Journal on Interactive Systems, 2024, 15:1, doi: 10.5753/jis.2024.3727

Gamification Techniques and Contribution Filtering in Crowdsourcing Micro-Task Applications

Ana Maria Amorim [**Federal University of Bahia, UFBA** | ***ana.amorim@ufba.br***]
Ailton Ribeiro [**Federal University of Bahia, UFBA** | ***ailton.ribeiro@ufba.br***]
Murilo Guerreiro Arouca [**Federal University of Bahia, UFBA** | ***murilo.guerreiro@ufba.br***]
Iago Meijon [**Federal University of Bahia, UFBA** | ***iagomeijonm@gmail.com***]
Victor Cavalheiro [**Federal University of Bahia, UFBA** | ***victor.cavalheiro17@gmail.com***]
Maria Clara Pestana [**Federal University of Bahia, UFBA** | ***mpestana@ufba.br***]
Vaninha Vieira [**Institute of Computing, Federal University of Bahia, UFBA** | ***vaninha@ufba.br***]

Institute of Computing, Federal University of Bahia, Av. Milton Santos, s/n, Ondina, Salvador, BA, 40.170-110, Brazil.

Received: 30 September 2023 • **Accepted:** 22 April 2024 • **Published:** 15 May 2024

Abstract: The rapid expansion of the internet has led to a growing worldwide user base, with Brazil alone having approximately 83% of households connected to the network, equating to around 61.8 million households. Crowdsourcing, a production model that harnesses collective wisdom for problem-solving, has gained prominence in this digital landscape. Challenges in crowdsourcing include improving people's motivation and engagement and verifying the quality of a high number of contributions. In our research, we investigated the hypothesis that using gamification techniques, including recognition badges, feedback mechanisms, and user rankings, improves users' engagement and motivation in crowdsourcing micro-tasks applications. This paper presents ConTask, a crowdsourcing micro-task application, and two studies conducted to investigate the impact of using gamification techniques and contribution filtering as motivational factors in crowdsourcing. The first was a case study comparing two versions of ConTask: the original version and a gamified one. The second was an experimental study to evaluate the developed contribution filtering mechanism. Findings suggest that the use of gamification and contribution filtering can improve user participation in crowdsourcing micro-task applications.

Keywords: Gamification, Crowdsourcing, Micro-Tasks, Content Filtering, Motivation.

1 Introduction

The Internet has been expanding its influence over the years, with the number of users growing worldwide. In Brazil, 83% of households have some type of connection to the network, representing approximately 61.8 million Brazilian households, according to the TIC Domicílios survey [CETIC.BR-NIC.BR, 2021]. Therefore, it becomes logical the increase in interest in how to harness the creative and productive capacity of this community [Brabham, 2008].

The term crowdsourcing was introduced by Howe [2006], who defined it as: "a production model and process structuring that utilizes collective wisdom and learning for problem-solving or solution development". Brabham [2008] classified crowdsourcing systems into four different types: (i) discovery and knowledge management, (ii) distribution of Human Intelligence Tasks (HIT), (iii) broadcast research, and (iv) peer-assessed creative production. In this research, we used HIT in our studies.

The crowdsourcing model allows solving various problems using the online crowd, but good performance depends on an active and willing community to contribute to the system [Morschheuser *et al.*, 2016]. Keeping users active is not always easy, as it is a natural tendency for crowdsourcing system users to become demotivated and reduce the number of contributions made. Several motivational stimulus approaches have been studied, including the use of gamification, which introduces elements common to digital games into systems, such as recognition badges, feedback utilization, and user ranking [Morschheuser *et al.*, 2016].

A particular type of crowdsourcing is known as distributed human intelligence tasking (HIT), where a task is divided into micro-tasks [Brabham, 2013]. These micro-tasks are distributed through an online platform, such as Amazon's Mechanical Turk (MTurk)[1], and people perform them in exchange for small rewards (usually a few cents). These micro-tasks are usually not simple for computers to execute, but are easy to perform by humans, such as labeling images, translating texts, transcribing texts, transcribing scanned historical documents [Brabham, 2013; Gadiraju *et al.*, 2014], and others.

One of the main challenges faced by micro-tasks crowdsourcing systems lies in the effective management of the immense volume of contributions characteristic of this collaborative approach. Essential in this context is the ability to discern the most relevant and valuable contributions to ensure the quality of the outcome. The crowdsourcing model can enhance collaboration and promote collective actions, allowing a diverse group of people (with different skills, knowledge, and experiences) to work together on projects in a flexible and accessible manner. This means they can contribute to the project in their time and according to their skills and interests. Furthermore, the crowdsourcing model is often based

[1] https://www.mturk.com/. Accessed on: January 25, 2024.

on online platforms, making it easier for people to participate in projects regardless of their geographical location or technical skills. This can help create an inclusive environment that fosters collaboration and collective action.

This work investigates the impact of gamification as a motivation factor in crowdsourcing systems and implements ranking, notification, and feedback techniques in a crowdsourcing application prototype called ConTask [Pestana and Vieira, 2018a]. Additionally, it presents a filtering approach to support the identification of relevant contributions within the application through the implementation of a contribution filtering module.

The ConTask application aims to assist in the management of large spaces by implementing Human Intelligence Tasks (HIT) that allow users to report observed issues in these spaces. HIT stands for "Distributed Human Intelligence Tasking" and is a form of crowdsourcing in which larger tasks are divided into micro-tasks that are easy for humans to perform but are typically challenging for computers. Examples of micro-tasks include labeling images, translating texts, and transcribing scanned documents [Brabham, 2013; Gadiraju *et al.*, 2014; Stol and Fitzgerald, 2014].

To evaluate the proposed solution and gather data from real users, two investigations were conducted: (i) a case study comparing the original version of the application with the gamified version, where the experiment's results indicate that the gamification techniques used can stimulate user participation in micro-task crowdsourcing systems, and (ii) an experimental study to evaluate the filtering mechanism.

This article presents an expanded and in-depth version of the previous publication [Meijon *et al.*, 2023]. While the previous paper focused only on the gamification techniques to motivate users, this contribution expands the scope of the work and presents original findings on the investigation of contribution filtering techniques in crowdsourcing systems. Further details on the ConTask application are also reported.

The remainder of this article is organized as follows: Section 2 provides research background and presents related work. Section 3 describes the methodology used in this work, including how data was collected and analyzed. The ConTask application is presented in Section 4, both the original version and the version with gamification techniques, and the filtering mechanism. Section 5 presents the studies conducted with end users. Section 6 presents the results and discussion. Finally, Section 7 conclusions and future work.

2 Background and Related Work

In this section, we address intrinsic and extrinsic motivational factors. Additionally, we explore the application of gamification as a motivational strategy to enhance contributions in crowdsourcing systems and to filter contributions, which tend to appear in high volumes in those systems.

2.1 Motivational Factors: Intrinsic and Extrinsic

Motivation is a driving force that prompts an individual to act and can be classified as intrinsic or extrinsic, depending on the reasons or objectives that underlie the action [Deci and Ryan, 2010]. In the self-determination theory, Deci and Ryan [1980] differentiates between these types of motivation. The nature of motivation can vary, even if the intensity remains the same, as in the case of a student who may be motivated to do homework out of curiosity or to gain approval [Deci and Ryan, 1985].

Several researchers are focused on motivation in crowdsourcing [Lopes *et al.*, 2024; Wu and Gong, 2021; Alam and Sun, 2023]. Lopes *et al.* [2024] analyzed the motivational factors of workers (crowdworkers) in crowdsourcing projects that offer payment for tasks and identified that intrinsic motivations are driven by personal interests and internal emotions, such as pleasure, autonomy, and reputation, while extrinsic motivations are influenced by the work context, such as learning and financial rewards.

The sustained participation of crowdworkers is crucial for the success and sustainability of the online crowdsourcing community. However, this issue has not received adequate attention in the research community [Wu and Gong, 2021]. In Alam and Sun [2023], the results suggest that the motivation of crowdsourcing participants is shaped by an evolving combination of three basic components (i.e., contextual condition, outcome, and intensity of action) and mediated by two types of system usage practice (i.e., passive and active). Passive usage practices facilitate motivation sustainability from initiation to progression, while active usage practices play a fundamental role in sustaining it. Understanding the motivational factors and sustained participation of contributors in crowdsourcing projects is crucial for the effectiveness and longevity of online communities.

2.2 Gamification Techniques

The use of various gamification techniques in an attempt to generate greater interest in the utilization of crowdsourcing systems was the subject of study in [Feng *et al.*, 2018; Congcong Yang and Feng, 2021; Feng *et al.*, 2022; Bastanfard *et al.*, 2023]. In Feng *et al.* [2018], the authors aimed to strengthen intrinsic motivations such as self-presentation, self-efficacy, social bonds, and fun through gamification artifacts like point rewards and the provision of feedback. According to the authors, a scoring system can promote competition and enhance participants' self-efficacy. Meanwhile, the use of feedback increases users' sense of usefulness, making them feel valued and recognized by the platform. The study used data from 295 solvers on a large Chinese crowdsourcing platform[2] Results demonstrated that self-presentation, self-efficacy, and enjoyment are positively affected through gamification techniques and feedback.

Subsequently, the study by Feng *et al.* [2022] provides guidelines for collaborative knowledge crowdsourcing platforms on how to motivate solvers. Gamification mechanics related to immersion, such as avatars and customized interfaces [Meliande *et al.*, 2024; Ribeiro *et al.*, 2024], enhance solvers' self-esteem and stimulate their knowledge contribution. Gamification elements related to performance (e.g., points, badges, and performance feedback) should be

[2] Available at: `http://zbj.com`. Accessed on: January 24, 2024.

provided to strengthen solvers' beliefs in the value of their knowledge [Arouca *et al.*, 2024], which, in turn, will encourage their voluntary knowledge contribution. In the work by Bastanfard *et al.* [2023], gamification of data collection was presented as a solution to review, categorize images, and motivate people. The proposed method considers various challenges, such as motivation, financial costs, and delays.

