no specific actor role
role	String	Every user has a certain role	no specific actor role
score	Integer	A user has a score. This score can increase through participation in the discussions	no specific actor role
preferences	List<Tags>	A list of all preferences of the user.	no specific actor role
favorites	List<Groups>	A list of all favorite groups of the user	no specific actor role

Table 6.6: List of agents

Agent Activities

Table 6.7 provides an overview of the activities, the types of activities (read and write), their description, as well as the associated agents (actor or observer) and the actor role that is allowed to perform this activity. The actor role can be user, moderator, external or admin.

Act. #	Type	Activity	Activity Description	Agent	Actor Role
1	W	Create Group	Creates a new group to collect and share information.	Actor	Advocacy(external) or moderator, admin
2	W	Request Group	Create a request to create a new group.	Actor	user, moderator, admin

3	R	Show Group	Access a specific Group and show it's details	Obs.	
4	W	Modify Group	Modify the details of a group.	Actor	Advocacy (external) or moderator, admin
5	W	Tag Group	Assign a list of tags to a group.	Actor	Advocacy (external) or moderator, admin
6	W	Delete Group	Remove a group so that it is not accessible anymore.	Actor	Advocacy (external) or moderator, admin
7	W	Add Article to Group	Add a certain article to a specific group.	Actor	Advocacy (external) or moderator, admin
8	R	Read an Article of Group	Access an certain Article of a specific Group and reading it's content	Obs.	
9	W	Modify Article of Group	Modify the content and title of a certain article of a specific group	Actor	Advocacy (external) or moderator, admin
10	W	Delete Article of Group	Remove a certain article of a specific group so that it is not accessible anymore.	Actor	Advocacy (external) or moderator, admin
11	W	Post on Article of Group	Create a comment under an certain article of a specific group.	Actor	user, moderator, external and admin
12	W	Report a Post of User	Report an invalid post of a user under an article.	Actor	user, moderator, external and admin
13	W	Add Discussion to Group	Add a certain discussion to a specific group to discuss about it with other users.	Actor	user, moderator, external and admin
14	R	Read a Discussion of Group	Access a certain Discussion of a specific Group and reading it's content	Obs.	
15	W	Modify Discussion of Group	Modify the content and title of a certain discussion of a specific group including tags.	Actor	Advocacy (external) or moderator, admin or user who has created the discussion

16	W	Assign Tag to Discussion of Group	Assign a list of tags to a discussion.	Actor	Advocacy (external) or moderator, admin
17	W	Delete Discussion of Group	Remove a certain discussion of a specific group.	Actor	Advocacy (external) or moderator, admin or user who has created the discussion
18	W	Post on Discussion of Group	Post under a certain discussion of a specific group.	Actor	user, moderator, external or admin
19	W	Vote on Discussion of Group	Add a vote to a certain discussion of a specific group.	Actor	user, moderator, external and admin
20	W	Vote on Post of Discussion	Add a vote to a certain post of a specific discussion	Actor	user, moderator, external and admin
21	W	Report a post of User	Report an invalid post of a user under a discussion.	Actor	user, moderator, external and admin
22	W	Report Discussion of User	Report an invalid discussion of a user in a certain group.	Actor	user, moderator, external and admin

Table 6.7: List of Agent Activities

Artifact Input/Output Flows

Table 6.8 shows for every defined activity the client that can be used to perform this activity, the data input and output and if the activity is logged in the actor record or not. Input and output show what kind of data is provided by the actors and stored in the CI artifact and what kind of information from the CI artifact is consumed by the actors.

Act. #	Client	Artifact Input	Artifact Output	Actor Record Log
--------	--------	----------------	-----------------	------------------

1	Browser Client	<ul style="list-style-type: none">• group name (string)• group description (text)• group unlocked (boolean)• group category (string)• group tag list (set<string>)• group created at (timestamp)• group created by (user)		Yes
2	Browser Client	<ul style="list-style-type: none">• group name (string)• group description (text)• group category (string)• group tag list (set<string>)• group unlocked (boolean)		Yes

3	Browser Client		<ul style="list-style-type: none">• group name (string)• group description (text)• group category (string)• group tag list (set<string>)• group article list (set<article>)• group discussion list (set<discussion>)• group userfav list (list<user>)• group created at (times-tamp)	Yes
4	Browser Client	<ul style="list-style-type: none">• group name (string)• group description (text)• group category (string)• group unlocked (boolean)• group tag list (set<string>)• group updated at (times-tamp)• group updated by (user)		Yes
5	Browser Client	<ul style="list-style-type: none">• tag list (set<string>)		Yes

6	Browser Client	group deleted = yes. Deleted Flag is set		Yes
7	Browser Client	<ul style="list-style-type: none"> • article title (string) • article description (text) • article private (boolean) • article comments allowed (boolean) • article created at (timestamp) • article created by (user) 		Yes
8	Browser Client		<ul style="list-style-type: none"> • article title (string) • article author (string) • article text (text) • article created at (timestamp) • article list of posts (List<Pos>) 	Yes
9	Browser Client	<ul style="list-style-type: none"> • article title (string) • article description (text) • article private (boolean) • article comments allowed (boolean) • article updated at (timestamp) • article updated by (user) 		Yes

10	Browser Client	article deleted = yes. Deleted Flag is set		Yes
11	Browser Client	<ul style="list-style-type: none">• post title (string)• post text (text)• post created at (timestamp)• post created by (user)		Yes
12	Browser Client	<ul style="list-style-type: none">• report reason (text)• reported at (timestamp)• reported by (user)		Yes
13	Browser Client	<ul style="list-style-type: none">• discussion title (string)• discussion text (text)• discussion tag list (set<string>)• discussion created at (timestamp)• discussion created by (user)		Yes

14	Browser Client		<ul style="list-style-type: none"> • discussion name (string) • discussion author (string) • discussion text (text) • discussion created at (timestamp) • discussion list of post (list<post>) • discussion number of votes (integer) • discussion number of votes on posts (integer) 	Yes
15	Browser Client	<ul style="list-style-type: none"> • discussion title (string) • discussion description (text) • discussion tag list (set<string>) • discussion updated at (timestamp) • discussion updated by (user) 		Yes
16	Browser Client	<ul style="list-style-type: none"> • discussion tag list (set<string>) 		Yes
17	Browser Client			No

18	Browser Client	<ul style="list-style-type: none">• post title (string)• post text (text)• post created at (timestamp)• post created by (user)• post number of votes (integer)		Yes
19	Browser Client	<ul style="list-style-type: none">• vote (integer: 1 or -1)		Yes
20	Browser Client	<ul style="list-style-type: none">• vote (integer: 1 or -1)		Yes
21	Browser Client	<ul style="list-style-type: none">• report reason (text)• reported at (timestamp)• reported by (user)		Yes
22	Browser Client	<ul style="list-style-type: none">• report reason (text)• reported at (timestamp)• reported by (user)		Yes

Table 6.8: List of artifact input/output flows

6.2.3 Dissemination Model

The dissemination model describes the rules what content and how this content is disseminated back to the users. Table 6.9 describes the dissemination rules comprising the used trigger, the occurring frequency (schedule), the component which performs the analysis of content for dissemination (analyzer), the kind of data used for the analysis

and the expected filtered output to be disseminated, the communication channel used to show the information to the user, the purpose of the rule and the recipients.

Trigger	Schedule	Analyzer	Data Source / Filtered Output	Channel	Purpose	Recipient
Show personal recommendations on the dashboard	Every time a new group is added or the user make a recommendations	Recommender System	Data Source: Actors Records: Preferences, Artifact Content: Tags Filtered Output: Recommended groups based on interests selected by the user.	Platform	Show personal recommendations to a user so that she can discover easier groups of interest.	All actors
Show notifications of new posts and new alerts	A new post to a discussion or a new task on a platform or a new alert	Notification Builder	Data Source: Artifact Content: new posts, new tasks Filtered Output: Notifications about new post and new alerts	Platform	User should be aware about new posts and alerts to be always up-to-date and do not miss a lucrative task.	All actors

Table 6.9: List of dissemination rules

6.3 CI Operation Viewpoint

The CI operation viewpoint takes a closer look at the initial content for the CI artifacts and a possibly needed transformation of the sources, which is shown in Table 6.10. In addition, the initial actor profile is described where we clarify expertise and recruitment. Finally, we describe the metrics and probes to capture data of interest of our CI analytics

model (see Table 6.11).

6.3.1 Initial Content Acquisition Model

Initial Data and Transformation

Table 6.10 clarifies from where the initial data are coming and how they need to be transformed to fit into the CI artifact.

Artifact Content	Transformation	External Source	Initial Data
Group Metadata	The content of some discussion are manually transformed into our system	TurkerNation	All existing groups and their articles and discussions
Group Metadata	The content of some discussion are manually transformed into our system	TurkerNation ¹	All existing groups and their articles and discussions
Group Metadata	The content of some discussion are manually transformed into our system	CloudMeBaby ²	All existing groups and their articles and discussions
Group Metadata	The content of some discussion are manually transformed into our system and translated into English.	faircrowd.work ³	All existing groups and their articles and discussions
Alerts Meta-data	The content of alerts are manually transformed into our system	Reddits HWTF ⁴	All existing alert-groups

Table 6.10: Sources for initial CI artifact content and needed transformations

Initial Actor Profile

In the following the required expertise of an initial actor base and their recruitment is described.

- Expertise
 - New Crowd worker
 - Crowd Worker with more expertise
 - Advocacy
- Recruitment

¹<http://turkernation.com> Accessed 28.5.2018

²<http://www.cloudmebaby.com/forums/portal.php> Accessed 6.6.2018

³<http://faircrowd.work/de/>. Accessed 28.3.2018

⁴<https://www.reddit.com/r/HITsWorthTurkingFor/> Accessed 28.5.2018

- Advertisement in different crowd worker forums for our system. Asking worker organization to mention our new system. This strategy is to get more new members and also experienced crowd workers.
- Write to different worker organization and present our system and the advantages for them. This strategy is to get more advocacy members.

6.3.2 CI Analytics Model

Metrics

The CI analytics model gives more insights into the measurement design that is used to analyze the actor activities and feedback loop and provide status metrics of the CIS. The model is presented in Table 6.11 and describes the probes, the time of measurement, which system component is responsible for the analysis, which properties are measured and the metrics.

Metric	Probe	Time of Measurement	System Component	Property
Amount of new groups	Preferences Probe	After creating a new group	Controller Component	Actor activity "new group created"
Amount of users	Preferences Probe	After register a new user	Controller Component	Actor activity
Amount of requested groups	Preferences Probe	After requesting a new group	Controller Component	Actor Activity
Amount of new preferences	Preferences Probe	After adjusting preferences	Controller Component	Actor Activity
Amount of new Posts per week	Notification Probe	After new post under discussion	Controller Component	Actor Activity
Amount of new Alerts per week	Notification Probe	After sending new alert	Controller Component	Actor Activity

Table 6.11: List of metrics

WorkerHub Prototype

WorkerHub is a social platform where crowd workers can share their knowledge, interact with each other or self-organize themselves. The platform will support the crowd workers and provide them all information they need with involvement of advocacy. WorkerHub follows all the defined design principles in a good structured way.

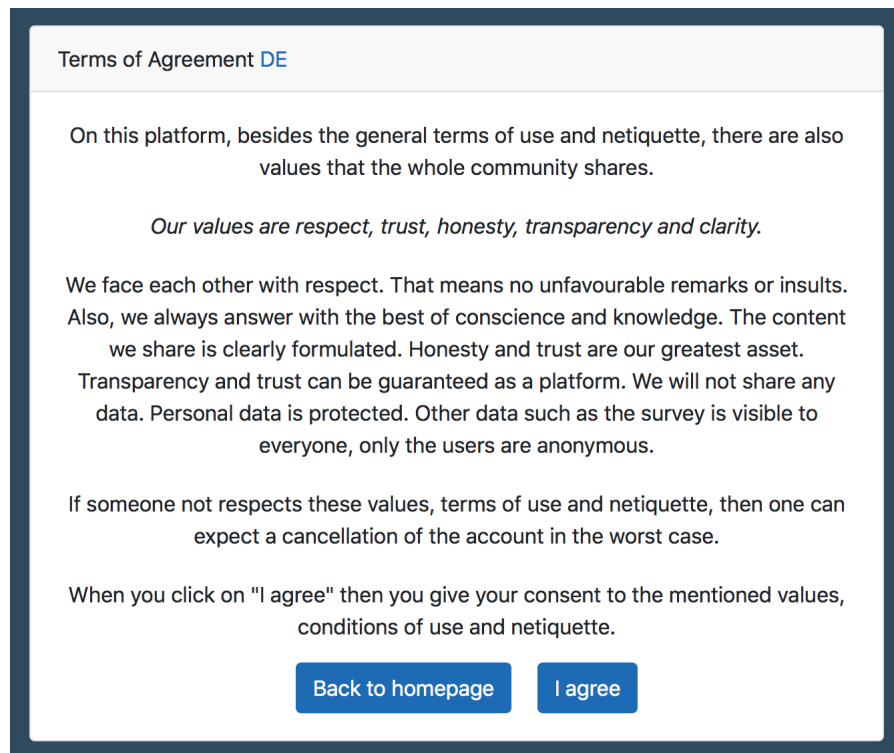
Based on the investigated design principles and the architecture design of a CIS for the crowd work domain, we developed a prototype that implements the proposed design so that we can evaluate it as well as answer RQ 3: "*How do the identified principles support the design of collaborative crowd worker platforms?*". With the development of the prototype "WorkerHub", we aim to investigate the application and usefulness of the proposed design principles and architecture models. WorkerHub should represent a contact point for crowd workers across all crowd work platforms to get in contact with each other and to get reliable and trustworthy information. On WorkerHub workers have the opportunity to get informed, share information and connect with each other in a save and trustworthy environment. It also offers statistics about the current situation of workers and a contact list of official institutions. With the integrated alert system workers get informed about lucrative tasks and can share their experiences. Before going into more detail of all features and how we implemented the defined principles, we present relevant implementation details.

7.1 Register & Terms of Use

WorkerHub is non-exclusive closed. A user who is not logged in can see certain articles, which are marked as public. So, other users get a taste of the content of the platform.

Because of the *Privacy* principle, most of the content is locked for anonymous users. Users have to register on WorkerHub to see more information. Before registration they have to sign our terms of agreement, which are shown in figure 7.1. With our own

terms of use we fulfill the *Fairness and Objective Exchange* as well as the *Distribution of Information and Awareness* principle. We remind the user to behave in a fair and objective way as well as highlight awareness and a community with defined values. Also, we point out the importance of data and how we treat their personal data in the terms of agreement, which also fits into the *Privacy* principle.



Terms of Agreement [DE](#)

On this platform, besides the general terms of use and netiquette, there are also values that the whole community shares.

Our values are respect, trust, honesty, transparency and clarity.

We face each other with respect. That means no unfavourable remarks or insults. Also, we always answer with the best of conscience and knowledge. The content we share is clearly formulated. Honesty and trust are our greatest asset. Transparency and trust can be guaranteed as a platform. We will not share any data. Personal data is protected. Other data such as the survey is visible to everyone, only the users are anonymous.

If someone not respects these values, terms of use and netiquette, then one can expect a cancellation of the account in the worst case.

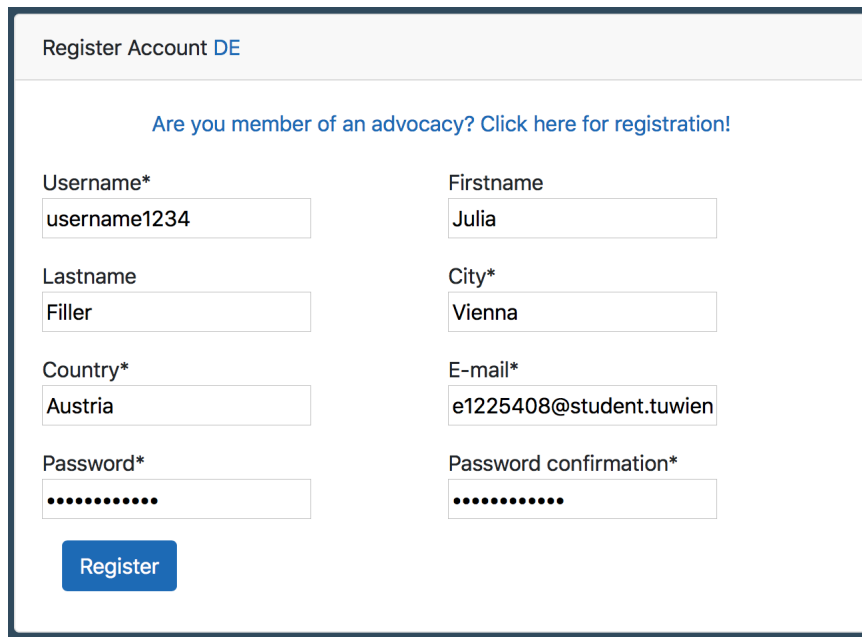
When you click on "I agree" then you give your consent to the mentioned values, conditions of use and netiquette.

