POLISH-JAPANESE ACADEMY OF INFORMATION TECHNOLOGY

Designing Crowdsourcing Systems for Older Adults

Kinga H. Skorupska
Student number: S18181

ICT & Psychology
Polish-Japanese Academy of Information Technology
Thesis prepared under the supervision of:
Prof. Grzegorz Sędek, Ph.D.
Auxiliary supervisors:
Radosław Nielek, Ph.D.

Warsaw, June 1, 2022

Abstract. Designing crowdsourcing systems for older adults is a significant challenge in the domain of Informatics. This multidisciplinary problem is related to multiple issues in the area of HCI, such as accessibility, user motivation as well as crowdsourcing-specific considerations, like task decomposition and presentation. The existing design paradigms within all of these constitute barriers to older adults' productive participation in crowdsourcing initiatives. Therefore, crowdsourcing systems designed with older adults in mind ought to not only mitigate the aforementioned barriers and answer older adults' needs related to human-human interaction (HHI) and human-computer interaction (HCI), providing them with a positive experience, but also produce valuable contributions, by capitalizing on older adults' unique characteristics and strong suits, such as their lifelong experience and skills. To develop such systems, it was crucial to verify different aspects of crowdsourcing systems, such as diverse use scenarios and modes of interaction with traditional – PC-based, and non-traditional computer systems as well as various crowdsourcing tasks and their presentation. In the course of this research a Smart TV-based system, a chatbot and web-based solutions were proposed and tested to explore barriers and opportunities for older adults' involvement in crowdsourcing projects. Those systems were evaluated both in terms of the manner of engaging with them and the attractiveness of the crowdsourcing tasks themselves. Based on (1) the design and verification of these systems in various exploratory studies as well as (2) a meta-analysis of other research related to relevant design heuristics and (3) the exploration of design guidelines which emerged at the intersection of the areas of older adults' use of technology and crowdsourcing **I propose the comprehensive AFFORCE framework for designing crowdsourcing systems for older adults**. This framework constitutes a design and development guide for researchers and practitioners who would like to build engaging crowdsourcing experiences that would encourage, especially, older adults to become active contributors and could act as a gateway to empower them to use other ICT-based solutions.

ACM Computing Classification:
• Information systems - **Crowdsourcing**
• Human-centered computing - **Interaction paradigms, Interaction design**
• Social and professional topics - **Seniors**

Streszczenie Projektowanie systemów crowdsourcingowych dla osób starszych jest istotnym wyzwaniem w domenie nauk informatycznych. Ten multidyscyplinarny problem związany jest z wieloma zagadnieniami w obszarze HCI, takimi jak dostępność, motywacja użytkowników, a także czynnikami specyficznymi dla crowdsourcingu, jak dekompozycja zadań i ich prezentacja. Wykorzystywane obecnie paradygmaty projektowania mogą stanowić bariery dla produktywnego udziału osób starszych w inicjatywach crowdsourcingowych. Dlatego systemy crowdsourcingowe zaprojektowane z myślą o osobach starszych powinny nie tylko niwelować bariery ale także odpowiadać na potrzeby osób starszych związane z interakcją człowiek-człowiek (HHI) i człowiek-komputer (HCI) oraz korzystać z mocnych strony osób starszych, takich jak ich doświadczenia, wiedza i umiejętności wypracowane przez całe życie. W pierwszym etapie prac zweryfikowano różne aspekty systemów crowdsourcingowych, zarówno tradycyjnych, bazujących na komputerach osobistych, jak i nietradycyjnych, bazujących na urządzeniach mobilnych i urządzeniach IoT. Zbadano scenariusze ich użytkowania, tryby interakcji z tymi systemami, a także różne zadania crowdsourcingowe i ich prezentację. W ramach badań przedstawionych w niniejszej pracy zaprojektowano i przetestowano autorskie rozwiązania: system oparty na Smart TV, chatbota i systemy webowe. Zbadano bariery i możliwości zaangażowania osób starszych w projekty crowdsourcingowe, zarówno pod względem ich sposobu angażowania się w tego typu projekty, jak i atrakcyjność samych zadań. Na podstawie (1) zaproponowanych systemów oraz ich weryfikacji w różnych badaniach eksploracyjnych, a także (2) metaanalizy innych badań związanych z odpowiednimi heurystykami projektowymi oraz (3) eksploracji wytycznych projektowych, które pojawiły się na przecięciu obszarów wykorzystania technologii przez osoby starsze i crowdsourcingu opracowano AFFORCE: kompleksowy framework do projektowania systemów crowdsourcingowych dla osób starszych. Framework ten stanowi przewodnik projektowania i rozwoju systemów crowdsourcingowych zarówno dla badaczy jak i praktyków, którzy chcieliby tworzyć angażujące doświadczenia zachęcające użytkowników, a zwłaszcza osoby starsze, zarówno do aktywnego udziału w projektach crowdsourcingowych jak i do efektywnego korzystania z innych rozwiązań opartych na ICT.