Both intrinsic and extrinsic motivations positively affect solvers' participation in crowdsourcing [Congcong Yang and Feng, 2021]. Amorim and Vieira [2023], propose a motivation model aimed at elderly contributors to crowdsourcing initiatives. The authors argue that the elderly are motivated by dynamic tasks that stimulate their curiosity, increase their knowledge and skills, help them take care of their physical and mental health, and by altruistic tasks. Some demotivating factors include irrelevant or time-consuming microtasks.

The increasing use of smartphones worldwide has led to the creation of various systems that employ crowdsourcing in a mobile environment. Considering this new setting, Chi and colleagues conducted a study on the design of smartphone applications that encourage users to complete micro-tasks in everyday contexts [Chi *et al.*, 2018]. They utilized the mobile app Crowdsource[3], developed by Google and known as Collaborative Contribution. The objective of the study was to assess how gamification, recognition, curiosity, and fun influenced user participation in the system. The results suggest that gamification and recognition are essential in motivating users.

In their research, Kobayashi *et al.* [2015] established methods to keep users motivated in a crowdsourcing system focused on social contributions, such as digitizing a collection in a public library. The authors applied gamification techniques like ranking and feedback. In addition to individual feedback, the authors also implemented a community feedback technique, where the portal's page displayed the number of tasks completed by the community. These techniques were effective in keeping users engaged in the system.

Walter *et al.* [2022] assessed the impact of gamification in a graphical interface with a progress bar element. Several triangles were displayed, and the participant's task was to label the triangles according to their type. As the task was performed, the progress bar showed the percentage of labeled triangles. After completing all the tasks, a final score was provided, indicating the number of correctly labeled triangles. If participants were not satisfied with the result, they could repeat the work. Other gamification elements were also employed. In addition to background music, visual and auditory effects were introduced as feedback on task execution. When a triangle was labeled correctly, a pleasant tone played, and the screen turned green for a while. Conversely, if labeled incorrectly, an unpleasant tone sounded, and the screen turned red temporarily. After the work was completed, a list of the highest scores, along with participant IDs, was displayed. All the game elements used in Walter and colleagues' article are typical of games and are often utilized in the context of gamification, as described by Sailer *et al.* [2013]. A list of other possible game elements can be found in [Deterding *et al.*, 2011].

Results of the research by Walter *et al.* [2022] demonstrated that participants can be motivated to generate more data through the use of gamification. However, it was observed that this effect diminished when participants performed the tasks multiple times. The improvement in the quality of the data collected through the use of game elements could not be demonstrated. The quality of task execution with and without game elements did not differ significantly.

Unlike the related works described in this section, our research aimed to analyze qualitatively the incorporation of gamification elements (ranking, notification, and feedback) into a crowdsourcing micro-task application.

2.3 Contributions Filtering

Filtering is a technique used to select and separate elements within a specific domain. For example, searching for a science fiction documentary on a movie and series platform, sorting files on a computer by their type, or narrowing down restaurant choices based on the type of cuisine. In the micro-task crowdsourcing environment, the use of filtering is important because it is a way to select the most relevant contributions for task resolution, considering the high number of contributions expected in such systems.

In this research, we employed two distinct techniques for micro-task contributions filtering: text-based filtering, performed using an algorithm, and human-based filtering, carried out through a voting mechanism.

For text-based filtering, we chose to use the algorithm by Boyer and Moore [1977], which is an approach used to search for specific patterns within texts. The decision to use this algorithm was made because it is widely used and well-accepted in the literature [Tanalol *et al.*, 2023]. According to Cole [1994], this algorithm performs well in string searching and is used in the majority of applications. This algorithm initially aligns the target pattern "P" with the text "T". Following this, it preprocesses the alphabet of letters that may be encountered in the text "T" to avoid unnecessary character comparisons. This algorithm underwent adjustments in [Navarro and Tarhio, 2000], where the authors proposed a study to identify patterns in texts compressed using the LZ78 and LZW models. This allowed the use of the text directly in its compressed form, eliminating the need to decompress the file for text searching.

In the filtering performed by humans, we utilized majority voting, which is a voting mechanism where the "winner" is the option that receives the majority of votes. In crowdsourcing applications, this approach leverages the crowd to vote on contributions they deem more or less suitable for solving a particular micro-task [Chen *et al.*, 2020; Tao *et al.*, 2020; Chen *et al.*, 2022b]. The advantage of using this mechanism, as highlighted in Chen *et al.* [2020], lies in reducing costs by eliminating the need to hire experts to evaluate contributions and increasing crowd engagement and participation.

Current label aggregation strategies do not consider the differences in the quality of workers labeling different instances. The study by Tao *et al.* [2020] argues that a worker may exhibit varied labeling qualities across different instances. They propose four strategies that assign weights to work-

[3]Available at: `https://crowdsource.google.com/`. Accessed on: January 24, 2024.

ers when labeling different instances, using the similarity between workers' labels to estimate their specific quality in each instance. These strategies combine the worker's specific quality in each instance with their overall quality across all instances to determine their weight when labeling a specific instance. Chen *et al*. [2022b] further suggest that when label quality surpasses random classification, the performance of label integration methods tends to improve as the number of labels increases.

In this work, we explored gamification techniques for user motivation in a crowdsourcing system for micro-tasks, as well as investigated two contribution filtering techniques.

3 Research Method

The methodology of this research covers two main investigations: (i) gamification techniques, and (ii) contributions filtering. Below, we present the methodological structure adopted for each of these domains.

3.1 Gamification Techniques

Initially, a literature review was conducted to explore the basic concepts of motivation, gamification, and to identify related research in the field. Next, gamification techniques were incorporated into a prototype of a crowdsourcing application for micro-task distribution. Finally, we evaluated the prototype through a study involving eight users who used the prototype for six days. The following details the steps:

1. The initial literature review and discussion of related works allowed us to define the gamification techniques (feedback, ranking, and notifications) to be implemented. These techniques, when used together, have been identified as the most commonly utilized in many studies such as Vaughan [2018]; Law *et al*. [2016]; Park *et al*. [2018]. The search for articles was conducted on Google Scholar[4], Scopus[5], IEEE Xplore[6] and through manual searches on conference and congress websites;
2. The second stage involved the selection of the crowdsourcing system for micro-task distribution to be used in the user study and the implementation of the gamification techniques defined in the previous stage. The primary criterion for selection was having access to the source code to be able to carry out the necessary implementations.
3. In the final stage of the research, we evaluated the proposal with eight volunteers over a period of six days. For participant selection, we employed convenience sampling, a non-probabilistic sampling technique, where subjects who are most accessible and convenient are selected [Wohlin *et al*., 2012]. We contacted colleagues and former classmates (graduates) who volunteered to participate in the research. For the evaluation, we used two versions of the application: the original version and the version with the gamification techniques implemented in the previous stage. The data obtained during the use of both versions of the application were analyzed, considering the number of tasks made available, accepted, and declined. Finally, we administered a questionnaire[7] consisting of three open-ended questions to gather the volunteers' opinions on the use of the two versions of the application. Each question was related to one of the gamification techniques integrated into the application. The questions were as follows: (i) Did you feel encouraged to use ConTask at any point during the study due to the daily notification you received? (ii) During the experiment, did you have an interest in completing tasks to improve your ranking position? (iii) Did know the total number of contributions (tasks completed) and the number of pending tasks influence your desire to contribute? Subsequently, we conducted a content analysis of the responses, relating and discussing the results with those found in the literature. We analyzed the questionnaires using qualitative coding procedures based on Lazar *et al*. [2017] and Charmaz [2014]. Due to the limited number of participants, we chose not to use tools and conducted the analysis manually. We adopted an inductive process based on the researcher's interpretation, reading each participant's questionnaire, highlighting and recording relevant parts.

When inviting people to participate, we explained the research's objective and theme. Since many individuals were not familiar with the term "crowdsourcing", we used examples to illustrate the concept. We also introduced ConTask and how the application fits within the research context. In doing so, we aimed to clarify any questions people had and make the research proposal more understandable and appealing to them.

3.2 Contributions Filtering

The filtering mechanism is designed to assist in screening relevant and valid contributions from the substantial volume typically present in a crowdsourcing system. To achieve this, a literature review was conducted to search for filtering techniques, and two of them were selected and implemented in a crowdsourcing application. To evaluate the use of the filtering mechanism, an experimental study was conducted, involving data generated by participants and other data generated automatically to simulate the large volume of contributions in a crowdsourcing system.

1. In the first stage, a literature review focused on searching for filtering strategies in crowdsourcing systems was conducted. Following the analysis of related works, two filtering techniques were selected. The selected filtering techniques are as follows: (i) Keyword Filter: This filter implements the modified Boyer-Moore algorithm and is responsible for validating contributions

[4]Available at: `https://scholar.google.com.br/`. Accessed on: January 24, 2024.
[5]Available at: `https://www.scopus.com/`. Accessed on: January 24, 2024.
[6]Available at: `https://ieeexplore.ieee.org/`. Accessed on: January 24, 2024.
[7]Available at: `https://github.com/ctamp2000/questionnaires.git`. Accessed on: January 25, 2024.

made in the micro-task based on keyword matching. (ii) Majority Voting Filter: This filter utilizes the crowd to evaluate contributions through a voting mechanism. These techniques were chosen based on their suitability for the task of filtering contributions in the micro-task crowdsourcing system.