[Back to homepage](#) [I agree](#)

Figure 7.1: Terms of agreement on WorkerHub

In the registration process we distinguish between two different users: crowd workers and advocacy member. Figure 7.2 shows the registration form for crowd workers. In both registration processes we ask the user to indicate a username, first name, last name, city, country, e-mail address and password. First name and last name is not obligatory. If the user is from an advocacy, then the platform additionally asks for the specific institution. Members of advocacy have more rights and will be listed at the advocacy list.

After the registration, the user gets an e-mail where she has to confirm her e-mail address. Without the confirmation the user cannot log in.



Register Account [DE](#)

Are you member of an advocacy? [Click here for registration!](#)

Username*	Firstname
<input type="text" value="username1234"/>	<input type="text" value="Julia"/>
Lastname	City*
<input type="text" value="Filler"/>	<input type="text" value="Vienna"/>
Country*	E-mail*
<input type="text" value="Austria"/>	<input type="text" value="e1225408@student.tuwien"/>
Password*	Password confirmation*
<input type="password" value="....."/>	<input type="password" value="....."/>

Figure 7.2: Registration form for crowd workers

We only ask important data from our users as described in the *Privacy* principle. For our platform it is important to know the city and country because of the recommendations we create and provide to the user and because users tend to connect with other users from the same country/city [24]. We protect the identity of a user with the username and encrypt the password with BCrypt. In figure 7.3 the encryption process is shown. We generate a salt value, which helps to encrypt the password of the user and store this value in the database.

```
def encrypt_password
  if normalpassword.present?
    self.salt = BCrypt::Engine.generate_salt
    self.password = BCrypt::Engine.hash_secret(normalpassword, salt)
  end
end
```

Figure 7.3: Code snippet from the encryption process

7.2 Users, Role Management & Reporting System

The platform applies a role management system that addresses the principle *Trust and Reputation*. Not every user has the same permissions to do activities within the CIS and thus this system guarantees a safe space for the user base. We defined four roles: (1)

user, (2) moderator, (3) admin, (4) external. The admin role alone has the permission to perform every possible action including delete or edit a user, and edit a contribution of another user. In addition, this role can change the roles of other users. The roles external and moderator have the same permissions. They can create a group or an article and check if a requested group is appropriate or not and activate a group as shown in figure 7.4. In this example a user has suggested a specific platform group named Foodora. A click on "Show" delivers an overview of the asked group and especially a link to the platform. The link is obligated and important for the decision whether the group fits in the context of the CIS.

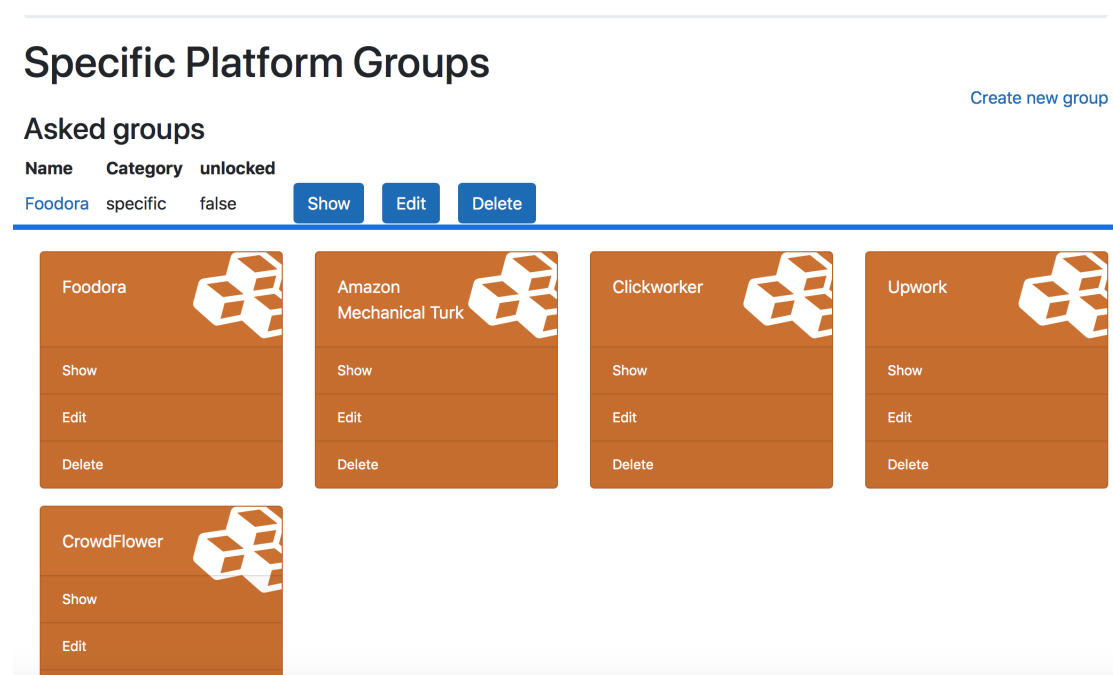


Figure 7.4: Interface of handling requested groups

In addition, these roles are responsible for handling the reporting system and check if the content of the groups is up to date. Furthermore, they need to check if every specific group has an alert, delete or edit an alert group. The difference between both roles is that users with external role have a label that mark them as a member of advocacy and they are shown in the list of advocacy.

A normal user can

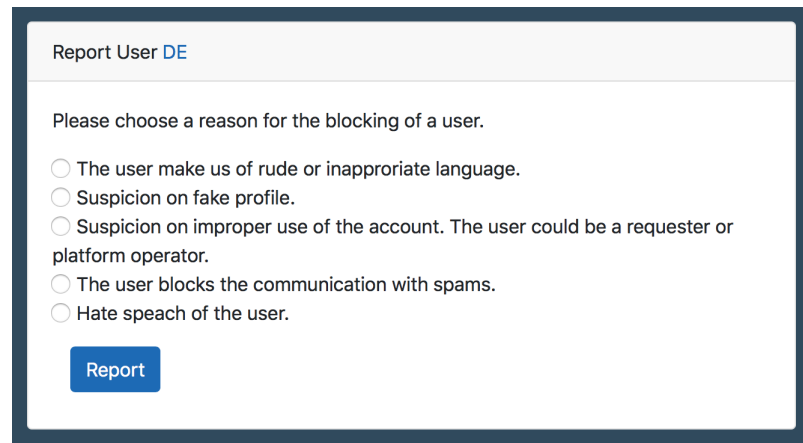
- read all provided information,
- request a group or an alert group,
- add a group to favorites,

- create a discussion or participate in a discussion,
- read all discussions and articles,
- vote on discussions and posts,
- report a user,
- view a user profile,
- do a survey and take a look on the results of the survey,
- get a list of all members of the advocacy,
- set an alert of a lucrative task,
- subscribe to an alert group and view the alert history,
- view all of her notifications and alerts,
- configure the own preference system.

A person who is not logged in, can only see certain articles and get a list of all members of the advocacy. Every person who would like to contribute to the CIS needs to register and provide some personal information. An overview of all registered users is accessible for every role except the role user. The view also helps to handle reported users, which is explained in the following.

The reporting system serves two principles: *Trust and Reputation* and *Fairness and Objective Exchange*. A user can report another user if a discussion or post of her has been inappropriate. Moderators, admins or externals can review then the respective post and, if confirmed, block the user, otherwise the user report is rejected and deleted. The reporting button can be found in the right corner of every contribution. Figure 7.5 shows the reasons to report a user that every user has to select as rationale. The reason helps to categorize whether this reporting is appropriate or not. Worth to mention is that we implemented the reason "Suspicion on improper use of the account. The user could be a requester or platform operator". We want to prohibit that requesters or platform operators distribute wrong information.

Figure 7.6 shows the reporting overview in detail. It shows what user is involved, which post or discussion is reported and in which context. It is also possible to go to the discussion of interest. Based on this information, a user with respective rights can easy decide if it is an appropriate reporting or not. In figure 7.6 the user worker23 is reported for the reason "The user make use of rude or inappropriate language." The post of the reported user is shown as well as the discussion to which the reported post is associated.



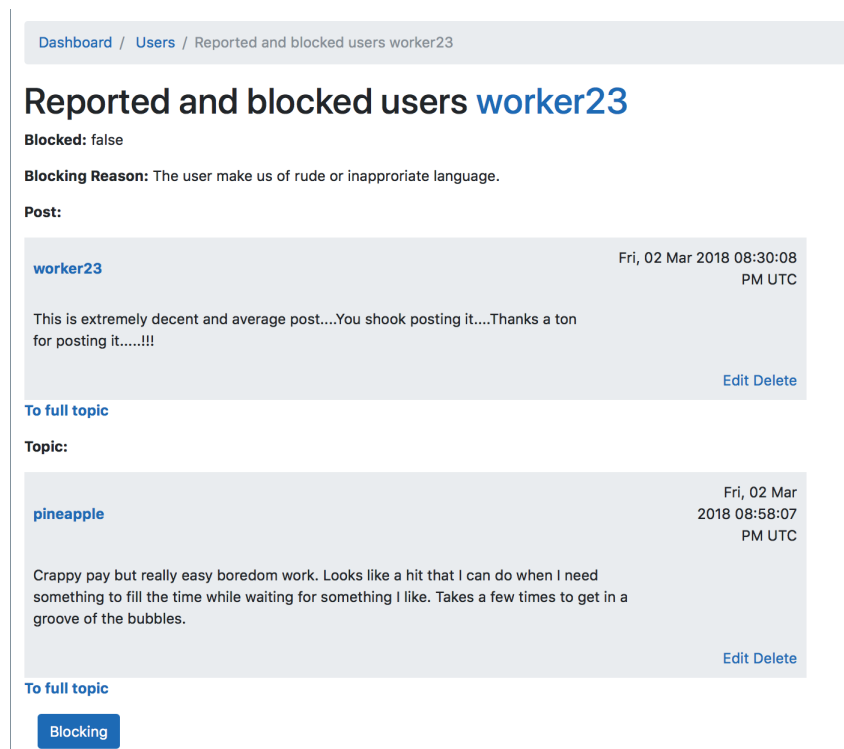
Report User DE

Please choose a reason for the blocking of a user.

- ☐ The user make us of rude or inapproriate language.
- ☐ Suspicion on fake profile.
- ☐ Suspicion on improper use of the account. The user could be a requester or platform operator.
- ☐ The user blocks the communication with spams.
- ☐ Hate speech of the user.

[Report](#)

Figure 7.5: Reporting reasons



[Dashboard](#) / [Users](#) / Reported and blocked users worker23

Reported and blocked users worker23

Blocked: false

Blocking Reason: The user make us of rude or inapproriate language.

Post:

worker23 Fri, 02 Mar 2018 08:30:08 PM UTC

This is extremely decent and average post....You shook posting it....Thanks a ton for posting it.....!!!

[Edit](#) [Delete](#)

[To full topic](#)

Topic:

pineapple Fri, 02 Mar 2018 08:58:07 PM UTC

Crappy pay but really easy boredom work. Looks like a hit that I can do when I need something to fill the time while waiting for something I like. Takes a few times to get in a groove of the bubbles.

[Edit](#) [Delete](#)

[To full topic](#)

[Blocking](#)

Figure 7.6: Interface of handling reporting a user

7.3 Preference System & Dashboard

The preference system does not directly address any of the design principle. It helps a user to find groups of interests very fast and in our CIS architecture design it acts

as a stimulus. figure 7.7 shows a list of different terms. These terms are derived from the assigned tags to a group. Terms, which are greyed out are tags that a user already has chosen. At any time, a user can change her preferences. According to the chosen preferences the user get suggestions of groups which have one or more preferences in their tag list that could be interesting for the user to discover new information.

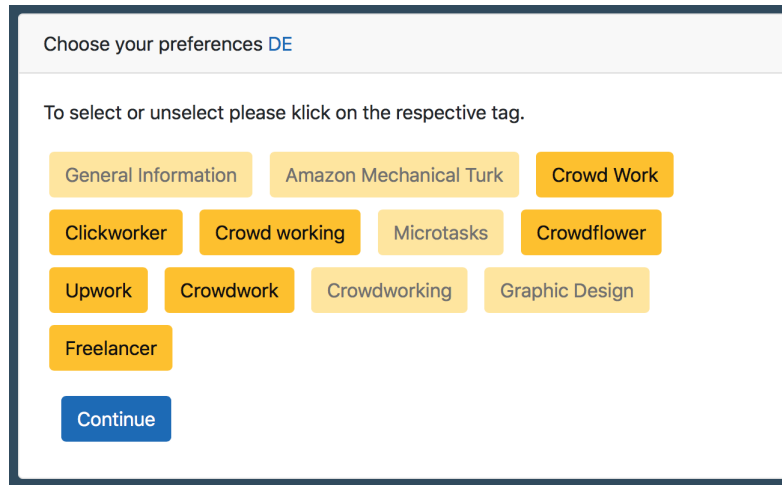


Figure 7.7: Interface of the preference system

The dashboard is a central element of the user interface of the platform. It shows on one place all relevant groups and information for the user based on her preferences. First of all, the dashboard offers an overview of all received notifications, all set alerts and provide the opportunity to do a survey when ever the user likes. In addition, the user can optimize her preferences and favorite groups of the user are highlighted in green color. The alerts where the user is subscribed are presented in red and at the bottom of the dashboard interesting groups are recommended to her. An example of a user dashboard is shown in figure 7.8.

7.4 Groups & Notifications

As established in chapter 6, the CI artifact is the group. Groups are used for the distribution of information and therefore they address the *Distribution of Information and Awareness* principle. We differentiate between two kinds of groups. One is the General Information Group where only common discussions are discussed like the legal situation in Austria or crowd working in general. The other one focuses more on specific crowd work platforms like all discussions about Amazon Mechanical Turk or Clickwork. This section is called Specific Platform Groups. A screenshot of an overview of specific platform groups is presented in figure 7.9.

A normal user cannot edit or delete a group, nor create one. For normal users it is

7. WORKERHUB PROTOTYPE

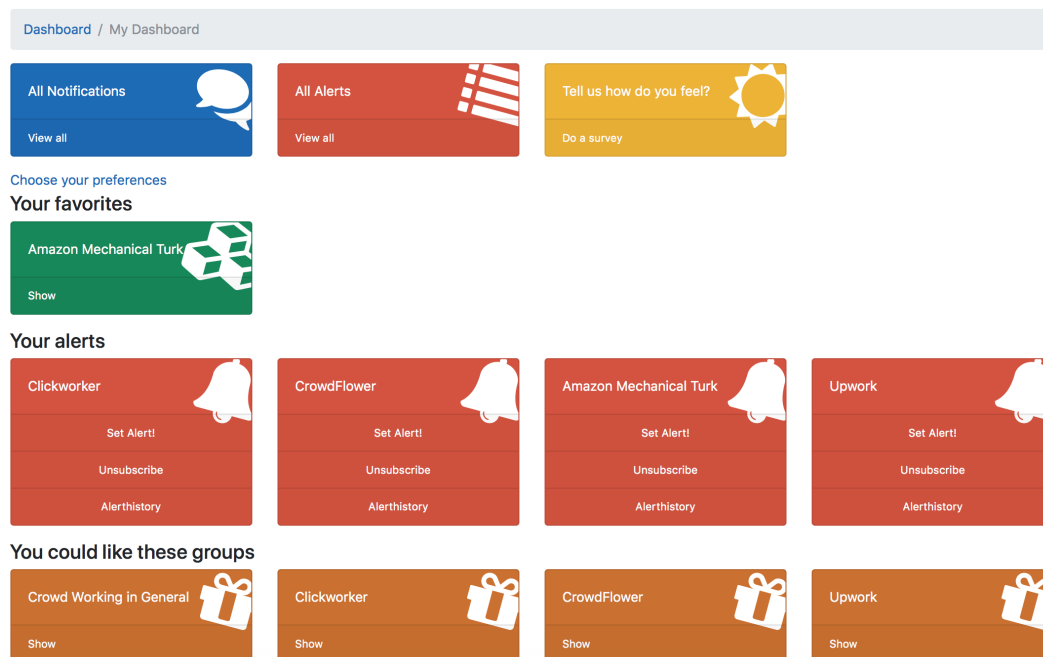


Figure 7.8: Dashboard on WorkerHub

possible to request a group and then an administrator or external can check the proposal and accept it or not. This action should prevent to have double groups or groups that make no sense. figure 7.9 shows the navigation elements on the side and every interface provides breadcrumbs, which are shown over the heading of the site.