ACM Computing Classification:
• Information systems - **Crowdsourcing**
• Human-centered computing - **Interaction paradigms, Interaction design**
• Social and professional topics - **Seniors**

Table of Contents

1 Introduction

Demographic trends, shaped by decreasing birth rates and longer lifespan, signal an accelerating increase in the share of older adults in western societies. In 2020 people aged 65+ comprised 20.6% of the EU population [5] and projections for 2100 establish their future share at 31.3%. The same trend persists in other parts of the world [1]. U.S. Census Bureau also projects that by 2050 the population of people aged 65+ will almost double in the U.S. [99]. In 2019 there were over 700 million older adults aged 65+ in the world, and by 2050 this number is expected to double worldwide [4]. [1]

In the EU-27 [2] the potential of this increasingly numerous group of older adults remains largely untapped as in 2019 only 8.1% of men and 3.9% of women aged 65+ had full time employment. At the same time slightly over half of them remained active in the labour market on part-time basis [3]. There is a shortage of both jobs and projects allowing older adults to comfortably contribute their unique skills, knowledge and experience [85]. Similarly, opportunities for their remote engagement with the use of ICT-tools are overlooked because of stereotypes around older adults' ICT-learning, ICT-proficiency and presence online. Research and initiatives that would encourage older adults to remain active, both mentally and physically to foster their health, engagement and well-being are often skipped in favour of a disproportionate focus on stereotypes about health, social life, tech skills [135] and broadly understood accessibility [66]. This stereotype threat also has an adverse effect on the performance of older adults themselves, as they, as the affected group, expect to perform worse [16] and in effect they often do. So, despite significant

[1] Even though the COVID-19 pandemic has predominantly posed threat to older adults, due its effects on birth rates in developed countries [6] these trends remain timely.

[2] EU-27 refers to the number of member states of the European Union at the time of the creation of the relevant statistic.

progress in the area of designing ICT solutions for older adults over the last 20 years there is still room for further research [65,64,97]. Especially the interdisciplinary approach, making use of the findings in the area of cognitive and behavioral psychology, psychology and aging, as well as human factors psychology holds promise of supporting ICT research in the efforts to design ICT solutions which have a greater chance of engaging older adults.

In particular, crowdsourcing platforms could benefit from older adults' massive untapped potential, including their knowledge, skills and rich life-long experience as well as social and motivational compensatory mechanisms [133]. Increasingly, this area gains in significance with the rise of big data [35] and the potential of AI an ML solutions, making use of human-in-the-loop workflows [28,116]. Yet, crowdsourcing as implemented currently largely does work for older adults. There exist multiple barriers to older adults' inclusion in typical crowdsourcing projects. These are related to the experiences offered by crowdsourcing systems, their features and the design of crowdsourcing tasks and are outlined in detail in Section 2.6.

Encouraging older adults' contributions to crowdsourcing is a significant current and future (as explained in Section 3.4) challenge. It is beneficial to the society, because of the direct value of crowdsourcing contributions and the importance of building social capital [145,138]. It also benefits older adults individually, as it gives them opportunities to stay active and foster their sense of well-being [43,39,116], especially as more older adults are becoming active online.

Therefore, developing new crowdsourcing systems, making use of novel interfaces, for older adults, which would not only mitigate the aforementioned barriers but also make use of older adults' strong suits, such as their professional expertise [85], higher crystallized intelligence [55], good transitive reasoning abilities [117], pro-social motivation and diligence [81], is an important research problem. Such systems could empower older adults to become active and valuable contributors, also paving way to their increased engagement with other tasks and activities dependent on Information and Communication Technology (ICT).