2. The second stage involved selecting the crowdsourcing system for micro-task distribution to be used in implementing the filtering techniques and the experimental study. The primary criterion for system selection was also access to the source code for the implementation.
3. In the third stage, an experiment was conducted to evaluate the filtering mechanism. Participants were organized into an experiment divided into three stages: (i) preparation, (ii) contribution and voting, and (iii) data collection and evaluation. During the preparation stage, participants selected titles, keywords, and deadlines for the tasks. They accessed the application, entered their names, and provided responses. In the contribution and voting stage, participants contributed responses and voted on the contributions of others. The experiment included both automated and non-automated contributions. Automated Contributions: comprised of automated contributions, which will be displayed randomly immediately after the block of non-automated contributions. Non-automated Contributions: comprised of contributions made by participants, which will be displayed randomly and always precede the block of automated contributions. A questionnaire was administered to task requesters to assess the results of the contributions and filters. The focus of the questionnaire was to assess whether the filters implemented in the filtering module contributed to the requester user selecting the most coherent contributions. The questionnaire consisted of five open-ended questions: (i) Did majority voting help you select coherent contributions to solve your task? (ii) Did the keyword filter aid in selecting coherent contributions to solve your task? (iii) Based on the results of the majority vote filter and keywords, which filter brought the most coherent contributions to solving your task? (iv) Are filtered contributions more coherent than unfiltered contributions? (v) What improvements do you suggest for the applied filters? The analysis technique used for the form was the same as that used for the questionnaire regarding the study of gamification techniques.

In order for applicants to be able to answer the questionnaire, we clarified what the research was about and the concepts used. Thereafter, we present the results of the contributions without filtering and with filtering. Due to the advent of the Covid-19 pandemic, all processes related to this stage were conducted remotely.

4 ConTask

In this section, we present the two versions of the application used in the user study, as well as the overall architecture of the application after the implementation of the modules developed in this work. The original version will be referred to as ConTask 1.0, and the version with the gamification techniques implemented will be identified as ConTask 2.0.

ConTask 1.0 is a prototype application developed to assist in the management of large areas, especially university campuses, their buildings, rooms, bathrooms, and common areas [Pestana and Vieira, 2018a]. Monitoring and managing the structures of large spaces are challenges for any organization; however, with the assistance of the application, it is possible to minimize some administrative issues, such as the cleanliness of bathrooms and the proper functioning of equipment. For this purpose, the contribution of users who frequent these spaces is important. The application allows the creation of tasks that request information related to a specific area. In this way, common university spaces can be evaluated by the local community.

The application was developed using the Android operating system and uses contextual information composed of entities, elements, and sources to distribute micro-tasks. Contextual information interacts directly with the application, such as user, task, and problem. Contextual elements are relevant characteristics of entities that can be used in distribution, such as the user's location, the status of the requested task, and the type of task. Sources, in turn, correspond to the origin of the data used, such as GPS and a database [Vieira *et al.*, 2011].

Using the Global Positioning System (GPS) of the device, ConTask directs tasks to users based on their context and geographical location. Therefore, if a user is in an area of the campus where there are pending tasks, the application makes these tasks available to the user. Tasks can be of many types, such as assessing the level of illumination in a location or checking if any equipment is turned on. Based on user responses, the administration of the area collects information that assists in decision-making.

As shown in Figure 1, the general architectural diagram of the application includes the gamification and filtering modules developed in this work, with the contributions stemming from this development highlighted in gray. It is also possible to observe the arrangement of gamification techniques in the screen layer.

The ConTask user interface consists of four main functional areas: (i) Home page; (ii) Login page; (iii) Task section; and (iv) Notifications area. On the Home page, three gamification techniques are highlighted, and implemented to motivate users to make contributions within the application. As shown in Figure 1, the management layer's role is to capture the user's context through mobile device sensors such as GPS, process the context, and make tasks available to the user interface layer.

To support this work, we used a new server, and developed an API to facilitate communication between the server and the ConTask application. Through this API, the application can interact with the new server, which manages all the application's data and allows for the retrieval of data generated during the experiment.

The main contributions of this work are the filtering and the gamification module. The filtering module was included to perform contribution filtering. As can be seen in Figure 1, the filtering module includes two contribution filters: (i) keyword filter and (ii) majority voting filter.

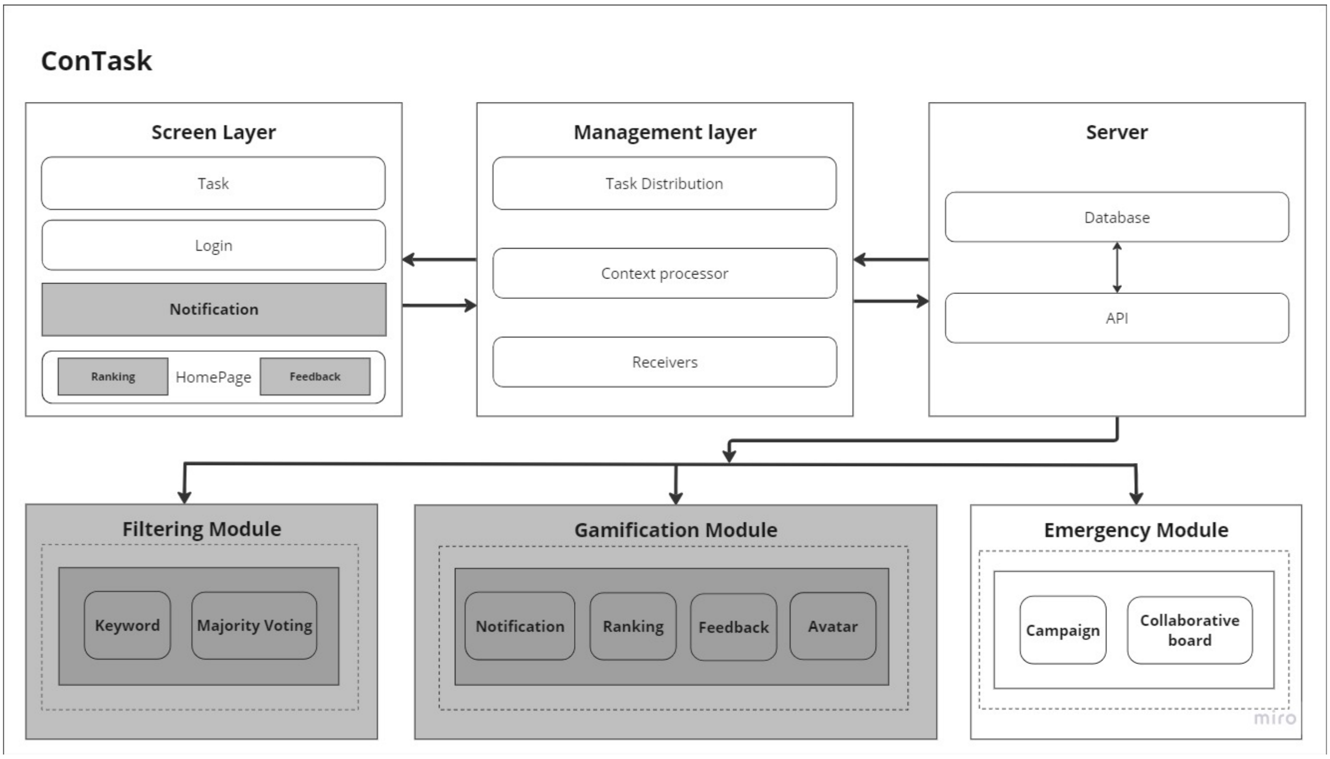

Figure 1. ConTask Main Architecture Diagram.

We are currently developing a new module in ConTask, the emergency module, where campaigns are created to provide support in disaster recovery and emergencies. The emergency module is intended to function as a solution for disaster recovery and emergencies, where contributors are individuals associated with institutions or non-governmental organizations (NGOs) engaged in humanitarian assistance to communities affected by disasters. The solution provides a pool of volunteers who are recruited by these agents to engage in humanitarian tasks. In this module, a campaign is created when a critical event occurs that requires the recruitment of volunteers. These volunteers are recruited based on their alignment with the context of the campaign. The tasks are of a humanitarian nature and aim to provide support to the affected community, such as collecting donations of material or financial resources, for example.

4.1 Gamification Techniques Implementation — ConTask 2.0

In the new version of ConTask (ConTask 2.0), three gamification techniques have been implemented: ranking, notification, and feedback. In ConTask 2.0, the ranking system was introduced on the welcome home page of the application, as illustrated in Figure 2. On this page, the system displays the user's position among other users who have also submitted their contributions and the number of tasks they need to perform to move one position in the ranking.

If the user has not yet completed any tasks, the application displays a message encouraging them to complete their first task and, from then on, be included in the user ranking. The system considers the number of contributions made for user ranking, and as a tiebreaker criterion, it considers the user who reached the required number of contributions first.

Another gamification element implemented in ConTask 2.0 was the use of notifications. Figure 3 displays the screen with the notification sent to the user's device and the welcome screen with the feedback submission. These elements were used to stimulate competitiveness and encourage the user to continue using the application, searching for new tasks. The information provided to the user as feedback includes: the overall total of tasks, the tasks available to the user, the percentage of the user's participation in the community compared to the overall total of tasks, the total number of tasks accepted, and the total number of tasks declined.

4.2 Filter Architecture in ConTask

In the current architecture of the ConTask application, the micro-task metamodel is composed of the following elements: question, answer, and context. The question represents the micro-task requested in the application. The answer represents the crowd's contribution to the proposed micro-task. The context is the desired scope for the task execution, for example: the potential users who can contribute to the micro-task and the location to which the task refers.

The filtering module operates on contributions. To implement the filtering module, it was necessary to modify the original architecture of ConTask, particularly the way data was persisted. In the new version, tasks and crowd responses are now stored in a database.

Furthermore, we have developed three new functionalities: (i) Micro-task validity: a period, in days, during which tasks can receive contributions from the crowd; (ii) Majority voting filtering: a filter that utilizes the crowd as evaluators of the contributions they propose; and (iii) Keyword filtering: a textual filter that uses keywords specified by the requester to validate contributions.