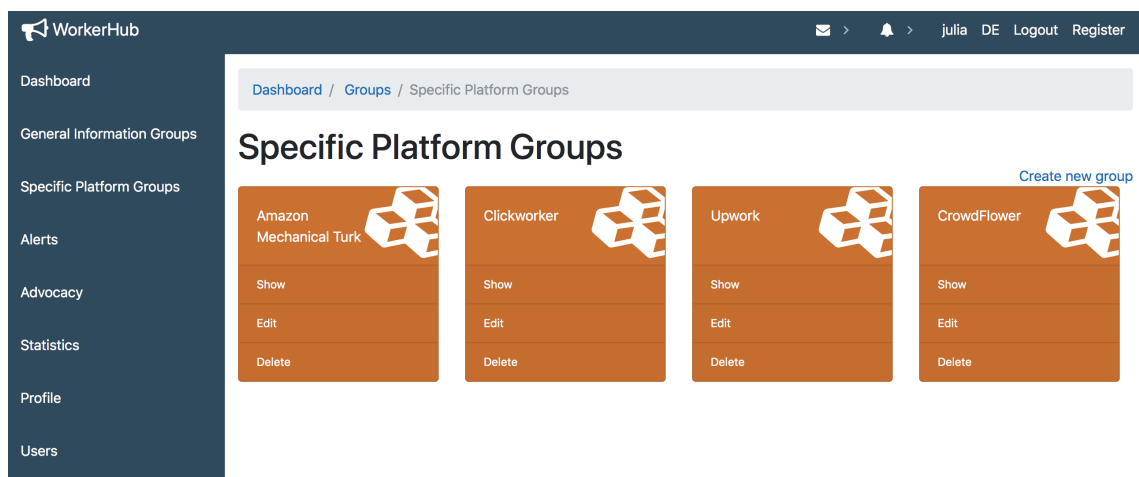


Figure 7.9: Detailed view of the section Specific Platform Groups

Figure 7.10 shows a group in detail. Every group has a name, description and a list of

tags, which are important for the preference system. Additionally, it is possible to add a group to user favorites for a faster access on the dashboard. A group contains two subsections: Articles and a Discussion Board. A normal user cannot create articles. They are contributions from members of an advocacy and include topics of general interest. The discussion board is for the information exchange between the crowd workers. An

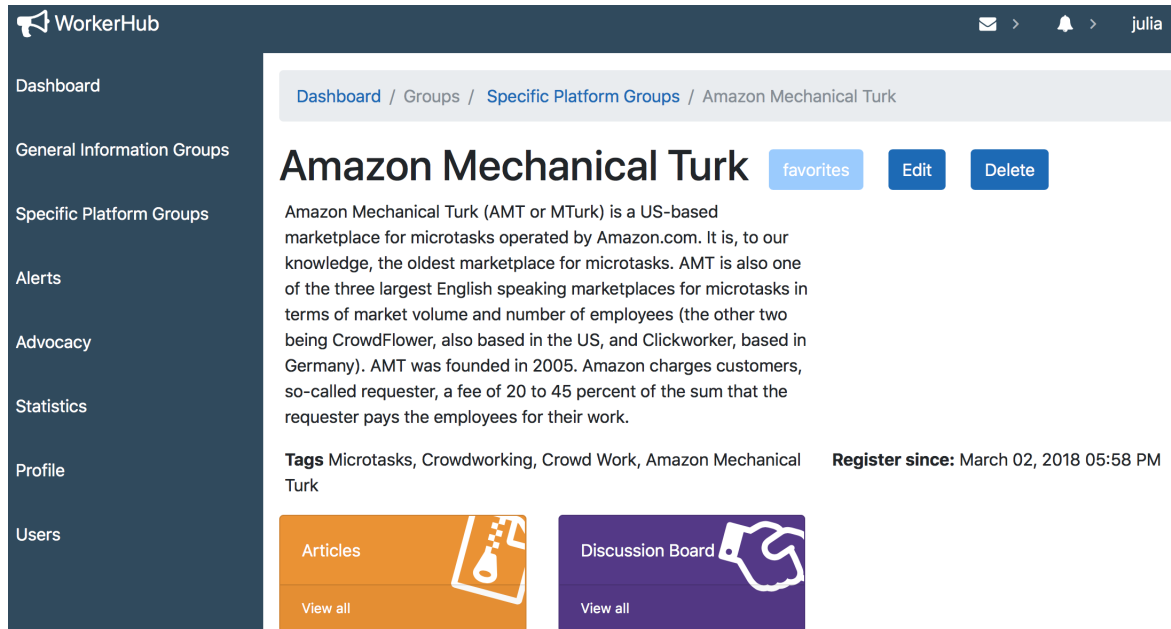


Figure 7.10: Group in detail

article can be private or public. If an article is private then it is only visible for registered users. An article always has an author, a creation date and content. Sometimes it is possible that users can comment via a post to an article. That is the decision of the author if she allows this feature or not. Articles are elements that address the principle *Trust and Reputation*. They are provided by a reliable source of information, e.g., an interest group. The content is trustworthy. It is an element for externals to present important information on one place. In this ways important information for every crowd worker can be aggregated here and is not spread over several platforms, discussions, and contributions and is easy to find.

The discussion board is a further important element for the crowd workers to communicate and interact with others and it acts similar to a forum. Every user can create a new discussion, which can be an issue or question and other users on the platform can answer. The latest discussion is always on top. A discussion consists of a title and a description of the problem and a list of posts as a reaction to the discussion created by users. figure 7.11 shows a discussion in detail with title, author and creation date, vote counter and content. Every user can add a new post to a discussion. A post consists of an author, creation date, vote counter and content.

The voting of posts is an important feature on WorkerHub, because it helps to highlight relevant posts and punish posts which disrupt the discussion. The vote counter is on the left side of every contribution. We fulfill with the vote counter the principle *Trust and Reputation*, because the vote is on one hand a gamification tool and on the other hand a tool to ensure that reliable and trustworthy contributions are weighted more. The user gets rewarded if she gives a negative or positive vote. If a user receives a positive vote for her post or discussion, then she will earn a point on her user account. If the user receives a negative vote, then she will lose one. It is not possible to vote on the same discussion or post more than one time. With this system we aim to filter out trolls or users who want to make troubles. It is not possible for a normal user to modify or delete every post or discussion of another user. The only one who is allowed to do this are the users with the role of administrator. So the normal user as well as externals can only modify and delete her own contributions.

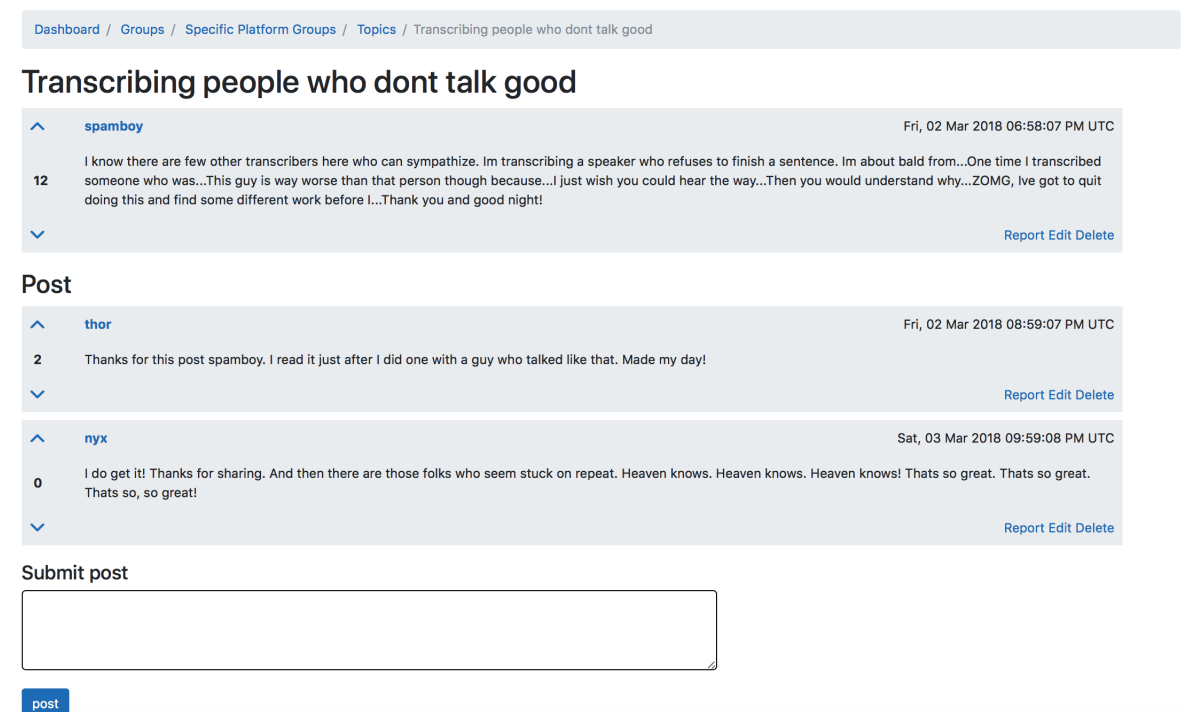


Figure 7.11: A discussion in detail

Notifications represent stimuli for subsequent actions of the users according to our CIS architecture models. There are two possibilities how a user can receive a notification. Either a user has opened a new discussion and another user adds a post to her discussion, or a user has posted something to a discussion and another user replies on a post. Figure 7.12 illustrates the header of the application where it shows an envelope, a bell, the username, a symbol to switch language and logout. The envelope notifies the user that another user reacted to her discussion or post. The bell shows all new alerts, which are

described in the next section. With a click on the envelope, a list of new notifications is opened and the specific discussion is viewed with the title of the topic and the creation date. It is also possible to show a list of all received notifications.

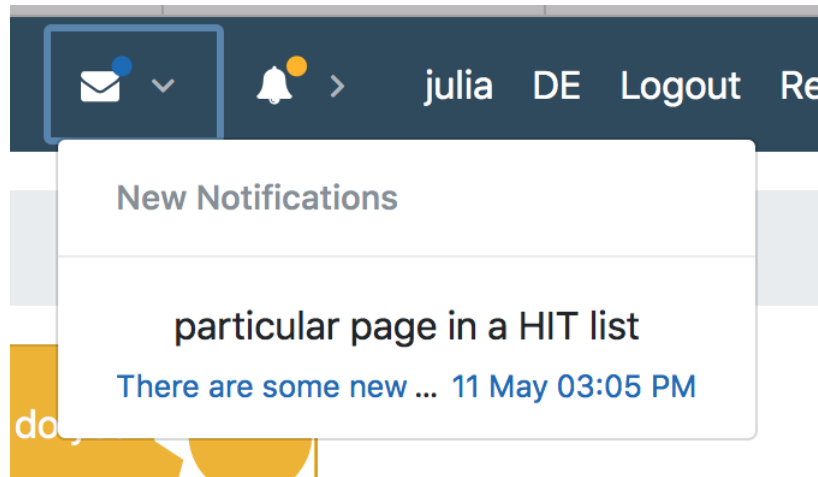


Figure 7.12: List of new notifications

7.5 Alert System

Besides the notifications, a user can receive alerts. A user can subscribe to an alert of a specific group representing a crowd work platform to receive a notice when another user has found a new interesting task. A user can also show a history of all dropped alerts and set an alert by her if she has found a lucrative task she wants to share with others. In figure 7.13 an alert in detail is shown. An alert consists of the name of the user who set the alert, the title of the task, requester, revenue, overhead and link to the task, creation date and other information which can be useful. The field for other information is relevant, since every crowd work platform delivers different additional information, e.g., in figure 7.13 the other information is ">90% easy MC" which means that the task includes easy multiple choice.

Alerts: Upwork

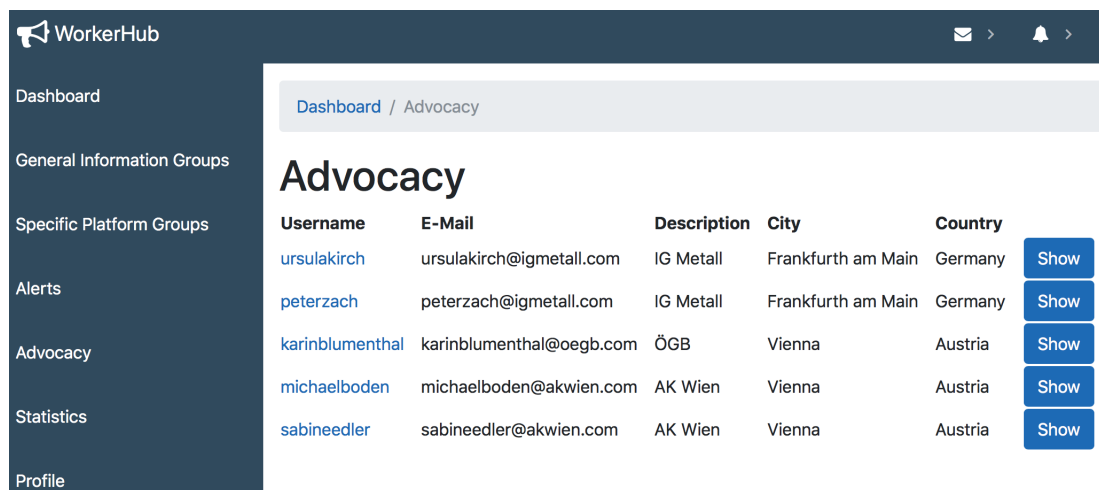
julia: Learning about new objective game		
Requester: Communication and Learning Lab	Revenue: 0,15\$	Overhead/Time: 0:30 min
>90 % easy MC	Sat, 12 May 2018 01:47:05 PM UTC	
Link: https://upwork/projects/3VPC65W52NMIHLNXOBFMQ52POEAMIK/tasks?ref=w_pl_prvw		

Figure 7.13: Alert in detail

If a new alert is dropped then every user who is subscribed to this alert group will get a notification. Then in the header appears a sign on the bell symbol. With the alert system we fit the principle *Distribution of Information and Awareness*, because we offer crowd workers a tool where they can easily and quickly share and identify new lucrative tasks.

7.6 Advocacy

WorkerHub provides all users a list of all members of an advocacy that are registered in this CIS, which is shown in figure 7.14. This list consists of the username, e-mail address, description, city and country as well as a link to their profile. With this feature we address the principle *Trust and Reputation*. Users should know who the people are behind the advocacy role and how they can get in contact with them.



Username	E-Mail	Description	City	Country	
ursulakirch	ursulakirch@igmetall.com	IG Metall	Frankfurth am Main	Germany	Show
peterzach	peterzach@igmetall.com	IG Metall	Frankfurth am Main	Germany	Show
karinblumenthal	karinblumenthal@oegb.com	ÖGB	Vienna	Austria	Show
michaelboden	michaelboden@akwien.com	AK Wien	Vienna	Austria	Show
sabineedler	sabineedler@akwien.com	AK Wien	Vienna	Austria	Show

Figure 7.14: List of all members of advocacy

7.7 Survey & Statistic

In our Prototype it is possible to do a survey. This feature fulfills the *Analyzability and Transparency* principle, because with regular surveys we give crowd workers the opportunity to share their working life, current conditions and experiences with others. In return, the advocacy has the possibility to collect data about the current situation of the crowd workers. The survey collects the data in an anonymous way, which is important for the crowd workers. A user is reminded two times a month to do a survey and she can access the survey at the dashboard. The survey consists of five questions and the users can provide an answer according to their experiences. Figure 7.15 shows one question of the survey. We can see that the users should grade their working conditions. All questions have the same answers. Under the statistics menu the users can view the

results of all survey participants. figure 7.16 shows a few results from a survey of March 2018. Over 77% of the users have participated and over 40% of the participants rated their working conditions as sufficient. For an external it is possible to download the results in form of a CSV or XLS for further analysis. The statistics show how many users participated, the questions and the number of answers and every answer possibility in percentage.