Moreover, thanks to the curb-cut effect, which is established in the digital space [53] [3] designing for older adults has a great chance to improve the crowdsourcing experience for other populations as well. As Alan Newell put it "Attention paid to developing technology with a specific consideration of the needs of older people now and in the future will not only bring social and economic dividends in itself but will help in the design of technology which is more accessible for all users." [95]

[3] Curb-cut effect is based on the idea of designing for specific populations first, thereby being able to put design goals more clearly – only then the benefits of the design "trickle down" to other populations.

1.1 Research Problem: Developing Crowdsourcing Systems that are Engaging for Older Adults

The problem of designing and developing crowdsourcing systems for older adults is both underappreciated and underexplored, as there is very little research which directly touches upon this topic. The existing studies largely point to the barriers to crowdsourcing for older adults, but few explore the ways to mitigate them and even fewer focus on opportunities or present working solutions (as described in Section 2.1). It is more dire, especially in light of the multiple challenges specific to crowdsourcing systems and task design as well as ICT learning (see Table 1 in Section 2.6). There exists no successful ICT-based crowdsourcing system designed specifically to encourage older adults to contribute, so the area of designing working crowdsourcing systems for older adults constitutes a significant research niche.

The design, development and verification of such crowdsourcing systems and the formulation of an actionable and relevant system design framework would constitute a timely contribution to the domain of Informatics, especially in the area of Human-Computer Interaction (HCI) and help realize important social goals as well as foster the individual well-being of older adults.

1.2 Research Objectives

To explore this research problem I have formulated three specific research objectives, of which objective 1 and 2 overlapped in time, as designing ICT solutions is an iterative process.

1. Analyse older adults' interaction with diverse crowdsourcing systems, including ICT-related barriers to participation and motivational paradigms especially prominent for older adults
To gather insights for the creation of actionable design guidelines for crowdsourcing systems it is necessary to verify older adults' interaction and engagement with both novel crowdsourcing systems and interfaces, and traditional web-based crowdsourcing solutions. For this reason, the proposed novel interfaces ought to be further tested in exploratory studies alongside tasks, which can be most commonly found as part of crowdsourcing projects.

2. Develop, design and verify novel crowdsourcing interfaces for older adults
Existing crowdsourcing interfaces are largely reliant on computer skills, are browser-based and built on the WIMP paradigm. This alone may constitute a barrier as a comfortable computer setup is a must, as well as learning unfamiliar metaphors, which commonly function in the ICT space. Therefore, designing and verifying crowdsourcing systems that explore other interaction paradigms, especially relying on prior familiarity with everyday activities, such as controlling a TV-set via a remote or giving voice commands, would constitute a marked novelty.

3. Formulate a framework for designing crowdsourcing systems for older adults

On the basis of (a) the experience and results of prior exploratory studies, (b) a related literature review in the area of crowdsourcing for older adults and (c) the exploration of design guidelines which emerged at the intersection of the areas of older adults' use of technology and crowdsourcing an actionable framework for designing crowdsourcing systems for older adults will be developed. It would enable system designers to reference a comprehensive source of knowledge on this topic and to create more accessible and motivating experiences for older adults who are an increasingly prominent group of potential contributors. As will be shown by the review of the state of the art in the following section, this framework will constitute a significant innovation in the domain of Informatics, especially in the area of HCI.

2 Overview of the State-of-the-art

2.1 Crowdsourcing Systems

In 2013 Kittur et al. [61] postulated a set of possible advancements in the area of crowdsourcing, both in terms of the new trends of using ML and human-in-the-loop systems, possible workflows and the motivational and socioeconomic aspects of crowd work. Since then, too little has changed in terms of the value of human labour and caring for the well-being of crowd workers, who, in turn may provide poor quality contributions. Crowdsourcing tasks are often designed in a way which is far-removed from the experience of the general population [19] while being too short to allow for the understanding of their purpose [61], especially if they originate as smaller chunks of more complex tasks [62]. On top of this, there are often harsh deadlines for tasks' completion [148] which result in dropouts or abandonment, discouraging contributors [45]. One way of improving task quality is by allowing for users' self-assessment [38], as faced with it the users solve the tasks 15% more accurately and providing a way to become better contributors with clear advancement paths [137]. Also support systems are valuable as they help users find appropriate tasks, either with the support of the platform or the community [26].