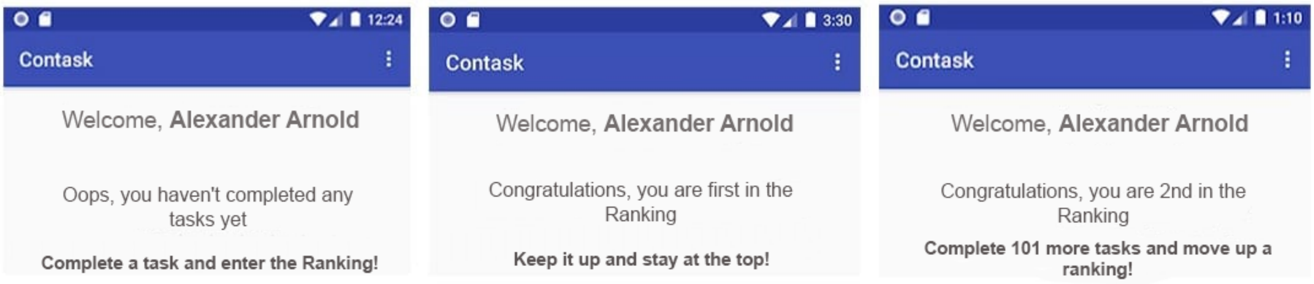

Figure 2. Home page of ConTask 2.0 with the use of the ranking mechanism.

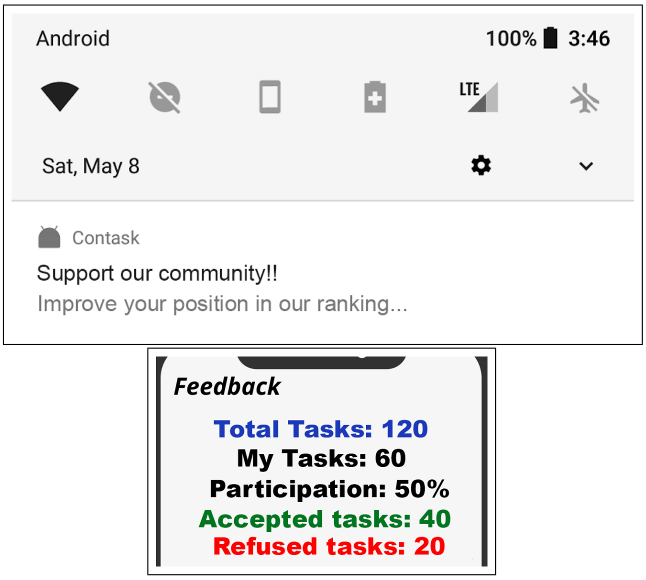

Figure 3. Notification mechanism implemented in ConTask 2.0 and the feedback screen.

The original micro-task metamodel of ConTask has been adapted as follows, as illustrated in Table 1. In addition to the pre-existing elements (question, answer, and context), we have added the elements keywords and task validity, determined by the task requester at the time of its creation. Keywords will be used in the contribution filtering process. Task validity will specify the number of days during which the crowd can contribute.

Table 1. ConTask Micro-Task Metamodel. Adapted from (Pestana and Vieira [2018b]).

Element	Description
Question	Any topic or assessment the requester wishes to be involved in
Answer	Yes/No responses, limited responses, open-ended responses, or media files
Context	Specify a limited group of users who can respond to the micro-task
Keyword	Specify keywords that contributions for this micro-task should contain
Task validity	Specify the number of days the task will be valid to receive contributions

Figure 4 shows the proposed architecture for ConTask, where we have added an API component, a database component, and a filtering module, named Contribution Filter, composed of two filters: keyword filter and majority voting filter, described as follows:

1. **API:** REST Application Programming Interface (API) is responsible for the operations related to the creation, updating, and removal of micro-tasks, contributions, and votes within ConTask.
2. **Database:** a new database has been modeled to persist the data of micro-tasks, contributions, and votes consistently, enabling operations through the API. The database used in the original version of ConTask has been migrated to this new database.
3. **Keyword Filter:** a textual filter implemented using a modified Boyer-Moore algorithm and responsible for validating contributions.
4. **Majority Voting Filter:** a filter implemented to use the crowd as evaluators of contributions.

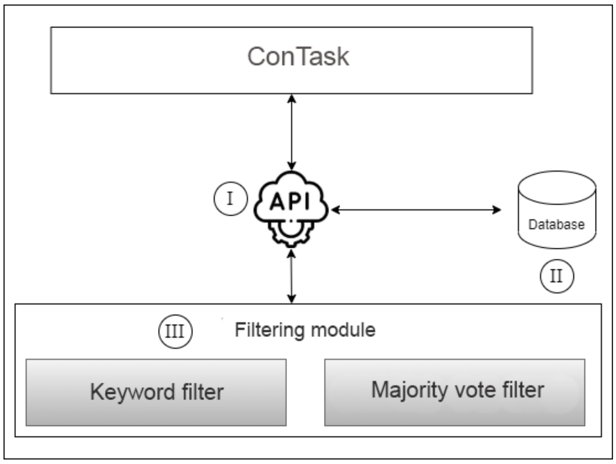

Figure 4. Proposed filtering architecture for ConTask.

Given the proposed architecture, we can define the interaction flow as follows: ConTask initiates the interaction with the API (Item I), and the API performs and persists the requested operation in the database (Item II). If the operation involves a contribution, a filtering process will be carried out (Item III).

4.3 Filtering Module Implementation

For the implementation of the API and the filtering module, the Java programming language was used. Figure 5 details the flowchart of the proposed architecture for ConTask, from task creation to the storage of the result. The requester should define the task title, the keywords to be used in the filter, and the task's validity in days. The task's validity will determine the deadline for receiving contributions. The insertion of contributions will trigger the operation of the filtering module, composed of the keyword filter and the majority voting filter.

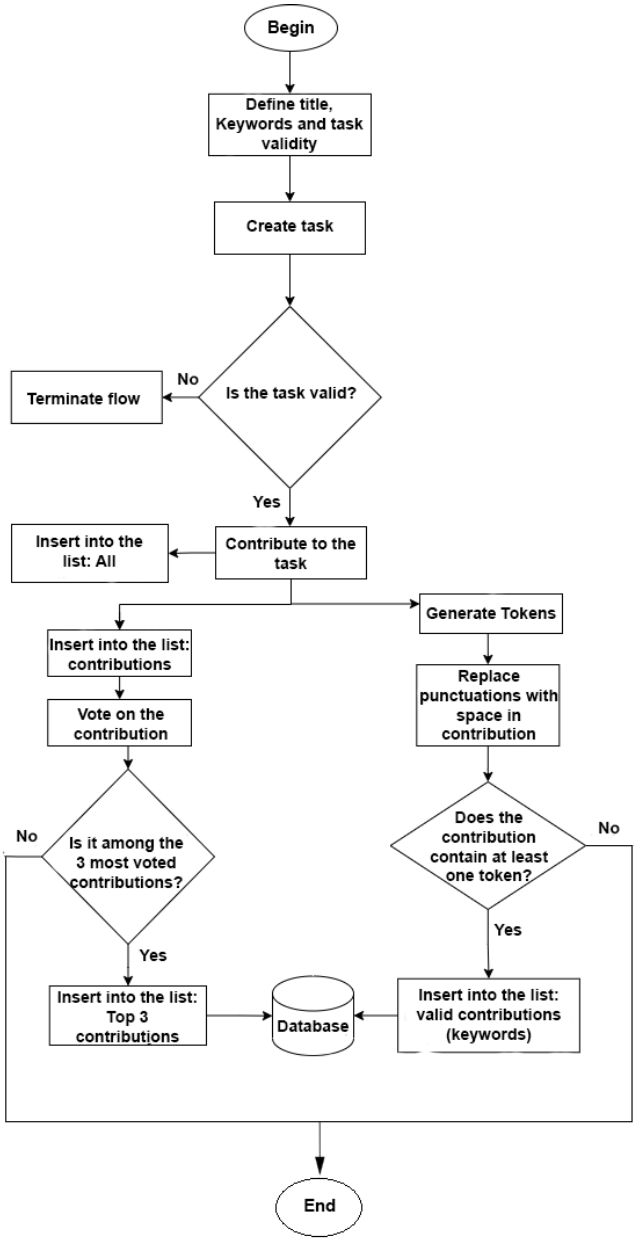

Figure 5. Complete flowchart of the contributions filtering module.

Figure 6 illustrates the architecture of the proposed API, to which the ConTask application is connected. ConTask initiates a request to the controller via the HTTP protocol, and based upon the verb enclosed within it "GET" or "PUT", the controller executes the corresponding operation, employing the pertinent model, and persists the resource within the database. This is the main component for maintaining the consistency of the data that is created, modified, and consumed.

Upon receiving a request through the API, the system will be responsible for selecting the appropriate controller for execution. The chosen controller will then leverage the HTTP verb included in the request to define the operation to be performed on the database. If the request contains the "GET" verb, it will trigger the "Retrieve" operation, which will retrieve and provide ConTask with the query results from the

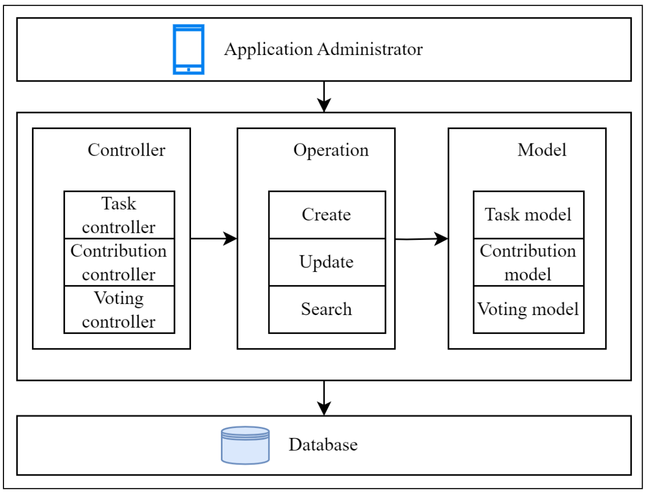

Figure 6. Proposed API architecture.

database. Conversely, if the request uses the "PUT" verb, it will activate either the "Update" or "Create" operation, which will be used to update or create information in the database.