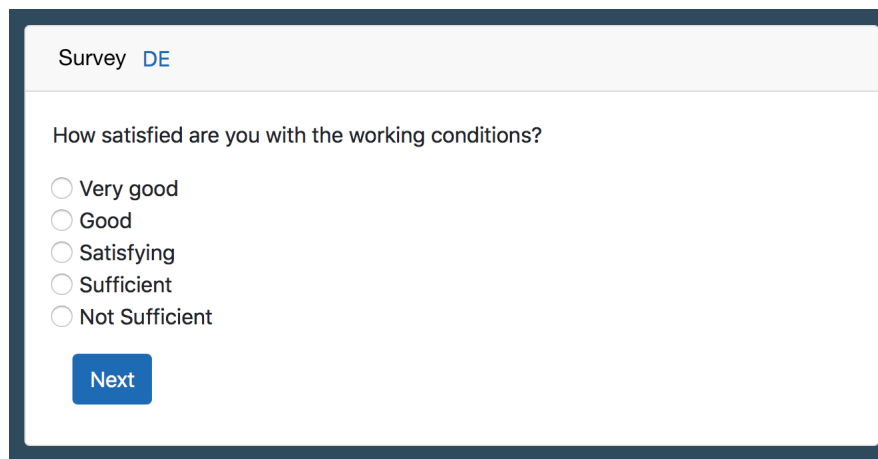


Figure 7.15: A question of the survey

3. 2018

Participants: 77.19 % of 57

How satisfied are you with the working conditions?

2.27%	9.09%	27.27%	43.18%	18.18%
Very good	Good	Satisfying	Sufficient	Not Sufficient

How is the treatment on the platform?

6.82%	18.18%	27.27%	40.91%	6.82%
Very good	Good	Satisfying	Sufficient	Not Sufficient

How satisfied are you with the working situation?

0.0%	11.36%	27.27%	18.18%	43.18%
Very good	Good	Satisfying	Sufficient	Not Sufficient

Figure 7.16: Interface of the statistic site

7.8 User Profile

Every user has a profile. The profile helps users to get to know each other and to establish the *Trust and Reputation* principle. Each user profile shows current status of gamification elements. On the basis of these elements, which include levels, points and number of discussions, posts and alerts, other users can judge the user and assess the relevance of her contributions. There is a difference between the information the user herself is allowed to see on her own profile and the information that other users sees on her profile. The information on the own profile shows every detail the user has stated in the registration process. The profile of another user only shows the gamification elements, username, registration date, city and country as can be seen in figure 7.17. The user pineapple lives in Innsbruck, Austria, and she is either a newbie or not very active on the platform which can be retraced from her level and points but also from her badges (gamification elements).

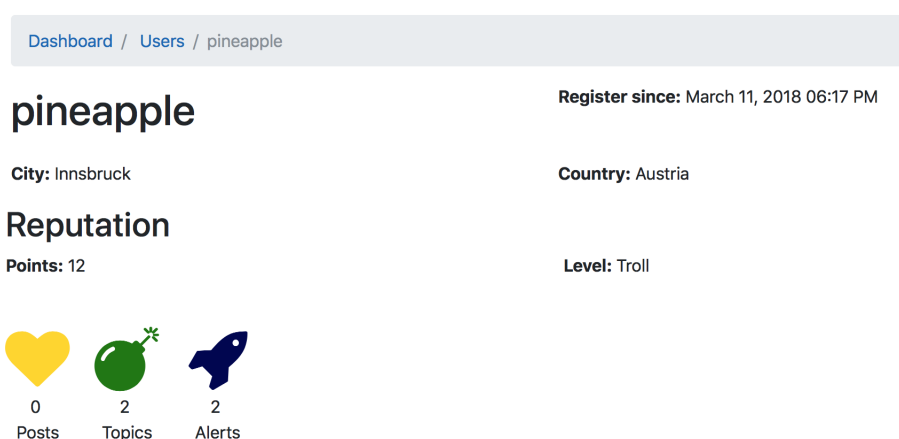


Figure 7.17: Interface of another user's profile

7.9 Implementation Details

WorkerHub is implemented as a web application that is accessible by using a browser client, either on a PC or mobile device. WorkerHub was built as a Ruby on Rails application with a PostgreSQL database. We chose Ruby on Rails, because it allows a very fast and easy prototyping process. As template we used a Bootstrap Template¹ to ensure responsiveness and FontAwesome² was used for the icons. We encrypted the password with BCrypt which is a password hashing scheme[64]. The source code was managed by a Git repository provided by GitHub³ to have versioning and to regularly

¹<https://startbootstrap.com/template-overviews/sb-admin-2/> Accessed 8.5.2018

²<https://fontawesome.com> Accessed 8.5.2018

³<https://github.com/> Accessed 8.5.2018

push new features. For production we chose to push it on Heroku⁴ and WorkerHub can be accessed using this link: <https://workerhub.herokuapp.com>.

Figure 7.18 shows the class diagram of WorkerHub. The central elements in our implementation are the user and the group. Groups are related with articles and discussions, and these contributions allow posts. Also, the scoring system, which increases trust on the platform is linked to discussions, discussion posts and users. The scoring system is responsible for the motivation of a user by giving users points when they are interacting on the system. Then, we implemented a preference system, which helps the users to find their groups of interest very fast. The preference system depends on tags assigned to groups. The alert system consists of a relation between groups, alerts and alert subscriptions. Every alert belongs to an alertgroup and a user can subscribe to an alertgroup. Users can select their favorite groups for faster access on the dashboard. A notification is shown when a discussion has changed. The statistic element manages the survey questions and answers. Finally, reported users are represented by the block element as well as the blocking reason and the problematic post.

In this chapter, we introduced the prototype WorkerHub and explained its features in detail. Furthermore, we showed how we implemented the design principles and the CIS architectural models. Table 7.1 shows us a summary of all features and also, which feature belongs to which design principle. The principle *Independence Centralization* is fulfilled through the platform itself. The register process and the way the data is processed is for the *Privacy* principle. Groups, Advocacy, Profile and Role Management support the *Trust Reputation* principle. The principle *Fairness Objective Exchange* is involved in our Terms of Use and Reporting System. Furthermore, the principle *Distribution of information Awareness* is complied through the Terms of Use, Groups and Alert system. Survey and Statistic meet the principle *Analyzability Transparency*. The last principle *Work Type-Specific Software Clients* is WorkerHub itself, because it is Web App. Our Stimuli on this platform are the Notifications and Preferences.

Design Principle	Feature
Independence & Centralization	WorkerHub itself
Privacy	Register Process
Trust & Reputation	Groups, Advocacy, Profile (Voting System, Gamification Elements), Role Management
Fairness & Objective Exchange	Terms of Use, Reporting System
Distribution of Information & Awareness	Terms of Use, Groups, Alert System
Analyzability & Transparency	Survey, Statistic
Work Type-Specific Software Clients	WorkerHub itself - Web App
Stimuli	Notifications (Posts, Alerts), Preference System

Table 7.1: Overview over design principles and their features on WorkerHub

⁴<https://www.heroku.com> Accessed 8.5.2018

In the next chapter, we present an evaluation of the prototype regarding to the design principles and aim to derive improvements of the prototype and its design.

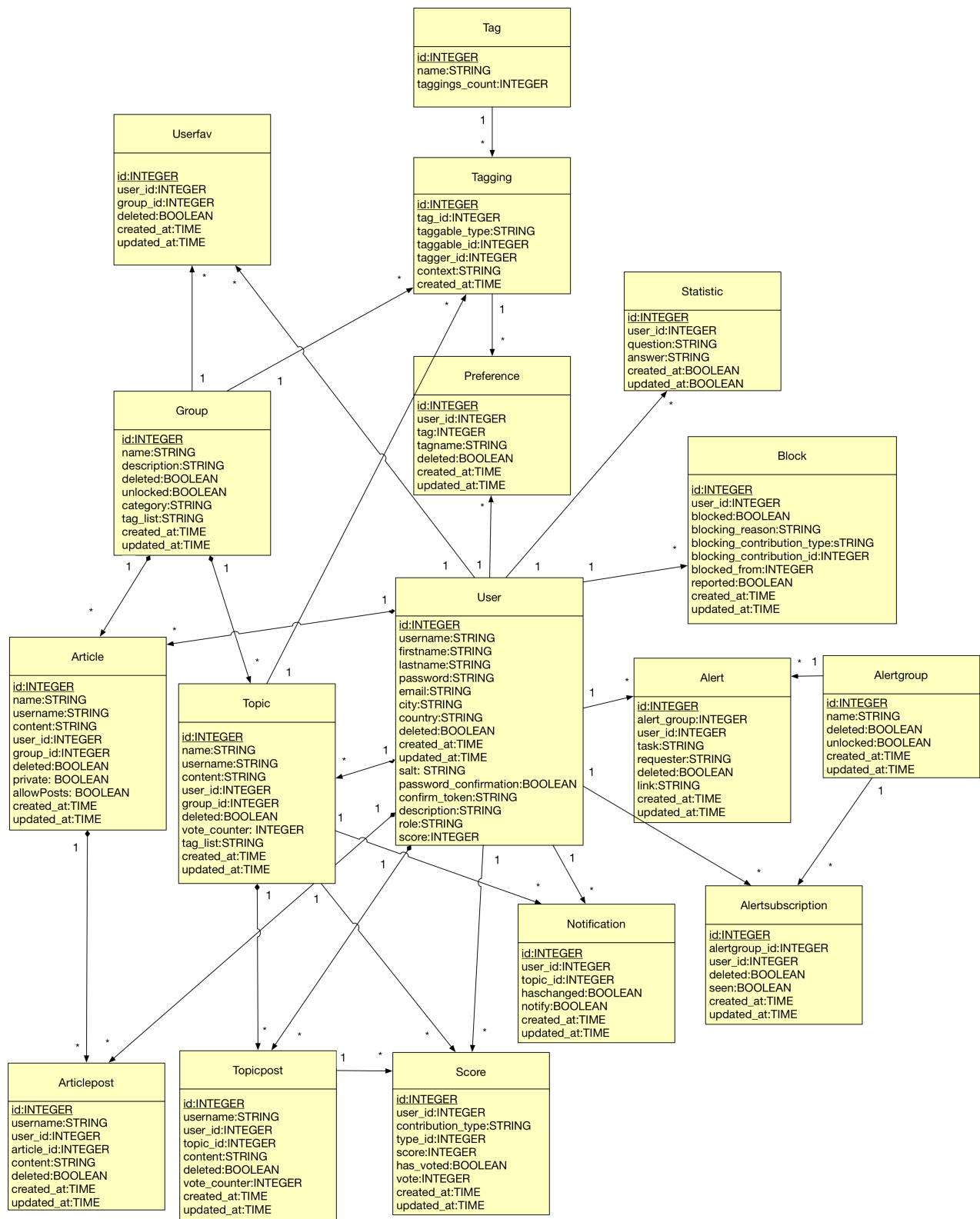


Figure 7.18: Database diagram of WorkerHub

Evaluation & Discussion

This chapter describes the evaluation concept and presents the evaluation results as well as a discussion of these results. The evaluation is performed by conducting a qualitative case study. We asked several stakeholders to test our prototype and then do a survey. Finally, we present improvements and summarize limitations.

8.1 Case Study Design

In this section we explain the design of the case study including the method, participants, the survey and setting. Also, we state our target that we want to achieve with our questions and the way we ask the questions. The interpretation of the results follows in the next section.

8.1.1 Method

WorkerHub is implemented to test the design principles and the chosen CI architecture models. Therefore, we implement a case study. A case study is an empirical method aimed at investigating contemporary phenomena in their context [65]. It has a design phase where the objectives are chosen and the case was designed [65]. Afterwards the data collection starts [65]. There the technique has to be chosen like interviews, observation etc and the data source is determined [65]. Then, the analysis phase starts [65]. We evaluated questions to every principle and wrapped them into a survey, which was our design phase. Then, we have done the data collection where we asked four leading stakeholders to test our prototype and then to answer the survey questions.

8.1.2 Participants

We focus on a representative sample of crowd workers and advocacy in our case study. We asked people from the chamber of labor in Austria and ÖGB via e-mail to evaluate

our prototype. These people are experts in the topic of crowd work and are aware of the poor working conditions. Also, we post in the forum of TurkerNation to attract crowd workers to do our study.

8.1.3 Data Collection

Our data collection is a second-degree data collection, which means that the researcher is not in direct contact with the subjects but also we directly collect raw data during the data collection [65]. We choose for our data collection the survey. The survey comprised three steps, which are presented in figure 8.1. The first step was to read the guidelines for the survey (see appendix A .1). The guidelines highlighted the goal of WorkerHub, an overview of all features and tips when to use the prototype. In the second step, the participant had to register on the platform and to get familiar with the system and its elements. The third step was the completion of the questions of the survey (see appendix B .2). For the questionnaire we used a Google Form ¹. Google Form enabled us to collect all data and get summarized results. Also, the data is there stored.

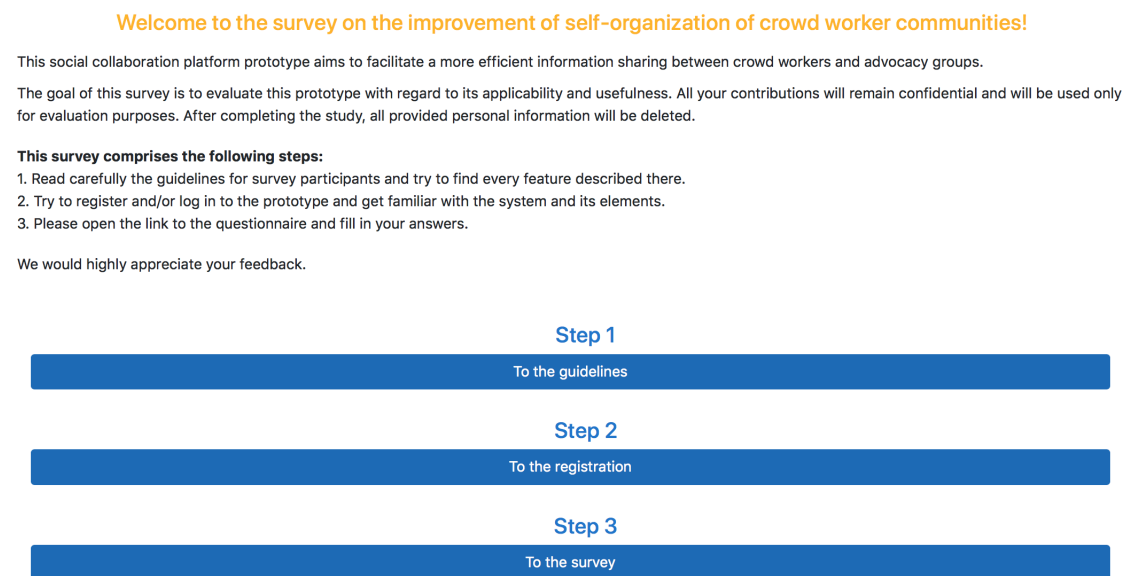


Figure 8.1: Steps of the survey

In the questionnaire, we started with demographic and general questions:

- Which role do you have?
- In which country do you live?

¹<https://www.google.com/forms/about/>. Accessed 6.10.2018

- What kind of crowd work do you typically pursue?
- How long do you pursue crowd work?

Then we asked questions regarding the usability and usefulness of WorkerHub:

- How would you define the level of difficulty to use the platform for your purposes?
- Do you think the structure of groups is useful?
- How would you define the level of discoverability to find useful and relevant information?
- Do you think that using this platform could have a benefit for involved stakeholders and improve the current crowd work situation?

In addition, we prepared questions regarding our design principles:

- Distribution of Information and Awareness:
 - Did you experience the information provided by the platform and its artifacts are sufficient and useful for your purposes and goals?
 - Is it easy to participate in a discussion or to create a new one?
- Analyzability and Transparency:
 - Do you think the “Survey” feature provided on the platform is useful? (You can do a survey on the dashboard and then take a look on the “Statistics” page)
- Trust and Reputation:
 - Do you think the “Reporting a user” and “Voting on topics and posts” features are useful? (You can go to a discussion and vote on a topic or post or report another user. If you are an external user, please check out under „Users“ the provided overview of reported and blocked users)
 - Do levels and badges of other users increase your trust in the reliability of the information they provide?
- Privacy:
 - Does the platform provide a sufficient level of anonymity? If not, where do you see areas of improvement to provide users with an increased level of anonymity?

We also asked the participants to rate the key features of WorkerHub regarding their usefulness. All these features address one or more principles, which are referred to in brackets. These features are:

- General Information Groups (Distribution of Information, Trust and Reputation)
- Specific Platform Groups (Distribution of Information, Trust and Reputation)
- Alerts (Distribution of Information)
- Profile site (Trust and Reputation)
- Survey and Statistics (Analyzability and Transparency)
- Reporting/Blocking a user (Fairness and Objective Exchange)
- Voting on topics and posts (Trust and Reputation)

As a follow up question to this rating we asked the participants "Why do you think these elements are useful / not useful?"

To ensure that the prototype covers key findings, which were identified during the literature review, we asked the participants to rate them. The key findings are:

- Workers can engage with each other and exchange information about profitable tasks, problems they face when searching for tasks, difficulties with the platform or with task requesters.
- Workers can get to know other workers.
- Workers can search for help and support.
- Workers can look for information from trustful sources.
- Workers can engage in a respectful and trustful manner but with protection of their privacy.
- Other stakeholders can obtain more information about current working conditions of crowd workers, their working life and problems they have to deal with.
- Other stakeholders can get in contact with groups or individual workers and share important information to improve their situation.

In addition, we formulated questions to check the application of our CIS architecture models:

- After the first log-in you had to specify your preferences. Did it help you to find your groups of interest easier? You can change your preferences at the dashboard under 'Choose your preferences'

Finally, we asked the participants for ideas of improvements and what functionalists they might have missed on the prototype.

8.1.4 Data Analysis

The data analysis is quantitative data analysis where we used descriptive statistics such as mean value and standard deviation [65] to find out if our principles and architecture fit the needs of our stakeholders.

8.2 Results & Discussion

This section presents the discussion and results as well as derived improvements and limitations of the prototype and of this thesis.

As mentioned above, we created questions, which aims to check and confirm the proposed design principles and our CIS architecture models as well as the usability and usefulness of the prototype.

We transferred the answers from non free text questions into a grading system, which is shown in Figure 8.2 to interpret the results using descriptive statistics: *1* relates to strongly agree/very easy/extreme benefit/very useful, *2* relates to agree/easy/yes/yes, everything works fine/little benefit/useful, *3* relates to average/uncertain/it was easy to participate, but not to start a new discussion or it was easy to start a new discussion, but not to participate/I did not specify any specifications, *4* relates to disagree/difficult/no/no, it did not worked for me/hardly any benefit/not useful, and *5* relates to strongly disagree/very difficult/not very useful/no benefits.

In the survey we asked also several free text questions. These questions will be assigned to a principle or to usability and usefulness. At the end, we will consider all results and derive our interpretation.

The first principle *Independence and Centralization* was not asked in the questionnaire, because it is mostly a technical question. WorkerHub is launched on Heroku and so it is a self-sustained and independent platform where every crowd work platform can be represented. We provide workers a central location where they can exchange information, alert each other and get in contact with advocacy.

The second principle *Privacy* is addressed by the question "Does the platform provide a sufficient level of anonymity? If not, where do you see areas of improvement to provide users with an increased level of anonymity?". 75% of the participants agreed on this question. One participant said that she wishes different settings for anonymity in the

1	↔	strongly agree very easy extreme benefit very useful
2	↔	agree easy yes yes, everything works fine little benefit useful
3	↔	average uncertain it was easy to participate, but not to start a new discussion it was easy to start a new discussion, but not to participate I did not specify any specifications
4	↔	disagree difficult no no, it did not worked for me hardly any benefit not useful
5	↔	strongly disagree very difficult not very useful no benefits

Figure 8.2: Grading system conversion

profile. In addition, we asked for a rating regarding the key finding "Workers can engage in a respectful and trustful manner but with protection of their privacy." to assess the privacy of the users. In Figure 8.3 we can see that 50% of the participants provided a very good rating and 25% rated 2 and 3, which is good and average rating. The mean of these results is 1,4 and the standard derivation is 0,9. So, we can say that on WorkerHub the privacy principle is fulfilled even though the standard deviation is relatively high. But if we include the standard deviation to our mean, which will be 2,3 then we get also a positive outcome.

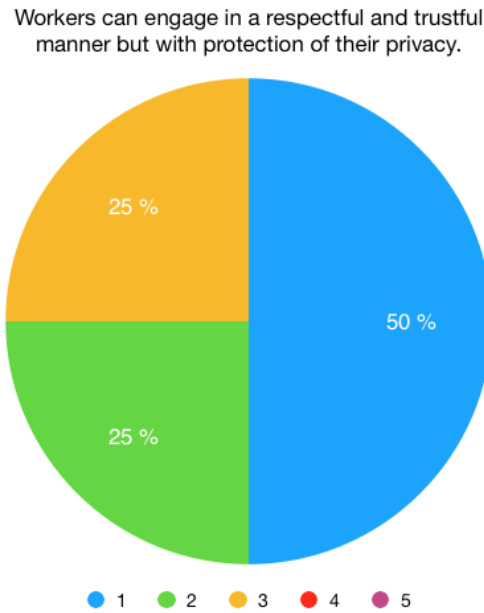


Figure 8.3: Rating of Privacy principle

The third principle is *Trust and Reputation*. With regard to this principle we asked several questions including rating on features or key findings and free text questions. In the rating part we asked these questions and findings:

- Do levels and badges of other users increase your trust in the reliability of the information they provide?
- Feature: Profile Site
- Feature: Voting on topics and posts
- Finding: Workers can get to know other workers.
- Finding: Workers can look for information from trustful sources.

There, we get this frequency distribution shown in Figure 8.4. We had in total 20 ratings and 85% are positive and the rest is average. Most of the participants rate these findings as more than satisfying and more than 40% give a rating of 1.

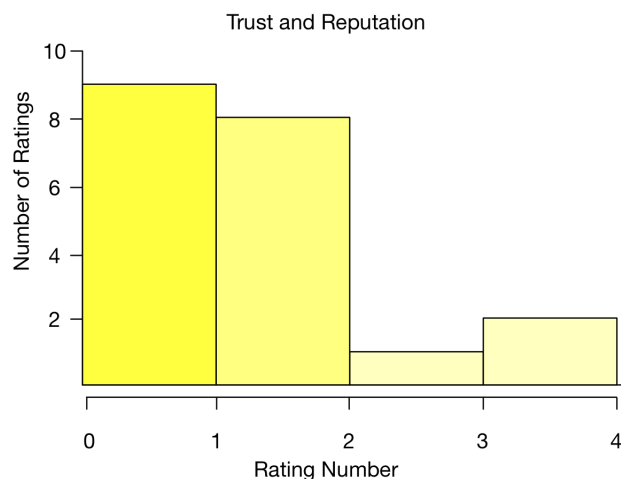


Figure 8.4: Rating of Trust and Reputation principle

Two participants said that trust and reputation is useful and it makes sense to improve trust. One participant mentioned the dependence on how much information will be provided and another one fears a confusion of the user with voting.

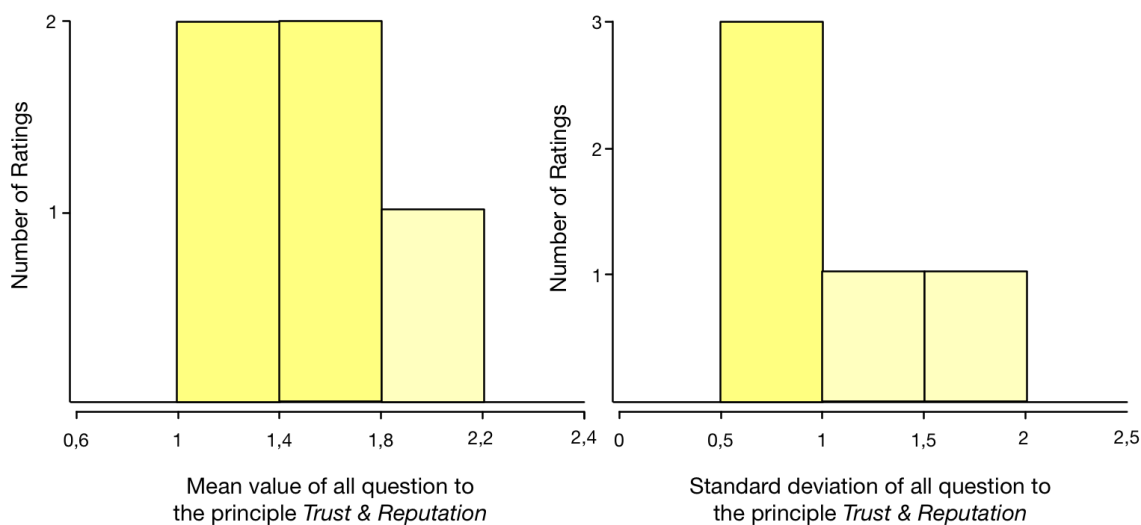


Figure 8.5: Mean and standard deviation of rating to questions of Trust and Reputation

Based on the answers we do not know why a participant rated "voting on posts and topics" as confusing. Maybe the participant assumes if a post gets a higher voting than the place of it will be changed. This is not the case on WorkerHub. The posts of a discussion are ranked according to their creation date and not the voting. As we can

see from the distribution and also the free text answers: WorkerHub meet this principle. Also, this interpretation is supported from the mean and standard deviation, which is shown in figure 8.5. The mean of each component is in the range of 0,5 and 1,25. Hence, in average the participants tend to rate the components very good.

The *Fairness and Objective Exchange* principle is only directly addressed by the rating of the feature "Reporting/Blocking a user". Furthermore, this principle is asked indirectly in one free text question. It is the same case for the principle *Trust and Reputation* where we asked about the usefulness of reporting a user. Two participants said that it is useful and the other two participants did not mention anything in context of the reporting feature. Another key finding which also occurs in the *Privacy* principle is whether users on WorkerHub can engage in a respectful and trustful manner. 50% of the participants strongly agree and 25% agree. In figure 8.6, we see the rating of the feature "Reporting/Blocking a user". The opinions of the participants are splitted. Half of them rated this feature as good to very good and the other half only thinks it is average. One participant warns of the misuse of the reporting feature. Also, this illustrates the mean of 1,8 and the standard deviation of 0,943.

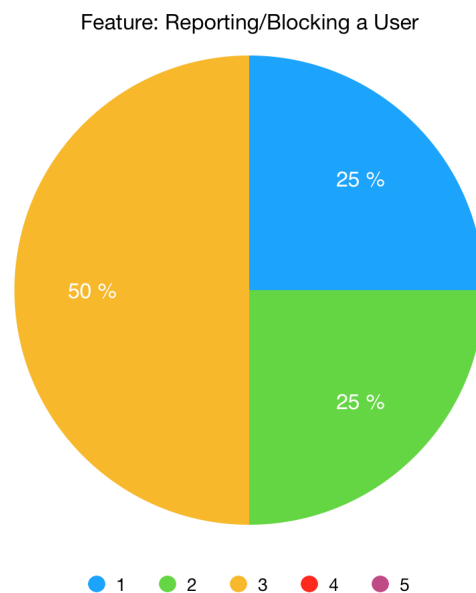


Figure 8.6: Rating of Fairness and Objective Exchange principle

There were also some comments from the participant with regard to unclear usability and also in connection with reporting. We think that they had a few problems with the implementation of our reporting system, which is not context of reporting itself. This assumption is based on the average rating of the reporting feature. On WorkerHub, there is an overview about the post, which is reported and the associated topic. The reported user will not be blocked immediately. Therefore, an administrator, external or moderator

is responsible. According to the answers of the participants, the principle itself is existent on WorkerHub and makes sense to improve language culture on platforms. But here is some more potential to test this again with a better implementation of the reporting feature.

Also, we asked many questions to the *Distribution of Information and Awareness* principle, which is one of our most important principles. Therefore, we have asked these questions, features and key findings:

- Did you experience the information provided by the platform and its artifacts are sufficient and useful for your purposes and goals?
- Is it easy to participate in a discussion or to create a new one?
- Feature: General Information Groups
- Feature: Specific Platform Groups
- Feature: Alerts
- Finding: Workers can engage with each other and exchange information about profitable tasks, problems they face when searching for tasks, difficulties with the platform or with task and requesters.
- Finding: Workers can search for help and support.
- Other stakeholders can get in contact with groups or individual workers and share important information to improve their situation.

The histogram of this rating consists of eight questions/features/findings, which are listed above and each of them has four different ratings. Overall, this makes 32 ratings in the histogram. In Figure 8.7, we can see that more than 80% rated all the questions as good and very good. Furthermore, in Figure 8.8 where the mean and standard deviation is shown, we see a small standard deviation, which means that the representativity of the mean is good. In addition, we asked the participants why they think that these features are useful or not. One participant mentioned the usefulness of the feature to exchange ideas between workers and interest groups.

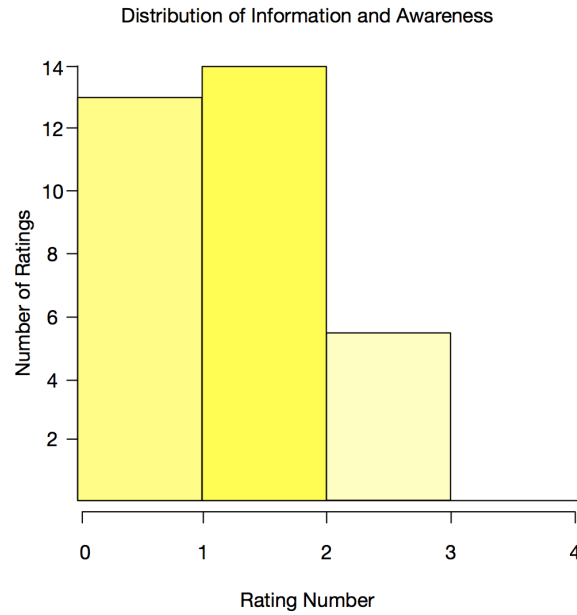


Figure 8.7: Rating of Distribution of Information and Exchange principle

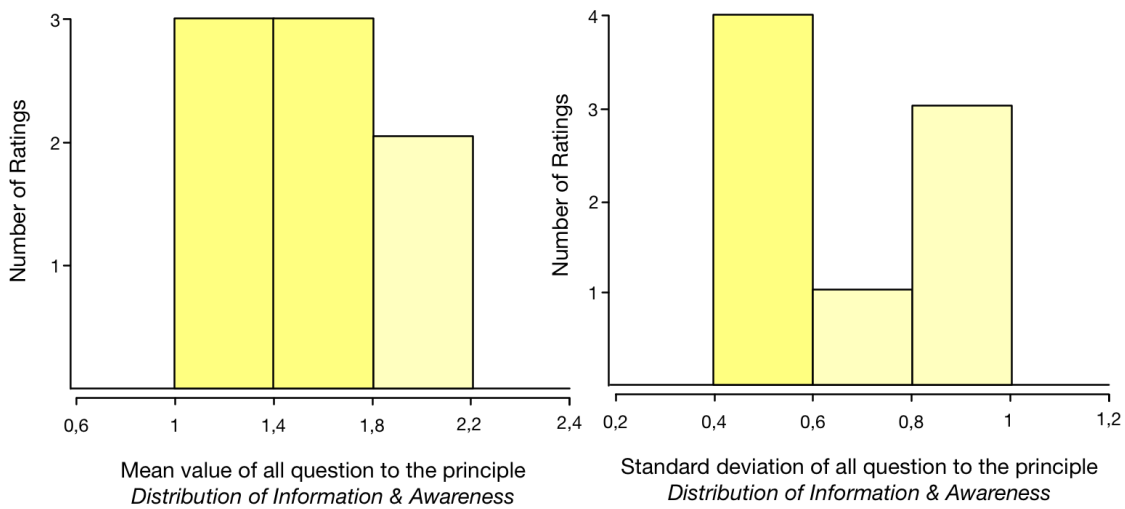


Figure 8.8: Mean and standard deviation of rating to questions of Distribution of Information and Exchange

We asked the participants if any feature needs further improvements to make it more useful. Some participants had no clue of the functionality of the alert system. This is not very surprising because our participants are members from advocacy and no crowd

workers. So they have not really a benefit of the alert system. Overall, we conclude that this feature is useful and very well implemented on WorkerHub.