The keyword filter encompasses the pattern recognition algorithm introduced in Boyer and Moore [1977], enabling the determination of contribution validity based on predefined keywords within the task. In this study, the proposed algorithm represents an adaptation of the Boyer-Moore algorithm, necessitated by the requirement for token generation. Tokens are automatically derived from the keywords defined by the micro-task requester, adding a space as both a prefix and a suffix. Consequently, a keyword like "house", for instance, becomes "house". Token creation is imperative to prevent the algorithm from erroneously identifying the target pattern as a suffix, prefix, or infix. For instance, the term "house" should not trigger a match when embedded within words like "marriage" or "casualty". Should the token be identified within the participant's contribution, it will be included in the "Valid Contributions" list; otherwise, it will be appended to the "All Contributions" list.

5 Experimental Studies

In this section, we present the experimental studies performed to evaluate the proposed mechanisms.

5.1 Contribution Filtering

To evaluate the contribution filtering mechanism proposed in ConTask, an experimental study was conducted. Below are the methodological details of the experiment, as well as the analysis of the results found.

Initially, a pilot study was conducted to fine-tune the process of applying the experiment and addressing any technical difficulties. This pilot study involved a 28-year-old volunteer.

Following the pilot study and corresponding adjustments to the protocol, the experiment was conducted with a group of 5 volunteers, all university students, aged between 21 and 25, including 4 males and 1 female (Table 2). The experiment lasted for 4 hours.

Table 2. Profile of Participants.

Participant	Age	Gender	University student
P1	23	Male	Yes
P2	21	Female	Yes
P3	25	Male	Yes
P4	23	Male	Yes
P5	24	Male	Yes

Due to the COVID-19 pandemic, the experiment was conducted virtually. Participants were asked to install the application through the provided APK via a WhatsApp group. After installing the application, participants entered a Google Meet room with their microphones muted, since the voice channel was only used to provide necessary instructions for the experiment.

For each task, participants were distributed as follows: one person was randomly chosen to participate as the requester (task owner), while the others were responsible for contributing and voting on contributions they deemed most suitable for solving the task (crowd role). For this experiment, three distinct tasks were created:

- **Task 1:**
  - Question: Do you think the university bus provides safety to students?
  - Validity: 1 day.
  - Keywords: dangerous, safe, and unsafe.
- **Task 2:**
  - Question: Is the street outside your house well lit?
  - Validity: 1 day.
  - Keywords: bright and dark.
- **Task 3:**
  - Do you consider the water supplied in the university's water fountains ideal for consumption?
  - Validity: 1 day.
  - Keywords: dirty, clean, and contaminated.

The experiment was split into three stages: (i) preparation; (ii) contribution and voting; (iii) collection and evaluation.

In the preparation stage, initially, the requester defined the title, keywords, and validity of the task. Then, the researcher conducting the experiment assisted the participants in logging into the application by entering their names and demonstrated the existing functionalities for contributing to a task and voting on the contribution they deemed best to answer the task.

The objective of the contribution and voting stage is for participants to contribute to the presented task and vote on the contributions of other participants. To simulate a real-life scenario that could be experienced on a university campus, a program was developed to generate 100 contributions using LeroLero[8], a tool for generating random sentences, and the NodeJS[9] framework to persist the content of the sentences in the database. Afterward, all participants were required to submit a contribution and vote on at least one contribution. Participants could not vote more than once on the same contribution. Each task received a total of 104 contributions, with 4 contributions from participants and 100 contributions generated automatically. To avoid influencing the participants' votes, contributions were displayed as follows: (i) first, the contributions from participants (non-automated) were shown randomly; (ii) then, the automated contributions were displayed, also randomly.

The collection and evaluation stage aims to gather the opinions of task requesters regarding the result of applying the filters to the selection of contributions made by participants. To collect the opinions of the requesters, a questionnaire based on Günther [2003] with 5 open-ended questions was developed using the Google Forms tool. Initially, three results of the received contributions were shown to the requester: one with the keyword filter, another with the majority voting filter, and the last one with the results of the contributions without any filtering. Then, a brief explanation of the concepts that the requesters should use to answer the questions was provided. Results are discussed in Section 6.

5.2 User Study - Gamification

The objective of this study was to assess if the gamification techniques introduced in ConTask 2.0 impacted the number of user contributions and to analyze whether users found these techniques relevant in motivating their participation in the system.

Eight participants were selected for the study, which spanned six days. The participants were randomly selected from a group of university students based on their availability and willingness to participate in the research. Table 3 presents the participants' profiles, including age range, gender, education level, experience with crowdsourcing applications, and mobile application usage. Those invited to participate in the research were informed about the concepts used, allowing them to determine their experiences with crowdsourcing and mobile technology. The participants were randomly divided into two groups of four individuals each. Each group was designated by a letter, forming Group A and Group B. The groups started with different versions of the application, with Group A using ConTask 2.0, and Group B using the original version, ConTask 1.0.

Table 3. Profile of the participants in the user study.

ID	Gender	Age	Experience *(crowdsourcing)*	Experience *(mobile)*
P1	Male	24	No	Yes
P2	Male	33	No	Yes
P3	Female	24	Yes	Yes
P4	Male	23	No	Yes
P5	Male	25	Yes	Yes
P6	Female	30	No	Yes
P7	Male	22	No	Yes
P8	Male	26	No	Yes

Figure 7 provides an overview of the created protocol. During the first three days, participants in Group A received ConTask 2.0 with gamification features, while participants in Group B received ConTask 1.0 without modifications. At

[8]Available at: `https://lerolero.com/`. Accessed on: January 24, 2024.
[9]Available at: `https://nodejs.org/en`. Accessed on: January 24, 2024.

the end of the initial three days, there was a swap of the application versions, with Group A using ConTask 1.0 and Group B using ConTask 2.0. To mitigate potential bias in the study, participants were never made aware of which version of the application they were using at any point.

The geographic context filter of the application was disabled in both versions during the study. This was necessary because it was crucial for the same task to appear for both groups of users, even if they were in different geographic contexts. Due to the COVID-19 pandemic[10], the study was conducted remotely, using virtual interaction resources. Initially, two groups were created on WhatsApp, respecting the divisions of participants into Groups A and B as established in the experiment. An installer (APK) was generated for the Android platform for each version of the application. The APK was made available in the WhatsApp chat[11] according to the version established for each group.

To ensure that the group swap had indeed taken place, a control task was created in which users were asked whether they had completed the version exchange. As soon as users switched the application version, they would respond to this control task. This way, it was possible to verify that all participants had completed the version exchange, maintaining the ratio of 4 participants for each version of the application.

Four tasks were made available each day at different times and without a predefined schedule. Once the tasks were loaded on the server, they became available to both groups simultaneously. All loaded tasks had a duration of 24 hours, so participants had this period available to complete the task. The decision to keep the task active for only a predefined time was intended to encourage participants to open the application daily and prevent all tasks from being completed only on the last day.

At the end of the six days, all participants were provided with a questionnaire containing seven questions. Three questions were about the study, aimed at understanding how the user interacted with the modifications made in ConTask 2.0 and whether these modifications had influenced their participation. The remaining four questions were intended to profile the user. The questionnaire was available for one day for the submission of responses and was developed using the Google Forms tool[12]

The questionnaire responses were meticulously examined and coded into distinct themes related to participants' interactions with notifications. Table 4 exemplifies this categorization process. Within the "Notifications" category, three main subcategories were identified: 1) notifications that led to task execution, 2) notifications that did not lead to task execution, and 3) irrelevant notifications. In the first subcategory, two codings were established: 1.1) participants who felt encouraged to use the app and 1.2) those who were reminded to use it.

Conversely, for notifications that did not result in task execution, four codings were created, including instances where the notification function was disabled (2.1), the device was in silent mode (2.2), or the participant was using the device with other apps (2.3), as well as cases where the content of the notification was not read (2.4). Finally, the "Irrelevant notification" subcategory encompassed responses that were coded as: 3.1) participants did not notice the notification or 3.2) were not available to contribute at the time it was received.

Table 4. Categorization and coding of questionnaire responses. Category: Notification.

Subcategory	Coding
1. Led to task execution	1.1 Felt encouraged to use the app 1.2 Was reminded to use the app
2. Did not lead to task execution	2.1 Notification function disabled 2.2 Silent mode 2.3 Device in use with other apps 2.4 Did not read the content of the notification
3. Irrelevant notification	3.1 Did not notice the notification 3.2 Were not available to contribute at that time

Categorizing and coding questionnaire responses regarding user interactions with notifications is crucial for comprehending mobile app user behavior. By recognizing patterns, such as notifications that efficiently trigger task execution or those frequently ignored or considered irrelevant, developers can customize notification strategies to enrich user engagement and enhance the overall user experience. These insights aid in identifying effective notification strategies and enhancing overall user engagement and experience.

6 Results and Discussion

In this section, the results and discussions of the studies conducted in this work are presented.

6.1 Contribution Filtering Results

With the questionnaire analysis, we aimed to verify whether the filters implemented in the filtering module contributed in any way to help the requester select more coherent contributions.

When asked if *majority voting helped in selecting coherent contributions to solve your task*, 2 out of the 3 requesters responded that they were aided by the filter, while 1 stated that they did not like it much because they did not agree with the contribution that reached the top position in the *"TOP 3 Contributions"* list.

Regarding whether *the keyword filter helped in selecting coherent contributions to solve your task*, all 3 requesters said they liked the idea, although they found the filter somewhat limited.

When asked *based on the results of the majority voting and keyword filters, which filter provided more coherent contributions to solve your task*, 2 participants responded that the majority voting filter yielded better results, while 1 participant preferred the keyword filter.

When asked if *filtered contributions are more coherent than contributions that were not filtered*, all the requesters

[10]Available at:`https://covid19.who.int/`. Accessed on: January 24, 2024.
[11]Available at:`https://www.whatsapp.com/`. Accessed on: January 24, 2024.
[12]Available at: `https://docs.google.com/forms`. Accessed on: January 27, 2024.