The next principle is *Analyzability and Transparency*. This principle is asked through a feature, a key finding and also a free text question. We see in figure 8.9 the rating of the feature "Survey and Statistic" and of the key finding "Other stakeholders can obtain more information about current working conditions of crowd workers, their working life and problems they have to deal with." The feature is rated from all participants as very good and 75% absolutely agree with the finding. The mean of the feature is 0,8 and the standard deviation is 0,2. Because it is a very low standard deviation, we can say that representativity of the mean is good. From the key finding the mean is 1,2 and the standard deviation is 0,917. There, the representativity of the mean is low.

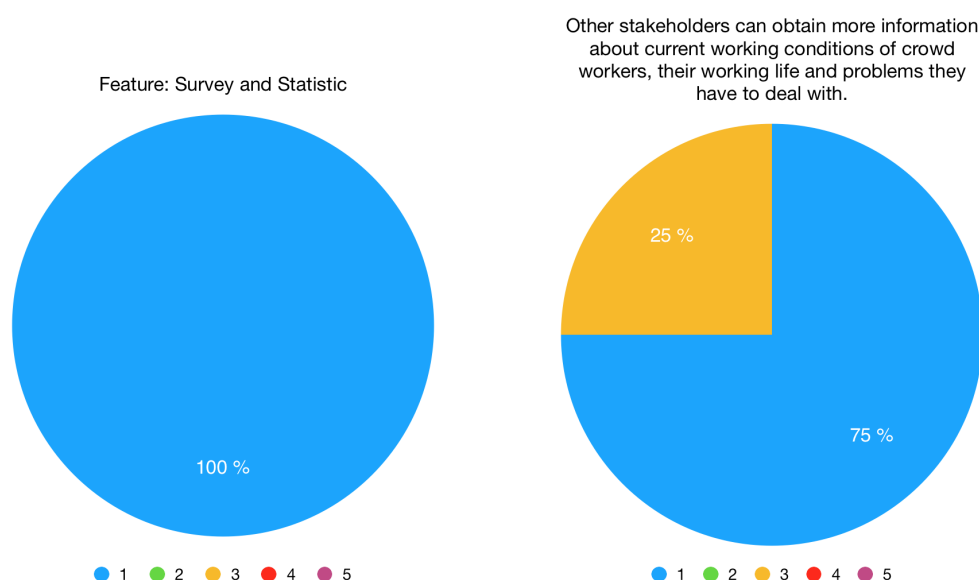


Figure 8.9: Rating of Analyzability and Transparency

Our participants had to answer this free text question "Do you think the "Survey" feature provided on the platform is useful?". 3 of 4 participants say that the feature is very useful especially for researchers and members from advocacy. The other participant rates it as very useful if a lot of crowd workers will do a survey. It is not surprising that this principle is fulfilled because it is created for members of advocacy. Therefore, it would be interesting to have the opinion of the crowd workers.

The last principle is *Work Type-Specific Software Clients*. This principle is partly fulfilled, because WorkerHub is only a web app, so we serve in the first instance only crowd workers and not mainly gig workers. For them it is better to have a native app. This point is more evaluated in the section "Improvements & Limitations".

Finally, we asked about usability and usefulness of WorkerHub. For the histogram in figure 8.10, we used 16 ratings. More than 87,5% of the rating results are between 1 and 2. Also, the mean in figure 8.11 is in average between 1 and 2. Plus, standard deviation is low which means a good representativity for the mean value.

With regard to further improvement and new features on WorkerHub, some participants wish more clarity in the user interface and for one of them the group structure requires too many clicks to come to the desired content. In the future, we will take more care on the usability of the prototyp because it is a limitation of the design. First of all, we

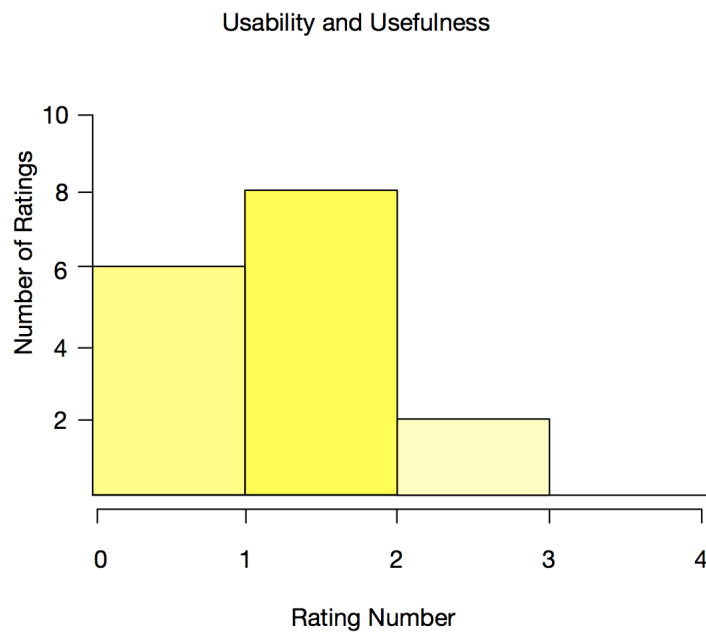


Figure 8.10: Rating of Usability and Usefulness

decided to have more structure in the groups. We know that this decision may lead to more clicks, but a better overview over all topics was more important than fewer clicks and more disorder in finding the desired content. If we have only one big forum then information will get lost, and we know from our analysis of other platforms that it is hard to find the right content. To reduce the number of clicks, we have established the concept of favorite groups. So, the user gets faster access to the desired content.

We asked one question about the CIS models, which gives us feedback if our stimulus works the right way. We asked "After the first log-in you had to specify your preferences. Did it help you to find your groups of interest easier?". We can see in figure 8.12 that 50% of the participants rated the preferences system with 1 and the other half with 3. The mean is 1,6 and standard deviation 1,077, which means a not optimal representativity for the mean value. But the participants who chose preferences were happy with this feature. So, we can assume that our stimulus worked.

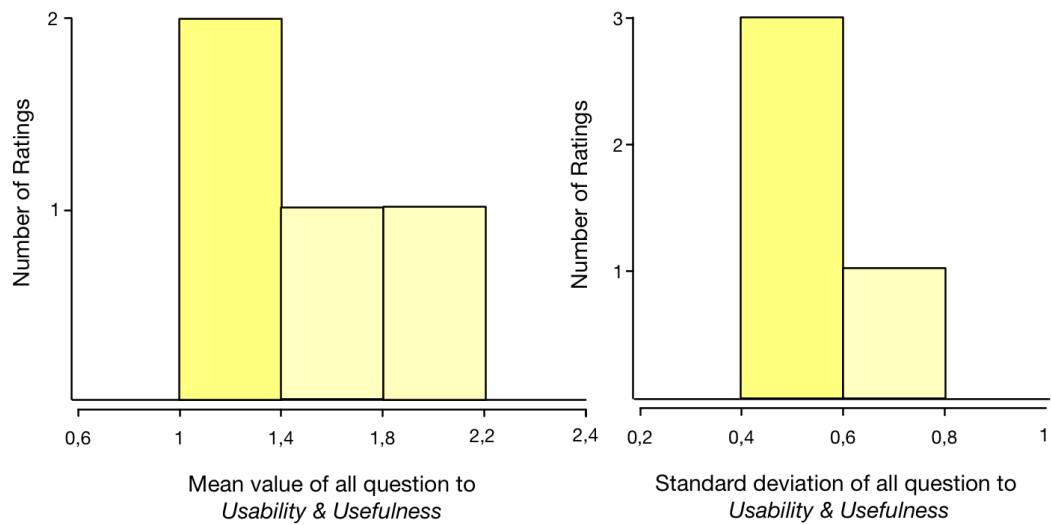


Figure 8.11: Mean and standard deviation of rating to questions of Usability and Usefulness

After the first log-in you had to specify your preferences. Does it helped you to find easier your groups of interest? You can change your preferences at the dashboard under 'Choose your preferences'

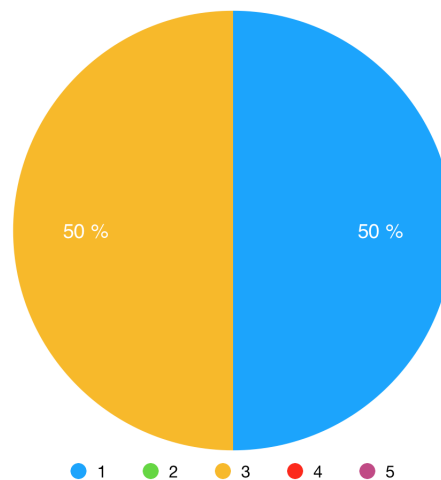


Figure 8.12: Rating of our stimulus in the CI Model

Finally, We have made a histogram about all mean values (Figure 8.13) and the relevant standard deviations (Figure 8.14) to interpret the results according to descriptive statistics. More than 72% of the mean values lie between 1 and 1,8 and more than 54% have a low standard deviation, which indicates a good representativeness of the mean values. The

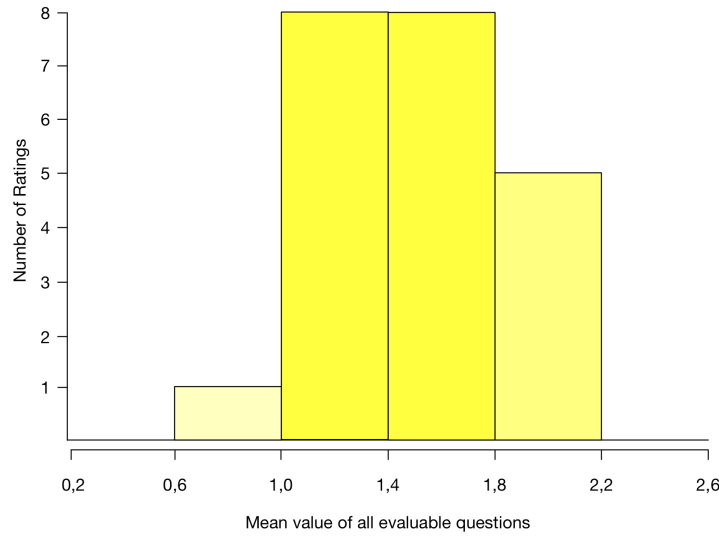


Figure 8.13: Histogram of all mean values from the survey

results confirm our assumption that with WorkerHub we built a platform to help crowd workers and advocacy to share knowledge and help them in self-organizations.

8.2.1 Improvements & Limitations

First of all, we speak about the improvements on WorkerHub, then about improvements in the design principles and limitations of WorkerHub and this thesis.

The most mentioned improvements were on the usability side. Participants wish a sorting by countries in the advocacy list and also a sorting by topics and posts. These suggestions make sense and if we further develop WorkerHub then sorting and a search function are very useful and necessary additions. Also, one participant wishes that a user can choose their level of anonymity which is a good objection and should be realized.

A few comments addressed the structure of features and the group structure. WorkerHub is limited to the countries: Germany, Austria and Switzerland. We had to make a choice about the scope of the thesis and the prototype. So we decided not to make it international because then we would need a bigger structure, which also means more implementation effort. In the future, WorkerHub should definitely be extended and include more countries, because workers are more likely to engage with others, when they are from the same country/region, and also gain more collective power. Furthermore, a structure based on countries makes sense to eliminate the load of content. Also, a space where all countries come together for the collective power would be useful. Another comment addressing the structure is that it is useful to have nested posts under a topic to have more clarity.

Some participants had problems with the reporting system. Unfortunately, they did not

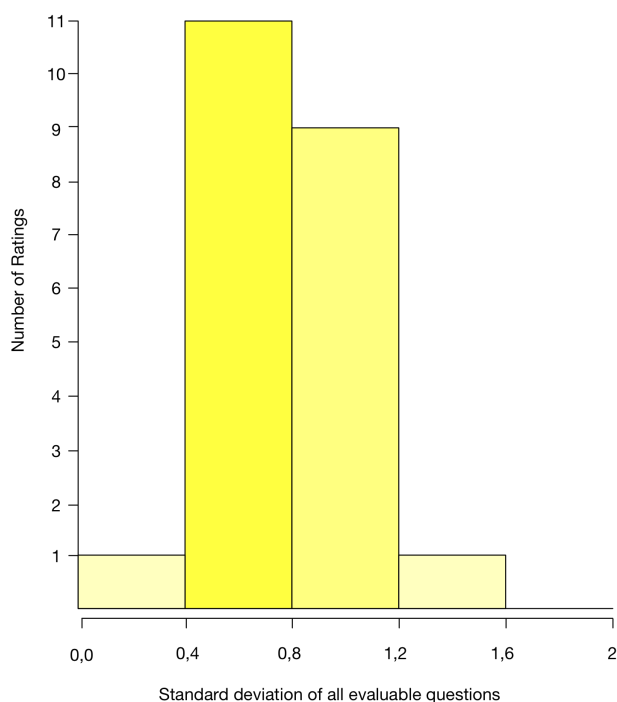


Figure 8.14: Histogram of all standard deviations from the survey

explain the reasons, but we plan to extract the reporting from the user's view and move it to a place on the side bar. Also, we plan to show which user has reported another user and save it in the statistics. If a user is only on the platform to report meaningful and useful content then the user can be blocked.

One participant suggested that WorkerHub needs to provide guidance to the main functionality at the first log-in. Our goal was to implement a prototype which is self-explanatory, but some of our participants had troubles to get along with WorkerHub. Therefore, the next step is to think about a small introduction, which highlights the main features.

The last note for further improvements were about a chat or messenger system. It was a conscious choice not to have such a system, because we wanted to avoid losing information or information that is locked up for a few people, which is the case on most forums and social media groups. A design principle is transparency and this value is followed by WorkerHub.

The results show very well a fit of the design principles according to the needs of crowd worker and advocacy. Though, we see two important improvements of our design principles. The first one is that a participant suggested a description about oneself in the profile. This suggestion would improve trust on the platform but also trust in a more general context. Therefore, we want to suggest in our design principles a self-description

to increase trust and reputation. This element is also very useful for advocacy because they can indicate their field of competence.

Another improvement to consider is an element of incentives for crowd workers to do a survey. The survey feature should ensure a regular collection of data of the working conditions of crowd workers. One points out the importance of the survey. One participant pointed out that this feature is only valuable when it is used by lot of crowd workers. This confirms our assumption to create even higher incentives for the crowd workers. In the case of WorkerHub, a user can get additional points when doing a survey or another badge. This information must be highlighted and promoted on the platform. It is important to ensure that many users will do a survey and to push more the intrinsic motivation.

There are also some limitations of this thesis. The research field is quite big and so are the possibilities to help crowd workers. We had to focus on crowd workers originating from the countries Germany, Austria, and Switzerland and also on advocacies. WorkerHub is web application which supports mostly crowd workers and not gig workers. Gig workers are location dependent and thus most of the time they are working with their phone. For them it is easier to have a native app than a web application. Furthermore, our prototype involves crowd workers, interest groups and regulators but not platform operators and requesters. One reason for this decision was the scope. If we include requesters then we also need a different structure, which increases design, implementation and evaluation effort. Moreover, there is already a system on the market where workers can get to know requesters better. In addition, it is possible to talk about requesters on WorkerHub, but they do not have an extra area. If WorkerHub is extended then it is to think about a certain role for requesters to improve the relationship between requesters and crowd workers.

Unfortunately, we did not make a big user study. For a master thesis there is no budget to pay workers to do a survey and crowd workers do nothing for free. Also, we want to attract more researchers and members of an advocacy to be involved in the survey and in the process to launch WorkerHub.

Conclusions & Future Work

The motivation for this thesis was to improve the bad working conditions for crowd workers and the need of a reliable system to cover all needs of crowd workers. During our research we identified that there is also a huge lack of information on the side of regulators and interest groups. Both parties are concerned with the current situation and would like to make changes and improvements, but they need more information. In the following we summarize this work, draw conclusions and give an outlook on future work.

The initial start point for this thesis was the poor working conditions of the crowd workers compared with the helplessness of advocacy to get in contact with crowd workers and help them. We pointed out the working life of a crowd worker and the disadvantages like the 80 hours work week for full-time compensation, unfair algorithm and rating on the crowd working platforms and the mistrust and irresponsibility on such systems. Additionally, crowd workers seek for help and so they founded forums, but now there are so many forums that they lost the overview. Advocacy groups has also no access to such forums and so, crowd workers and advocacy look for a central solution where they can get in contact with each other and share the knowledge. Therefore, we assumed at the beginning of our thesis that a central solution will solve their problem based on the following problems:

- On crowd working platforms for platform operators, requesters and crowd workers: No possibility to exchange information and get in contact with each other.
- In Forums and Chats for crowd workers: Too many channels available and no trustworthy information.
- On other information platforms for advocacy: No possibility to get in contact with crowd workers and collect information.