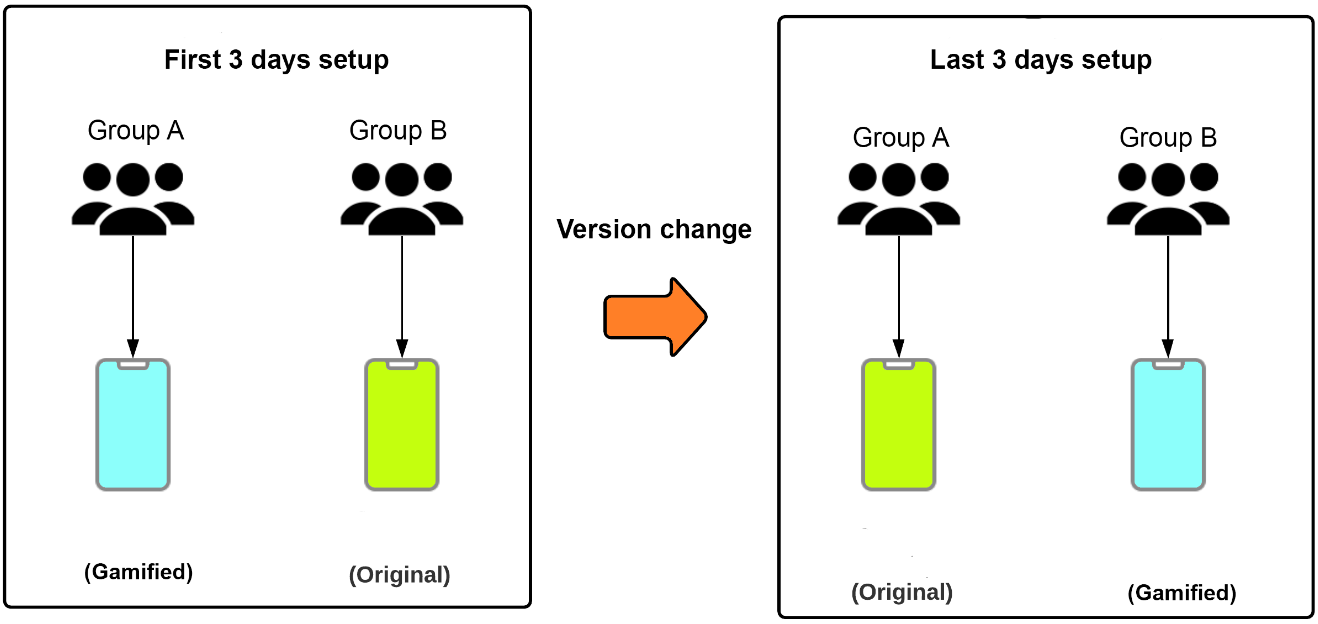

Figure 7. Overview of the protocol created with the configuration of groups x version of ConTask used in the Gamification User Study

preferred the results shown after applying at least one of the proposed filters. Finally, regarding *what improvements do you suggest for the applied filters*, the requesters provided suggestions for improvements that are documented in this text, in the section on future work.

In this study, we observed that automated contributions, even though they appeared in a block following the contributions made by participants, received multiple votes and occupied prominent positions. For instance, they took the second position in Task 1, the second position in Task 2, and the first position in Task 3. This observation can be explained by the motivation of participants to collaborate with the overall outcome of the work. In a voting scenario, items that are seen first tend to receive more votes, which could explain the automated contributions' rankings.

We observed that words that were part of the requester's keyword group, when used in plural form in participants' contributions, were not detected by the keyword filter. This limitation led the requesters to prefer the majority voting filter. However, in one of the tasks conducted in the experiment, the requester preferred the result shown by the keyword filter. The majority voting filter is subject to crowd behavior. Therefore, both filters have their strengths and limitations.

6.2 Gamification User Study

In the user study, both versions of the ConTask[13] application sent task solutions to the server. However, at the time of submission, in addition to the user's response, the application also sent information about the version used. This allowed for the collection and analysis of log data according to the version used. The experiment included a total of 24 tasks, with 4 tasks distributed each day in each group. Therefore, the maximum number of contributions expected per version of the application at the end of the experiment was 96 tasks executed.

Table 5 displays the data collected with indications of the application versions and the quantity of contributions received (micro-tasks executed) in the first 3 days of the experiment, before the version exchange between the groups. It also shows the quantity of contributions received in the last 3 days of the experiment, after the version swap between the groups, and the total tasks completed. It is important to note that the execution of the control task was not counted in the results as a completed task.

Table 5. Total tasks received according to the ConTask version and usage period.

Version	Received tasks (first 3 days)	Received tasks (last 3 days)	Total tasks
ConTask 2.0	47	44	91
ConTask 1.0	42	38	80

From the first to the third day of the experiment, 12 tasks were created. Since each group had 4 people, it was possible for each group to submit up to 48 contributions. At this point, Group A was using ConTask 2.0, and Group B was using ConTask 1.0. During this period, the gamified version, ConTask 2.0, recorded 47 contributions, while the original version, ConTask 1.0, had 42 contributions, representing an increase of 10.64% in the number of tasks executed by Group A.

In the last three days, after the version swap between the groups, 12 tasks were also made available, with an expected maximum of 48 contributions per version. Group B, using ConTask 2.0, completed 44 contributions, while Group A, with ConTask 1.0, completed 38 contributions. During this period, ConTask 2.0 recorded a 13.63% increase in the number of tasks executed. Out of the maximum expected 96 contributions by the end of the experiment, ConTask 2.0 obtained 91, while ConTask 1.0 obtained 80. Thus, the gamified version had a 12.08% higher contribution rate than the

[13]Available at: `https://github.com/mpestana/Contask2`. Accessed on: January 24, 2024.

original version, indicating greater participation by volunteers with the gamified version.

Upon analyzing the results for each group, it can be observed that Group A, during the first 3 days using the gamified version, completed 47 contributions, and during the last 3 days with the original version, they completed 38 contributions. Group B, which used the original version in the first period, completed 42 tasks, and with the gamified version in the second period, this number increased to 44 tasks. It is noticeable that the number of contributions remains higher for the gamified version, even though there is a decrease in the total number of contributions in the second period of the study, which was expected, as volunteers tend to be more active in the beginning.

These data suggest that the gamified version may have had more tasks completed due to the implemented gamification techniques (ranking, notification, and feedback). This is in line with the literature, as seen in the works Feng *et al.* [2018]; Chi *et al.* [2018]; Lee *et al.* [2013], which indicated the beneficial relationship for collaboration in micro-task crowdsourcing systems. This relationship can be understood by discussing the possible motivational factors of the sample subjects used Zhang *et al.* [2022].

The results of the studies indicate that the use of gamification can be effective in increasing user motivation and participation in crowdsourcing tasks. However, it is important to note that there is a tendency for greater user participation at the beginning of the experiment, which can affect the analysis of the results and suggest a greater impact of gamification than is actually the case. Therefore, it is necessary to exercise caution in interpreting the data and consider the context in which the studies were conducted. Additionally, it is important to assess the long-term effects of gamification on user motivation and participation.

6.3 Perceptions on using Gamification

In research involving crowdsourcing and motivation, establishing generalizations is a complex task, as intrinsic and extrinsic motivations vary from person to person, influenced by cultural, religious, and social factors [Toda *et al.*, 2022a]. Although the sample size is considered small for generalization and establishing correlations between gamification use and increased task completion, we can discuss the results and perceptions in line with those found in the literature. Both groups used both versions for three days, allowing all participants to respond to the same questionnaire. At the beginning of the study, participants were instructed to enable notifications from the application. Therefore, during the study, they received notifications about available tasks, information about their ranking position, and feedback on tasks completed. This allowed us to analyze users' perceptions regarding the use of ranking, notification, and feedback techniques in ConTask 2.0.

The first question in the questionnaire aimed to gauge users' perception of notifications: *"At any point during the study, were you prompted to use ConTask because of the daily notification?"*. Five participants responded that they were encouraged to use ConTask and contribute to a task when they received notifications, or that the notifications helped them remember to open the application at an opportune moment. The remaining participants (3) said they didn't notice the notification when they received it or weren't prompted to contribute at that time. Various reasons were cited: notification function disabled, focus or silent mode, device in use with other applications, or they received the notification but dismissed it without reading the content.

Some works in the literature report difficulties and resource limitations related to notifications. Considering that the user can take immediate actions or simply ignore a notification, it is necessary to use them effectively and responsibly [Fitz *et al.*, 2019]. To do so, it is necessary to analyze the importance of the notification in the user's context and not use this feature without proper relevance criteria. Therefore, the chosen time for participants to be notified was at 6:00 PM local time, a moment considered possibly opportune [Chen *et al.*, 2022a]. The timing of notification receipt significantly influences how they are received by users, and an overload of notifications can prevent important notifications from being viewed [Esteves *et al.*, 2022; da Silva and Vieira, 2018].

The second question of the questionnaire aimed to obtain users' perception regarding the ranking: *"During the experiment, did you have an interest in performing tasks to improve your ranking position?"* Six participants responded affirmatively, expressing their interest and willingness to contribute to enhance their position. The remaining two participants, on the other hand, did not feel motivated and did not actively seek to improve their rankings. These perceptions align with discussions in the literature, as motivation can be influenced by both intrinsic factors, related to psychological rewards, and extrinsic factors, linked to material rewards.

Some works in the literature investigate what motivates specific groups of people (e.g., the elderly, teenagers) to participate in crowdsourcing projects [Amorim and Vieira, 2019; Ooge *et al.*, 2020]. Personalization of these systems, targeting the use of specific groups, is proposed as an alternative to overcome deficiencies caused by generalization of use [Rodrigues *et al.*, 2021; Tondello and Nacke, 2020]. The classification mechanism depends on motivating factors, as one person may feel motivated by tracking their performance in the ranking, while another person may not enjoy competition [Vaughan, 2018; Brewer *et al.*, 2016; Chandler and Kapelner, 2013].