This resulted in our goal that is to build a central platform, which overcomes all identified problems. Therefore, we formulated three research questions:

Our first research questions was "*What are relevant requirements, capabilities and business processes of software systems that support crowd workers and worker organizations in knowledge sharing and self-organization?*". To answer this question, we had made a literature research where we found out that

- crowd workers are not a crowd of anonymous workers but also want to interact with each other,
- there are many studies who record the bad working situation of crowd workers
- there is a call for action on the politic side to help crowd worker
- a few researchers had implemented a solution and their findings and mistakes
- there is a need to worker organizations
- a collective intelligence system will fit the needs of crowd worker.

Then, we elicit the concerns and requirements of our stakeholders. Crowd workers want a safe space where they can interact with each other in trustful and respectful manner. Also, they want to simplify their work by knowledge sharing and by information from trustful resources. Platform operators do not want to change anything because they have the whole power. Requesters want a better relationship to workers. Interest groups and regulators want more or less the same. They want to collect information in order to have a foundation to change something in the political system. They want more transparency and more involvement with crowd workers.

The second research question was "*What are major conceptual software architecture design principles of social-collaborative platforms for crowd workers?*". This question builds up on the first research question. With the findings of the first research question, we designed our design principles. The design principles give every software developer guidance of the most important parts of a platform, which supports crowd worker in knowledge sharing and self-organization. The seven design principles are independence & centralization, privacy, trust & reputation, fairness & objective exchange, distribution of information & awareness, analyzability & transparency and work type-specific software clients.

Another part of this research question was to model our collective intelligence system (CIS). We choose a CIS because of the stigmergy which provides an indirect communication among individuals [43]. Therefore, we choose the CIS-AF. This framework gives us guidance to describe the key elements of our CIS and systematically model it. We looked at three different viewpoints: CI Context Viewpoint, CI Technical Realization

Viewpoint and CI Operation Viewpoint. In the first viewpoint we evaluated the current as-is work flow, the stigmergic coordination model and future to-be work flow. The second viewpoint illuminates the main artifacts, which are in our case the groups and the inputs and outputs. The third viewpoint carries out the initial data, how the transformations kicks start the feedback loop and also some metrics.

The last research question was *"How do the identified principles support the design of collaborative crowd worker platforms?"*. Therefore, we build our prototype WorkerHub where we implement our design principles according to our defined CI model. Then, we made a survey with four members of advocacy. First, the participants had to get to know the prototype better and then filled out our survey. The result was that the defined design principles and CI model support the design of collaborative crowd worker platforms. Two principles need additional improvement: Trust and Reputation will be extended with a description of the user at the profile and Analyzability will be extended to an extra intrinsic motivation for crowd workers. Also, the prototype needs further improvement on the usability side.

Overall, we conclude from the results of our survey that we found major design principles and models to design a platform for crowd workers and advocacy to support them in knowledge sharing and self-organization. With our solution we provided the users a central platform where they can get all the information they needed that include all relevant information for the daily working life of a crowd worker up to knowledge about specific platforms. Additionally, this information is accessible for all users in the system. Also, the list of advocacy members and the possibility that they can take part in conversation and post articles enables trustworthy and reliable contact to the crowd workers. The special benefit of the survey and statistic makes it possible that advocacy members are informed about the working life of the crowd workers at any time and they can collect them to do something on a political level.

With the knowledge of the design principles and the CI architecture model, it is possible for other developers or software engineers to build a platform, which helps the crowd workers and advocacy and in knowledge sharing and self-organization. Furthermore, the successful implementation of our prototype and the very good ratings for them, confirm that a central platform for the crowd workers and advocacy is very much needed.

A next step should be to conduct a larger case study with involvement of crowd workers to test the design principles and the architecture also from their side. Furthermore, it is necessary to improve the prototype with regard to better usability and involvement of more countries. Another improvement could be to find a mechanism or principle to involve requesters in a way that they get informed about the bad working conditions and also to engage with crowd workers. When some requesters see that there are other crowd work platforms where the working conditions are better, then maybe they will switch to them.

List of Figures

1.1	Origin of Crowd Work [1]	1
1.2	Examples of typical microtasks [3].	2
1.3	General process of the crowd work procedure with the involvement of the requester and the sorting of the results of the crowd workers by the algorithm of the respective crowd work platform.	4
1.4	Current state of knowledge sharing and self-organization of crowd workers	7
1.5	Target state of knowledge sharing and self-organization of crowd workers	9
2.1	The communication network among Amazon Mechanical Turk Workers [24].	15
2.2	Interaction between workers to generate high-quality outcomes [37]. . . .	16
2.3	The review process of the guild [26]	16
2.4	Overview of working platforms [3]	17
2.5	Five clusters and the analyzed platforms [18].	19
2.6	Voting system in Dynamo.	21
2.7	A screenshot of Turkopticon [16].	22
2.8	Elements of collective intelligence building "genes" [47].	25
2.9	Meta-Models for the Context Viewpoint [43].	26
2.10	Meta-Models for the Technical Realization Viewpoint [43].	27
2.11	Meta-Models for the Operation Viewpoint [43].	28
3.1	Overview over the research challenges	33
3.2	Overview over the research goal	34
3.3	Steps of the planned methodological approach	35
4.1	Overview of the key stakeholders and their relationships in the domain of crowd work	38
4.2	Communication channels between crowd workers	39
4.3	Snippet of the social area on TurkerNation	40
4.4	Model of the three laterality of a crowd work platform [3]	43
4.5	Overview of the principles of code of conduct from IG Metall	46
5.1	Overview of all characteristics of the design principle <i>Independence and Centralization</i>	52
5.2	Overview of all characteristics of the design principle <i>Privacy</i>	53

5.3	Overview of all characteristics of the design principle <i>Trust and Reputation</i>	55
5.4	Overview of all characteristics of the design principle <i>Fairness and Objective Exchange</i>	56
5.5	Overview of all characteristics of the design principle <i>Distribution of Information and Awareness</i>	57
5.6	Overview of all characteristics of the design principle <i>Analyzability and Transparency</i>	58
5.7	Overview of all characteristics of the design principle <i>Work Type-Specific Software Clients</i>	60
5.8	Report a user on TurkerNation	62
5.9	Forum structure of TurkerNation	62
5.10	A poll on TurkerNation	63
5.11	Profile of a user of Reddit HitsWorhtTurkingFor	64
5.12	Profile of a user at CloudMeBaby	65
6.1	As-Is workflow of current communication between crowd workers.	71
6.2	As-Is workflow of current communication between advocacy members.	72
6.3	Stigmeric coordination model	77
6.4	Activities of advocacy in to-be workflow	78
6.5	Activities of crowd workers in to-be workflow	79
7.1	Terms of agreement on WorkerHub	98
7.2	Registration form for crowd workers	99
7.3	Code snippet from the encryption process	99
7.4	Interface of handling requested groups	100
7.5	Reporting reasons	102
7.6	Interface of handling reporting a user	102
7.7	Interface of the preference system	103
7.8	Dashboard on WorkerHub	104
7.9	Detailed view of the section Specific Platform Groups	104
7.10	Group in detail	105
7.11	A discussion in detail	106
7.12	List of new notifications	107
7.13	Alert in detail	107
7.14	List of all members of advocacy	108
7.15	A question of the survey	109
7.16	Interface of the statistic site	109
7.17	Interface of another user's profile	110
7.18	Database diagram of WorkerHub	113
8.1	Steps of the survey	116
8.2	Grading system conversion	120
8.3	Rating of Privacy principle	121
8.4	Rating of Trust and Reputation principle	122

8.5	Mean and standard deviation of rating to questions of Trust and Reputation	122
8.6	Rating of Fairness and Objective Exchange principle	123
8.7	Rating of Distribution of Information and Exchange principle	125
8.8	Mean and standard deviation of rating to questions of Distribution of Information and Exchange	125
8.9	Rating of Analyzability and Transparency	126
8.10	Rating of Usability and Usefulness	127
8.11	Mean and standard deviation of rating to questions of Usability and Usefulness	128
8.12	Rating of our stimulus in the CI Model	128
8.13	Histogram of all mean values from the survey	129
8.14	Histogram of all standard deviations from the survey	130

List of Tables

5.1	Comparison of the analyzed platforms according to the design principles .	66
6.1	Overview of crowd worker activities in as-is workflow	74
6.2	Limitations in the current as-is workflow	75
6.3	List of improvements provided by the to-be workflow	80
6.4	Artifact Content	82
6.5	List of agents	83
6.6	List of agents	84
6.7	List of Agent Activities	86
6.8	List of artifact input/output flows	92
6.9	List of dissemination rules	93
6.10	Sources for initial CI artifact content and needed transformations	94
6.11	List of metrics	95
7.1	Overview over design principles and their features on WorkerHub	111

Bibliography

- [1] J. Howe, “The Rise of Crowdsourcing,” Wired Magazine, vol. 14, no. 06, pp. 1–5, 2006. [Online]. Available: http://www.clickadvisor.com/downloads/Howe{}_The{}_Rise{}_of{}_Crowdsourcing.pdf
- [2] E. Simperl, “How to Use Crowdsourcing Effectively: Guidelines and Examples,” LIBER Quarterly, vol. 25, no. 1, pp. 18–39, 2015.
- [3] F. A. Schmidt, “Arbeitsmärkte in der Plattform-ökonomie–Zur Funktionsweise und den Herausforderungen von Crowdwork und Gigwork,” Tech. Rep., 2016. [Online]. Available: <http://library.fes.de/pdf-files/wiso/12826.pdf>
- [4] S. T. Iqbal and B. P. Bailey, “Effects of intelligent notification management on users and their tasks,” in Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08, 2008, p. 93. [Online]. Available: <http://portal.acm.org/citation.cfm?doid=1357054.1357070>
- [5] W. S. Lasecki, J. M. Rzeszutarski, A. Marcus, and J. P. Bigham, “The Effects of Sequence and Delay on Crowd Work,” in Proceedings of the 33rd Annual {ACM} Conference on Human Factors in Computing Systems, {CHI} 2015, Seoul, Republic of Korea, April 18-23, 2015, 2015, pp. 1375–1378. [Online]. Available: <http://doi.acm.org/10.1145/2702123.2702594>
- [6] J. Teevan, “Productivity Decomposed : Getting Big Things Done with Little Microtasks,” in CHI Extended Abstracts on Human Factors in Computing Systems, 2016, pp. 3500–3507.
- [7] J. Surowiecki, The Wisdom of Crowds. Anchor, 2005.
- [8] J. Cheng, J. Teevan, S. T. Iqbal, and M. S. Bernstein, “Break It Down: A Comparison of Macro- and Microtasks,” in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15, 2015, pp. 4061–4064. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2702123.2702146>
- [9] A. Alkhatib, M. S. Bernstein, and M. Levi, “Examining Crowd Work and Gig Work Through The Historical Lens of Piecework,” in Chi 2017, 2017.

- [10] S. Kairam and J. Heer, “Parting Crowds: Characterizing Divergent Interpretations in Crowdsourced Annotation Tasks,” in Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16, 2016, pp. 1635–1646. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2818048.2820016>
- [11] M. S. Bernstein, G. Little, R. C. Miller, B. Hartmann, M. S. Ackerman, D. R. Karger, D. Crowell, and K. Panovich, “Soylent: a word processor with a crowd inside,” in Proceedings of the 23rd annual ACM symposium on User interface software and technology, 2010, pp. 313–322.
- [12] J. Kim, S. Sterman, A. A. B. Cohen, and M. S. Bernstein, “Mechanical Novel: Crowdsourcing Complex Work through Reflection and Revision,” 2016, pp. 233–245. [Online]. Available: <http://arxiv.org/abs/1611.02682>{%}0A<http://dx.doi.org/10.1145/2998181.2998196>
- [13] P. Kinnaird, L. Dabbish, and S. Kiesler, “Workflow transparency in a microtask marketplace,” in Proceedings of the 17th ACM international conference on Supporting group work - GROUP '12, 2012, p. 281. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2389176.2389219>
- [14] M. K. Lee, D. Kusbit, E. Metsky, and L. Dabbish, “Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers,” Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems, vol. 1, pp. 1603–1612, 2015. [Online]. Available: <http://dx.doi.org/10.1145/2702123.2702548>
- [15] N. Salehi, L. C. Irani, M. S. Bernstein, A. Alkhatib, E. Ogbe, K. Milland, and Clickhappier, “We Are Dynamo: Overcoming Stalling and Friction in Collective Action for Crowd Workers,” in Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems, vol. 1, 2015, pp. 1621–1630. [Online]. Available: <http://dx.doi.org/10.1145/2702123.2702508>
- [16] L. C. Irani and M. S. Silberman, “Turkopticon: Interrupting worker invisibility in amazon mechanical turk,” in Proceedings of the SIGCHI Conference on ..., 2013, pp. 611–620. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2470742>
- [17] J. Agell, “Why are small firms different? Managers’ views,” Scandinavian Journal of Economics, vol. 106, no. 3, pp. 437–452, 2004.
- [18] J. M. Leimeister, D. Durward, and S. Zogaj, “Crowd Worker in Deutschland,” Tech. Rep. 323, 2016. [Online]. Available: http://www.boeckler.de/pdf/p{__}study{__}hbs{__}323.pdf
- [19] L. Irani, “The cultural work of microwork,” New Media & Society, vol. 17, no. 5, pp. 720–739, 2015. [Online]. Available: <https://doi.org/10.1177/1461444813511926>

-
- [20] a. Kittur, J. Nickerson, and M. Bernstein, “The Future of Crowd Work,” in Proc. CSCW '13, 2013, pp. 1–17. [Online]. Available: [http://papers.ssrn.com/sol3/papers.cfm?abstract{ }id=2190946{ }5Cnpapers2://publication/uuid/AE6BF263-1DEF-4900-8C95-DC8BAD2DE4AF](http://papers.ssrn.com/sol3/papers.cfm?abstract={ }id=2190946{ }5Cnpapers2://publication/uuid/AE6BF263-1DEF-4900-8C95-DC8BAD2DE4AF)
 - [21] T. Mitra, C. J. Hutto, and E. Gilbert, “Comparing Person- and Process-centric Strategies for Obtaining Quality Data on Amazon Mechanical Turk,” in Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems, vol. 1, 2015, pp. 1345–1354. [Online]. Available: <http://dx.doi.org/10.1145/2702123.2702553>
 - [22] A. Aneesh, “Global Labor: Algocratic Modes of Organization*,” Sociological Theory, vol. 27, no. 4, pp. 347–370, 2009. [Online]. Available: <http://dx.doi.org/10.1111/j.1467-9558.2009.01352.x>
 - [23] J. Danaher, “The Threat of Algocracy: Reality, Resistance and Accommodation,” Philosophy and Technology, vol. 29, no. 3, pp. 245–268, 2016. [Online]. Available: <http://dx.doi.org/10.1007/s13347-015-0211-1>
 - [24] M. Yin, M. L. Gray, S. Suri, and J. W. Vaughan, “The Communication Network Within the Crowd,” in Proceedings of the 25th International Conference on World Wide Web - WWW '16, 2016, pp. 1293–1303. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2872427.2883036>
 - [25] J. J. Horton, L. N. Stern, and J. M. Golden, “Reputation inflation : Evidence from an online labor market,” in Working paper, 2015, pp. 1–31.
 - [26] M. E. Whiting, D. Gamage, S. S. Gaikwad, A. Gilbee, S. Goyal, A. Ballav, D. Majeti, N. Chhibber, A. Richmond-Fuller, F. Vargus, T. S. Sarma, V. Chandrakanthan, T. Moura, M. H. Salih, G. B. T. Kalejaiye, A. Ginzberg, C. A. Mullings, Y. Dayan, K. Milland, H. Orefice, J. Regino, S. Parsi, K. Mainali, V. Sehgal, S. Matin, A. Sinha, R. Vaish, and M. S. Bernstein, “Crowd Guilds: Worker-led Reputation and Feedback on Crowdsourcing Platforms,” 2016. [Online]. Available: <http://arxiv.org/abs/1611.01572{ }0Ahttp://dx.doi.org/10.1145/2998181.2998234>
 - [27] I. Metall, “Grundsätze für bezahltes Crowdsourcing / Crowdworking Leitfaden für eine gewinnbringende und faire Ziel und Zweck,” Tech. Rep., 2017.
 - [28] M. L. Gray, S. Suri, S. S. Ali, and D. Kulkarni, “The Crowd is a Collaborative Network,” in Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing - CSCW '16, 2016, pp. 134–147. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2818048.2819942>
 - [29] C. Forde, M. Stuart, S. Joyce, L. Oliver, D. Valizade, G. Alberti, K. Hardy, V. Trappmann, C. Umney, C. Carson, C. Centre for Employment Relations Innovation and Change, and University of Leeds, “The Social Protection of Workers in Platform Economy,” Brussels, p. 128, 2017.