The last question aimed to assess users' perception regarding the use of feedback: *"Did knowing the total number of contributions (tasks completed) and how many tasks were awaiting execution influence your desire to contribute?"* In response to this question, participants were divided in their opinions. Half of them indicated that the provided feedback, when read, did indeed stimulate their desire to continue performing tasks. They mentioned that the information about the number of available tasks also served as motivation because it compelled them to check which tasks were available. On the other hand, the other half did not perceive a relationship between the feedback and their desire to carry out tasks.

While gamified crowdsourcing has been investigated, there is still much to explore. In the work by Tsvetkova *et al.* [2022], the authors discuss the negative effects of feedback from a social perspective. They emphasize that if the individual effort and performance of people with different skills are not considered, inequality among individuals tends to increase. Furthermore, ethical discussions are held regarding the limits of competition and how to promote healthy collaboration [Toda *et al.*, 2022b]. These studies can provide context for the inconclusive results obtained in our study regarding the use of feedback as a motivating factor. Therefore, we align with the literature in recognizing the need for further research in this area.

6.4 Limitations and threats to the validity

Regarding the user study that investigates how the gamification techniques influence users of the crowdsourcing system, the threats to the validity of this study were analyzed and related to those identified based on Wohlin *et al.* [2012].

First, the experiment design. Despite working with two groups, neither of them functioned as a control group that could, for example, use only one version of the application. To mitigate this threat, the strategy of keeping the participants completely unaware of the specific version of ConTask they were using was adopted. However, it is worth noting that in future research, the inclusion of a strict control group could provide a more accurate assessment of the differences between the application versions.

Secondly, the experiment protocol remained unchanged during the first and last three days, which introduced the possibility of participants behaving differently in the second half of the study, possibly due to becoming more familiar with the procedure. To minimize this bias, we chose to restrict the disclosure of results for each stage until the end of the six days, keeping participants equally uninformed throughout the experiment. However, it is worth considering the implementation of variations in the experimental protocol in future research to address this issue more comprehensively.

Finally, participant selection deserves attention. Although our research has a qualitative approach and does not aim for generalization, it is relevant to note that participants were volunteers, which, according to Wohlin *et al.* [2012], may result in a more motivated sample that is better suited for new tasks. This characteristic of volunteers can affect the applicability of the results to broader contexts. Therefore, future research needs to consider strategies for involving a more diverse range of participants, including those who may be less motivated or experienced with the task at hand.

Regarding the user study that investigates contribution filtering techniques, the major challenges encountered are related to the enhancement of the Boyer-Moore algorithm for contribution filtering. This arose due to the necessity of handling punctuation and special characters, which significantly increased the complexity of the development process. Additionally, difficulties were faced when seeking ways to distinguish and randomize synthetically generated contributions from real contributions.

7 Conclusions and Future Work

In this paper, challenges related to the motivation of individuals and the effective management of a large volume of contributions in crowdsourcing are highlighted. The crowdsourcing model facilitates collaboration among diverse individuals, offering flexibility and accessibility, and allowing contributions based on individual skills and interests.

Two studies were presented. The first study was conducted to investigate how gamification techniques (ranking, notification, and feedback) can influence users in crowdsourcing. The study used a crowdsourcing application in two versions, one with gamification implementations and one without. The study's results indicated that the gamified version of the application showed more tasks completed by participants. At the end of the study, a questionnaire was administered where participants evaluated both versions of the application. The questionnaire analysis revealed that participants rated the three implemented gamification mechanisms positively and felt more motivated to perform tasks in their presence.

The second study presented aimed to investigate techniques to support requesters of micro-tasks in crowdsourcing systems in managing and selecting the most relevant contributions. This study also used a crowdsourcing micro-task application to conduct an experimental study to validate two contribution filtering techniques developed: the keyword-based contribution filter and the majority vote-based contribution filter. The experimental study involved the request and distribution of micro-tasks among participants, who were tasked with contributing and voting on contributions they deemed most suitable for solving a task. Subsequently, a questionnaire was administered to task requesters to gather their opinions on the results of applying the contribution filters. This experiment was divided into three stages and indicated that the majority of requesters found that the majority vote-based contribution filter assisted in the selection of the most relevant contributions. Additionally, all requesters expressed a positive reception to the proposal of using the keyword-based contribution filter.

We suggest as future work: i) to validate the use of ConTask 2.0 with a larger number of participants; ii) enhance the existing functionalities of the application, and analyze the impact of the artifacts on user participation, both in terms of the number of tasks completed and the quality of accepted contributions; iii) to diversify the types of micro-tasks offered to participants and expand their analysis by considering the specific type of task executed; iv) The need to explore ways to integrate synonym and antonym APIs to enhance the accuracy of the keyword filter in the matching process; v) The implementation of a feature to allow users to undo their votes once they have been submitted; vi) The creation of user profiles that assign different weights to votes, giving more influence to users with a high reputation in the application; vii) The adaptation of the text filter to recognize numerical variations, such as singular and plural forms, in keywords. Lastly, it's worth noting that the application's interface has not incorporated UX/UI techniques, suggesting the need for a future study with a focus on improving the interface design.

Declarations

Acknowledgements

We extend our gratitude to the participants for their valuable time and contributions, which are sincerely appreciated. Special thanks are also due to the Research Group on Collaborative Systems and Context-Aware Computing (CEManTIKA).

Funding

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and the Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB).

Authors' Contributions

Ana Maria Amorim: Investigation, Supervision, Research orientation, Writing – review, Methodology, Validation. **Ailton Ribeiro**: Conceptualization, Formal analysis, Writing – review & editing. **Murilo Guerreiro Arouca**: Conceptualization, Formal analysis, Writing – review & editing. **Iago Meijon**: Investigation, Writing – original draft, Data curation, Methodology, Validation. **Victor Cavalheiro**: Investigation, Writing – original draft, Data curation, Methodology, Validation. **Maria Clara Pestana**: Conceptualization, Writing – review. **Vaninha Vieira**: Project administration, Supervision, Writing – review & editing.

Competing interests

The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Alam, S. L. and Sun, R. (2023). The role of system-use practices for sustaining motivation in crowdsourcing: A technology-in-practice perspective. *Information Systems Journal*, 33(4):758–789. DOI: https://doi.org/10.1111/isj.12423.

Amorim, A. M. and Vieira, V. (2019). Exploratory study on the motivation of brazilian elderly people in crowdsourcing systems. In *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*, IHC '19, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3357155.3360478.

Amorim, A. M. and Vieira, V. (2023). Participation in crowdsourcing micro-tasks: what motivates brazilian older adults? *Universal Access in the Information Society*, pages 1–25. DOI: https://doi.org/10.1007/s10209-023-01023-9.

Arouca, M., Ribeiro, A., Amorim, A., Neves, I., Vieira, V., Barreto, M., Costa, F., and Brito, R. (2024). Gamification to support crowdsourcing and participatory mapping for signaling and spatialization of covid-19 transmission predictors. In *Proceedings of the 19th Brazilian Symposium on Collaborative Systems*, pages 42–55, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbsc.2024.238059.

Bastanfard, A., Shahabipour, M., and Amirkhani, D. (2023). Crowdsourcing of labeling image objects: An online gamification application for data collection. *Multimedia Tools and Applications*, pages 1–34. DOI: https://doi.org/10.1007/s11042-023-16325-6.

Boyer, R. S. and Moore, J. S. (1977). A fast string searching algorithm. *Commun. ACM*, 20(10):762–772. DOI: https://doi.org/10.1145/359842.359859.

Brabham, D. (2013). *Crowdsourcing*. The MIT Press Essential Knowledge series. MIT Press.

Brabham, D. C. (2008). Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence*, 14(1):75–90. DOI: https://doi.org/10.1177/1354856507084420.

Brewer, R., Morris, M., and Piper, A. (2016). Why would anybody do this?: Older adults' understanding of and experiences with crowd work. In *CHI 2016 - 34th Annual CHI Conference on Human Factors in Computing Systems*, pages 2246–2257. Association for Computing Machinery. DOI: https://doi.org/10.1145/2858036.2858198.

CETIC.BR-NIC.BR (2021). Executive summary - survey on the use of information and communication technologies in brazilian households - ict households 2020. Regional Center for Studies on the Development of the Information Society – Cetic.br; Brazilian Network Information Center – NIC.br.

Chandler, D. and Kapelner, A. (2013). Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior Organization*, 90:123–133. DOI: https://doi.org/10.1016/j.jebo.2013.03.003.

Charmaz, K. (2014). *Constructing Grounded Theory*. Introducing Qualitative Methods series. SAGE Publications.

Chen, K.-W., Chang, Y.-J., and Chan, L. (2022a). Predicting opportune moments to deliver notifications in virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3491102.3517529.

Chen, L., Xu, P., and Liu, D. (2020). Effect of crowd voting on participation in crowdsourcing contests. *Journal of Management Information Systems*, 37(2):510–535. DOI: https://doi.org/10.1080/07421222.2020.1759342.

Chen, Z., Jiang, L., and Li, C. (2022b). Label augmented and weighted majority voting for crowdsourcing. *Information Sciences*, 606:397–409. DOI: https://doi.org/10.1016/j.ins.2022.05.066.

Chi, P.-Y. P., Batra, A., and Hsu, M. (2018). Mobile crowdsourcing in the wild: challenges from a global community. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, MobileHCI '18, page 410–415, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3236112.3236176.

Cole, R. (1994). Tight bounds on the complexity of the boyer–moore string matching algorithm. *SIAM Journal on Computing*, 23(5):1075–1091. DOI: https://doi.org/10.1137/S0097539791195543.

Congcong Yang, H. J. Y. and Feng, Y. (2021). Using gamification elements for competitive crowdsourcing: exploring the underlying mechanism. *Behaviour & Information Technology*, 40(9):837–854. DOI: https://doi.org/10.1080/0144929X.2020.1733088.

da Silva, A. V. D. and Vieira, V. (2018). Towards an api for user attention prediction in mobile notification overload. In *Anais Estendidos do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web*, pages 13–17, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/webmedia.2018.4552.

Deci, E. L. and Ryan, R. M. (1980). Self-determination theory: When mind mediates behavior. *The Journal of mind and Behavior*, pages 33–43.

Deci, E. L. and Ryan, R. M. (1985). *Cognitive Evaluation Theory*, pages 43–85. Springer US, Boston, MA. DOI: https://doi.org/10.1007/978-1-4899-2271-7_3.

Deci, E. L. and Ryan, R. M. (2010). *Intrinsic Motivation*, pages 1–2. John Wiley Sons, Ltd. DOI: https://doi.org/10.1002/9780470479216.corpsy0467.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., and Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, page 2425–2428, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/1979742.1979575.

Esteves, B., Fraser, K., Kulkarni, S., Conlan, O., and Rodríguez-Doncel, V. (2022). Now, later, never: A study of urgency in mobile push-notifications. In Delir Haghighi, P., Khalil, I., and Kotsis, G., editors, *Advances in Mobile Computing and Multimedia Intelligence*, pages 38–44, Cham. Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-20436-4_4.

Feng, Y., Jonathan Ye, H., Yu, Y., Yang, C., and Cui, T. (2018). Gamification artifacts and crowdsourcing participation: Examining the mediating role of intrinsic motivations. *Computers in Human Behavior*, 81:124–136. DOI: https://doi.org/10.1016/j.chb.2017.12.018.

Feng, Y., Yi, Z., Yang, C., Chen, R., and Feng, Y. (2022). How do gamification mechanics drive solvers' knowledge contribution? a study of collaborative knowledge crowdsourcing. *Technological Forecasting and Social Change*, 177:121520. DOI: https://doi.org/10.1016/j.techfore.2022.121520.

Fitz, N., Kushlev, K., Jagannathan, R., Lewis, T., Paliwal, D., and Ariely, D. (2019). Batching smartphone notifications can improve well-being. *Computers in Human Behavior*, 101:84–94. DOI: https://doi.org/10.1016/j.chb.2019.07.016.

Gadiraju, U., Kawase, R., and Dietze, S. (2014). A taxonomy of microtasks on the web. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, HT '14, page 218–223, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/2631775.2631819.

Günther, H. (2003). Como elaborar um questionário. *Série: Planejamento de pesquisa nas ciências sociais*, 1:1–15.

Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14(6):1–4.

Kobayashi, M., Arita, S., Itoko, T., Saito, S., and Takagi, H. (2015). Motivating multi-generational crowd workers in social-purpose work. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '15, page 1813–1824, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/2675133.2675255.

Law, E., Yin, M., Goh, J., Chen, K., Terry, M. A., and Gajos, K. Z. (2016). Curiosity killed the cat, but makes crowdwork better. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 4098–4110, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/2858036.2858144.

Lazar, J., Feng, J., and Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction*. Elsevier Science.

Lee, T. Y., Dugan, C., Geyer, W., Ratchford, T., Rasmussen, J., Shami, N. S., and Lupushor, S. (2013). Experiments on motivational feedback for crowdsourced workers. DOI: https://doi.org/10.1609/icwsm.v7i1.14428.

Lopes, W., Augusto, P., Fernandes, I., and Madeira, C. (2024). Proposal for a gamification strategy applied to remote learning. *Journal on Interactive Systems*, 15(1):92–103. DOI: https://doi.org/10.5753/jis.2024.2700.

Meijon, I., Amorim, A. M., Ribeiro, A., Pestana, M. C., and Vieira, V. (2023). A study on applying gamification techniques to a crowdsourcing app for micro-tasks. In *Proceedings of the 18th Brazilian Symposium on Collaborative Systems*, pages 57–71, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbsc.2023.229082.

Meliande, R., Ribeiro, A., Arouca, M., Amorim, A., Pestana, M., and Vieira, V. (2024). Meta-education: A case study in academic events in the metaverse. In *Proceedings of the 19th Brazilian Symposium on Collaborative Systems*, pages 28–41, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbsc.2024.238057.

Morschheuser, B., Hamari, J., and Koivisto, J. (2016). Gamification in crowdsourcing: A review. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 4375–4384. DOI: 10.1109/HICSS.2016.543.

Navarro, G. and Tarhio, J. (2000). Boyer—moore string matching over ziv-lempel compressed text. In Giancarlo, R. and Sankoff, D., editors, *Combinatorial Pattern Matching*, pages 166–180, Berlin, Heidelberg. Springer Berlin Heidelberg. DOI: https://doi.org/10.1007/3-540-45123-4_16.

Ooge, J., De Croon, R., Verbert, K., and Vanden Abeele, V. (2020). Tailoring gamification for adolescents: a validation study of big five and hexad in dutch. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '20, page 206–218, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3410404.3414267.

Park, S., Kwon, S., and Lee, U. (2018). Campuswatch: Exploring communitysourced patrolling with pervasive mobile technology. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW). DOI: https://doi.org/10.1145/3274403.

Pestana, M. C. and Vieira, V. (2018a). Context-aware task distribution for mobile crowdsourcing. In *Pro-*

ceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems, IHC '18, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3274192.3274206.

Pestana, M. C. and Vieira, V. (2018b). Context-aware task distribution for mobile crowdsourcing. In *Proceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems*, IHC 2018, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3274192.3274206.

Ribeiro, A., Vieira, V., Alves, L., and Maciel, C. (2024). Vishnu: An approach to support the personalization of self-expressive avatars using context-awareness. *International Journal of Human-Computer Studies*, 185:103243. DOI: https://doi.org/10.1016/j.ijhcs.2024.103243.

Rodrigues, L., Palomino, P. T., Toda, A. M., Klock, A. C. T., Oliveira, W., Avila-Santos, A. P., Gasparini, I., and Isotani, S. (2021). Personalization improves gamification: Evidence from a mixed-methods study. *Proc. ACM Hum.-Comput. Interact.*, 5(CHI PLAY). DOI: https://doi.org/10.1145/3474714.

Sailer, M., Hense, J., Mandl, H., and Klevers, M. (2013). Psychological perspectives on motivation through gamification. *Interaction Design and Architecture(s) - IxD&A*, (19):28 – 37. DOI: https://doi.org/10.55612/s-5002-019-002.

Stol, K.-J. and Fitzgerald, B. (2014). Two's company, three's a crowd: A case study of crowdsourcing software development. In *Proceedings of the 36th International Conference on Software Engineering*, ICSE 2014, page 187–198, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/2568225.2568249.

Tanalol, H., Hashim, H., Turumogan, P., Noor, N. A. M., Baharum, A., and Deris, F. D. (2023). Identifying gamified teaching elements in computer science course. In *2023 International Conference on Platform Technology and Service (PlatCon)*, pages 18–23. DOI: https://doi.org/10.1109/PlatCon60102.2023.10255180.

Tao, F., Jiang, L., and Li, C. (2020). Label similarity-based weighted soft majority voting and pairing for crowdsourcing. *Knowledge and Information Systems*, 62:2521–2538. DOI: https://doi.org/10.1007/s10115-020-01475-y.

Toda, A., Klock, A., Pereira, F. D., Rodrigues, L. A., Palomino, P. T., Lopes, V., Stewart, C., Oliveira, E. H. T., Gasparini, I., Isotani, S., and Cristea, A. (2022a). Towards the understanding of cultural differences in between gamification preferences: A data-driven comparison between the US and Brazil. In Mitrovic, A. and Bosch, N., editors, *Proceedings of the 15th International Conference on Educational Data Mining*, pages 560–564, Durham, United Kingdom. International Educational Data Mining Society. DOI: https://doi.org/10.5281/zenodo.6853030.

Toda, A., Palomino, P. T., Rodrigues, L., Klock, A. C. T., Pereira, F., Borges, S., Gasparini, I., Teixeira, E. H., Isotani, S., and Cristea, A. I. (2022b). Gamification through the looking glass - perceived biases and ethical concerns of brazilian teachers. In *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium: 23rd International Conference, AIED 2022, Durham, UK, July 27–31, 2022, Proceedings, Part II*, page 259–262, Berlin, Heidelberg. Springer-Verlag. DOI: https://doi.org/10.1007/978-3-031-11647-6_47.

Tondello, G. F. and Nacke, L. E. (2020). Validation of user preferences and effects of personalized gamification on task performance. *Frontiers in Computer Science*, 2:29. DOI: https://doi.org/10.3389/fcomp.2020.00029.

Tsvetkova, M., Müller, S., Vuculescu, O., Ham, H., and Sergeev, R. A. (2022). Relative feedback increases disparities in effort and performance in crowdsourcing contests: Evidence from a quasi-experiment on topcoder. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2). DOI: https://doi.org/10.1145/3555649.

Vaughan, J. W. (2018). Making better use of the crowd: How crowdsourcing can advance machine learning research. *Journal of Machine Learning Research*, 18(193):1–46.

Vieira, V., Tedesco, P., and Salgado, A. C. (2011). Designing context-sensitive systems: An integrated approach. *Expert Systems with Applications*, 38(2):1119–1138. Intelligent Collaboration and Design. DOI: https://doi.org/10.1016/j.eswa.2010.05.006.

Walter, V., Kölle, M., and Collmar, D. (2022). A gamification approach for the improvement of paid crowd-based labelling of geospatial data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 5(4):113–120. DOI: https://doi.org/10.5194/isprs-Annals-V-4-2022-113-2022.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B., and Wesslén, A. (2012). *Experimentation in Software Engineering*. Computer Science. Springer Berlin Heidelberg.

Wu, W. and Gong, X. (2021). Motivation and sustained participation in the online crowdsourcing community: the moderating role of community commitment. *Internet Research*, 31(1):287–314. DOI: https://doi.org/10.1108/INTR-01-2020-0008.

Zhang, X., Xia, E., Shen, C., and Su, J. (2022). Factors influencing solvers' behaviors in knowledge-intensive crowdsourcing: A systematic literature review. *Journal of Theoretical and Applied Electronic Commerce Research*, 17(4):1297–1319. DOI: https://doi.org/10.3390/jtaer17040066.