- [30] U. Huws, S. Joyce, and University of Hertfordshire, “Österreichs crowdworkszene,” Tech. Rep. September, 2016.
- [31] D. L. Martin Risak, Arbeit in der Gig-Economy. ÖGB Verlag, 2017.
- [32] M. Leimeister, S. Zogaj, and I. Blohm, “Crowdwork – digitale Wertschöpfung in der Wolke: Ein Überblick über die Grundlagen, die Formen und den aktuellen Forschungsstand,” Crowdwork - zurück in die Zukunft? Perspektiven digitaler Arbeit, pp. 9–41, 2014.
- [33] J. M. Leimeister, S. Zogaj, and D. Durward, “New Forms of Employment And IT - Crowdsourcing-,” in 4th Conference of the Regulating for Decent Work Network, 2015. [Online]. Available: <https://ssrn.com/abstract=2736953>
- [34] F. A. Schmidt, “The good, the bad and the ugly: Why crowdsourcing needs ethics,” in Proceedings - 2013 IEEE 3rd International Conference on Cloud and Green Computing, CGC 2013 and 2013 IEEE 3rd International Conference on Social Computing and Its Applications, SCA 2013, 2013, pp. 531–535.
- [35] “Frankfurt Paper on Platform-Based Work,” Tech. Rep. December 6, 2016.
- [36] A. Felstinerf, “Working the Crowd : Employment and Labor Law in the Crowdsourcing Industry,” Berkeley Journal of Employment & Labor Law, vol. 32, no. 1, pp. 143–204, 2011. [Online]. Available: <http://search.ebscohost.com/login.aspx?direct=true{&}db=bth{&}AN=67233022{&}site=ehost-live>
- [37] S.-w. Huang and W.-t. Fu, “Don ’ t Hide in the Crowd ! Increasing Social Transparency Between Peer Workers Improves Crowdsourcing Outcomes,” in CHI 2013: Changing Perspectives, 2013, pp. 621–630.
- [38] A. Al-Ani and S. Stumpp, “Motivationen und Durchsetzung von Interessen auf kommerziellen Plattformen. Ergebnisse einer Umfrage unter Kreativ- und IT-Crowdworkern,” pp. 1–45, 2015.
- [39] D. Martin, B. V. Hanrahan, J. O’Neill, and N. Gupta, “Being a Turker,” in Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & #38; Social Computing, ser. CSCW ’14. New York, NY, USA: ACM, 2014, pp. 224–235. [Online]. Available: <http://doi.acm.org/10.1145/2531602.2531663>
- [40] D. Vakharia and M. Lease, “Beyond AMT: An Analysis of Crowd Work Platforms,” 2013. [Online]. Available: <http://arxiv.org/abs/1310.1672>
- [41] —, “Beyond mechanical turk : An analysis of paid crowd work platforms,” in iConference Proceedings, no. March 2015, 2014, pp. 1–17. [Online]. Available: <http://goo.gl/kvGHUx>
- [42] T. Gruber, “Collective Knowledge Systems: Where the Social Web meets the Semantic Web,” in Semantic Web and Web 2.0, vol. 6, no. 1, 2008, pp. 4–13.

-
- [43] J. Musil, A. Musil, D. Weyns, and S. Biffl, "An Architecture Framework for Collective Intelligence Systems," in Proc. of the 12th Working IEEE/IFIP Conference on Software Architecture (WICSA'15). IEEE, 2015, pp. 21–30. [Online]. Available: <http://dx.doi.org/10.1109/WICSA.2015.30>
 - [44] J. P. Bigham, M. S. Bernstein, and E. Adar, "Human-Computer Interaction and Collective Intelligence," The handbook of collective intelligence, 2014.
 - [45] J. Hollan, E. Hutchins, and D. Kirsh, "Distributed cognition: toward a new foundation for human-computer interaction research," in ACM Trans. Comput.-Hum. Interact., vol. 7, no. 2, 2000, pp. 174–196.
 - [46] M. Ackerman, "The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility," Human-Computer Interaction, vol. 15, no. 2, pp. 179–203, 2000. [Online]. Available: http://www.informaworld.com/openurl?genre=article&doi=10.1207/S15327051HCI1523{}_5&magic=crossref%7C%7CD404A21C5BB053405B1A640AFFD44AE3
 - [47] T. W. Malone, R. Laubacher, and C. Dellarocas, "Harnessing crowds : Mapping the genome of collective intelligence," in MIT Sloan School of Management, vol. 1, 2009, pp. 1–20.
 - [48] E. Bonabeau, "Decisions 2 .0 : The Power of Collective Intelligence," MIT Sloan Management Review, vol. 50, no. 2, pp. 45–52, 2009.
 - [49] T. Malone and M. Bernstein, Handbook of Collective Intelligence, ser. HANDBOOK OF COLLECTIVE INTELLI. MIT Press, 2015. [Online]. Available: <https://books.google.at/books?id=Px3iCgAAQBAJ>
 - [50] J. Musil, A. Musil, D. Winkler, and S. Biffl, "A first account on stigmergic information systems and their impact on platform development," in Proceedings of the WICSA/ECSCA 2012 Companion Volume on - WICSA/ECSCA '12, 2012, p. 69. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2361999.2362010>
 - [51] J. Musil, A. Musil, and S. Biffl, "Elements of software ecosystem early-stage design for collective intelligence systems," in 2013 1st International Workshop on Software Ecosystem Architectures, WEA 2013 - Proceedings, 2013, pp. 21–25. [Online]. Available: <http://dx.doi.org/10.1145/2501585.2501590>
 - [52] R. J. Wieringa, Design Science Methodology for Information Systems and Software Engineering, 1st ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014. [Online]. Available: <https://www.springer.com/gp/book/9783662438381http://link.springer.com/10.1007/978-3-662-43839-8>
 - [53] International Organization Of Standardization, "ISO/IEC/IEEE 42010:2011 - Systems and software engineering – Architecture description," ISO/IEC/IEEE 42010:2011E Revision of ISO/IEC 42010:2007 and IEEE Std 1471:2000, vol. 2011, no. March,

- pp. 1–46, 2011. [Online]. Available: [http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp={&}arnumber=6129467{&}contentType=Standards{&}sortType=asc{&}p{&}Sequence{&}filter=AND\(p{&}Publication{&}Number:6129465\)](http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp={&}arnumber=6129467{&}contentType=Standards{&}sortType=asc{&}p{&}Sequence{&}filter=AND(p{&}Publication{&}Number:6129465))
- [54] P. Runeson, M. Host, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering: Guidelines and Examples*, 1st ed. Wiley Publishing, 2012.
- [55] A. Josang, R. Ismail, and C. Boyd, “A survey of trust and reputation systems for online service provision,” in *Decision Support Systems*, vol. 43, no. 2, 2007, pp. 618–644. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S0167923605000849>
- [56] S. Billett, *Learning in the workplace: strategies for effective practice*. Sydney: Allen and Unwin, 2001.
- [57] H. Lee and B. Choi, “Knowledge Management Enablers, Processes, and Organizational Performance: An Integrative View and Empirical Examination,” *Journal of Management Information Systems*, vol. 20, no. 1, pp. 179–228, 2003. [Online]. Available: <https://doi.org/10.1080/07421222.2003.11045756>
- [58] R. T. Sparrowe, R. C. Liden, S. J. Wayne, and M. L. Kraimer, “Social networks and the performance of individuals and groups,” *Academy of Management Journal*, vol. 44, no. 2, pp. 316–325, 2001.
- [59] M. Dontcheva, R. R. Morris, J. R. Brandt, and E. M. Gerber, “Combining crowdsourcing and learning to improve engagement and performance,” in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, 2014, pp. 3379–3388. [Online]. Available: <http://dl.acm.org/citation.cfm?doid=2556288.2557217>
- [60] S. Pearson, “Taking Account of Privacy when Designing Cloud Computing Services 2 . Why is it important to take privacy into,” in *... Challenges of Cloud Computing*, 2009. CLOUD’ ..., 2009, pp. 44–52. [Online]. Available: <http://ieeexplore.ieee.org/xpls/abs{&}all.jsp?arnumber=5071532>
- [61] P. Kruchten, R. Capilla, and J. C. Dueñas, “The Decision View’s Role in Software Architecture Practice,” *IEEE Software*, vol. 26, no. 2, pp. 36–42, mar 2009. [Online]. Available: <http://ieeexplore.ieee.org/document/4786950/>
- [62] J. Hamari, J. Koivisto, and H. Sarsa, “Does gamification work? - A literature review of empirical studies on gamification,” in *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2014, pp. 3025–3034.
- [63] T. Weninger, X. A. Zhu, and J. Han, “An Exploration of Discussion Threads in Social News Sites : A Case Study of the Reddit Community,” in *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, vol. 579, no. 2, 2013.

- [64] K. Malvoni and J. Knezovic, “Are Your Passwords Safe: Energy-Efficient Bcrypt Cracking with Low-Cost Parallel Hardware,” in Proceedings of the 8th USENIX Workshop on Offensive Technologies - WOOT '14, no. Algorithm 1, 2014, p. 10. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2671293.2671303>
- [65] P. Runeson and M. Höst, “Guidelines for conducting and reporting case study research in software engineering,” in Empirical Software Engineering, vol. 14, no. 2, 2009, pp. 131–164.

.1 Appendix A

The prototypical platform *WorkerHub* aims to address the following identified needs:

- Workers like to engage with each other and exchange information about profitable tasks or problems they have with finding tasks, the platform or problem definitions.
- Workers inform each other when a lucrative task is online.
- Workers want to know how other workers do and what they can learn from each other.
- Workers want information from trustful sources.
- Workers want to engage in a respectful and trustful manner but with protection of their privacy.
- Stakeholders want to know about the working conditions of crowd workers. They want to have more information about working live of the worker and how they are treated.

In the following we summarize the main components of the *WorkerHub* platform:

- **Group structure.** The prototype has 2 different groups. (1) group with general information about crowd work (*General Information Groups*), and (2) groups related to specific platforms (*Specific Platform Groups*). Each group comprises *Articles* which are provided by advocacy groups and a *Discussion Board* where workers can interact with each other. In addition, users can vote on topics and posts.
- **Alerts.** Each platform has an alert to which you can subscribe to get informed when a new task is online. You can also setup a new alert if you want to share and inform about an identified lucrative task. Furthermore, a history of previous alerts is provided.
- **Advocacy.** A list of all members of advocacy groups who are providing information on the platform.
- **Statistics.** Each member can regularly do a survey about the current situation. Here you can view the results of the survey. A member will be noticed twice in a month to do a survey.
- **Profile.** The level of a user and her badges are shown there. The profile should provide information about the reputation of a member.
- **Dashboard.** The dashboard shows favorite groups, alert subscriptions, recommended groups and links to all notifications, alerts and the survey. A user gets a notification if a new alert pops up or another member added a new post relate to a post of the user.

Without registration many areas and contents of the platform are locked. A user who is not registered can only see articles which are explicitly public and the advocacy list. All other contents are not available.

There are different roles implemented on the platform. The normal user, moderator, admin, and extern. Moderators are experienced workers who should help by reporting users in case of disruptions and misuse. Externs are people from official institutions or worker organizations and this role is more visible than others, because *WorkerHub* is a place where workers get trustworthy information from reliable sources. They can see a list of users and also manage the reporting and blocking of users. In addition, they can create groups and articles without a request.

Please register a new account and be careful during the registration process. **As a member of an advocacy group, you need to choose another register form during the registration process.** Please

read the terms of conditions carefully. As an alternative you can also log in with the username *apollo* and password *123456*.

.1. Appendix A

Please get familiar with the prototype and its provided features. After your first login you will be asked to choose your preferences where you can choose whatever you like. Then you will be forwarded to your dashboard. If you are logged in with the mentioned username, you are automatically subscribed to an alert system.

If you feel ready, please fill in your answers in the questionnaire.

.2 Appendix B

Survey on Improvement of Self-Organization of Crowd Work Communities

Thank you for participating in this survey.

This survey investigates the applicability and usefulness of the WorkerHub platform prototype. Before you start to fill in your answers, please read carefully the user guidelines and get familiar with the prototype.

Answering the following questions should take not longer than 15 minutes. All your answers will remain confidential and will be used only for evaluation purposes. After completing the study, all your provided personal information will be deleted.

We highly appreciate your feedback

* Erforderlich

1. Which role do you have?*

- ☐ Worker
- ☐ Member of an advocacy group (trade union, policy maker, activist)
- ☐ Researcher

2. In which country do you live?*

3. What kind of crowd work do you typically pursue?*

- ☐ Cloud work (Amazon Mechanical Turk, CrowdFlower, etc.)
- ☐ Gig work (Foodora, Uber, MyHammer, etc.)
- ☐ None

4. How long do you pursue crowd work?*

- ☐ Less than 1 year
- ☐ 1-3 years

☐ More than 3 years

5. How would you define the level of difficulty to use the platform for your purposes?*

- ☐ Very Easy
- ☐ Easy
- ☐ Average
- ☐ Difficult
- ☐ Very Difficult

6. Do you think the structure of groups is useful?*

- ☐ Yes
- ☐ No

7. After the first log-in you had to specify your preferences. Does it helped you to find easier your groups of interest? You can change your preferences at the dashboard under ,Choose your preferences'?*

- ☐ Yes
- ☐ No
- ☐ I did not specify any specifications

8. How would you define the level of discoverability to find useful and relevant information? *

- ☐ Very Easy
- ¹⁵⁴☐ Easy
- ☐ Average

- ☐ Difficult
- ☐ Very Difficult

9. Did you experience the information provided by the platform and its artifacts are sufficient and useful for your purposes and goals? *

- ☐ Very Easy
- ☐ Easy
- ☐ Average
- ☐ Difficult
- ☐ Very Difficult

10. Is it easy to participate in a discussion or to create a new one? *

- ☐ Yes, everything worked fine.
- ☐ It was easy to participate, but not to start a new discussion
- ☐ It was easy to start a new discussion, but not to participate
- ☐ No, it didn't worked for me

11. Do you think the "Survey" feature provided on the platform is useful? (You can do a survey on the dashboard and then take a look on the "Statistics" page) *

- 12. Do you think the “Reporting a user” and “Voting on topics and posts” features are useful? (You can go to a discussion and vote on a topic or post or report another user. If you are an external user, please check out under „Users“ the provided overview of reported and blocked users) ***

- 13. Does the platform provide a sufficient level of anonymity? If not, where do you see areas of improvement to provide users with an increased level of anonymity? ***

- 14. Do levels and badges of other users increase your trust in the reliability of the information they provide? ***

- ☐ Yes
☐ No

- 15. Do you think that using this platform could have a benefit for involved stakeholders and improve the current crowd work situation? ***

- ☐ Extreme benefits
☐ Little benefits
☐ Hardly benefits
☐ No benefits

16. Please rate the following platform elements regarding their usefulness. *

	Very Useful	Useful	Average	Not useful	Not very useful
General Information Groups					
Specific Platform Groups					
Alerts					
Profil Site					
Survey and Statistic					
Reporting/Blocking a user					
Voting on topics and posts					

17. Why do you think these elements are useful / not useful? *

18. Please rate the following statements based on your opinion of the platform. *

	Strongly Agree	Agree	Uncertain/ not applicable	Disagree	Strongly Disagree
Workers can engage with each other and exchange information about profitable tasks, problems they face when searching for tasks, difficulties with the platform or with task requesters.					
Workers can get to know other workers.					
Workers can search for help and support.					

BIBLIOGRAPHY					
Workers can look for information from trustful sources.					
Workers can engage in a respectful and trustful manner but with protection of their privacy.					
Other stakeholders can obtain more information about current working conditions of crowd workers, their working life and problems they have to deal with.					
Other stakeholders can get in contact with groups or individual workers and share important information to improve their situation.					

19. Which features require further improvement to make them more useful*

20. What functionalities do you miss in the prototype? *