In general, crowdsourcing encompasses a broad range of activities online, from co-creation of content online to distributed work [30] and online work platforms, such as Upwork or Freelancer. However, in this research we will focus on crowdsourcing which hinges on many users performing tasks, which are part of a larger whole, encompassing a crowdsourcing project. Overall, despite diverse commercial and volunteer-based implementations of crowdsourcing systems, most of the advancements are related to the technical aspects of enabling crowd work, rather than human factors, such as access to crowdsourcing for groups excluded from the main technological discourse (older adults, people whose native language is not English, people with worse Internet access) or the totality of the crowdsourcing experience, with the focus on sustainability in terms of the workers' well-being [61]. At the beginning of the COVID-19 pandemic there was a large organic spike of interest in crowdsourcing, connected to the ability to work and volunteer remotely – however it was short lived, perhaps as not much has changed in terms of large-scale crowdsourcing project implementations, both for the general public and the groups most affected by the necessity of social isolation, such as older adults.

2.2 Select Crowdsourcing Platforms and Projects

What follows are descriptions of crowdsourcing projects which were used as case studies for the design and evaluation of crowdsourcing systems for older adults covered by this thesis.

2.2.1 Amara for TED Translators TED Translators, formerly TED Open Translation Program, allows volunteers to create subtitles for TED, TEDx and TED-Ed talks on a dedicated platform. The initiative allows contributors to make subtitles in over 100 languages. The translators' community of practice [4] [77] attempts to create good quality subtitles despite constituting of mainly non-professional volunteers (as of January 2022 over 41 thousand of them), who need to learn a lot of best practices. There are multiple issues, which affect this community, such as shortage of volunteers, especially reviewers, which are the first level of Quality Assurance (QA) and Language Coordinators (the second QA level). The main motivation of the members of this community is to "spread ideas" [37], which is aligned with the mission of TED. However, the QA interface, just as the translation interface is web-based (see Figure 1) and requires the use of keyboard shortcuts to be used efficiently.

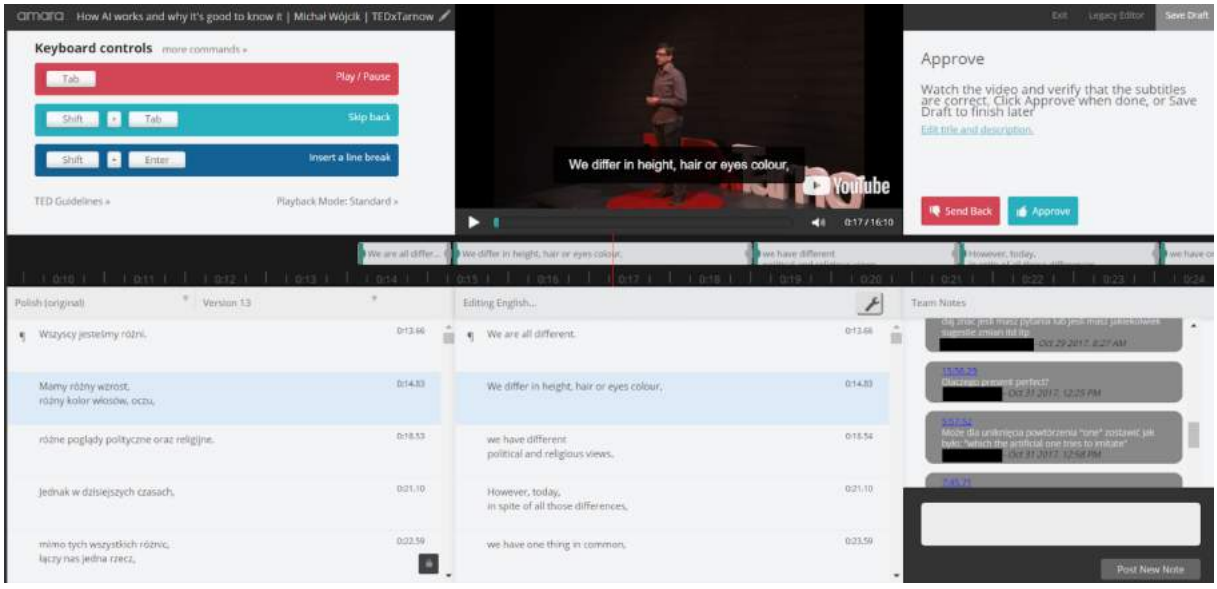


Fig. 1. The GUI of the Amara online subtitling platform.

It is a challenge, as there are over 3500 TED videos on TED.com website, and over 160 000 videos on the official TEDx YouTube channel, in addition to that there are special TED-Ed lessons, which also require subtitles. With a 2-step QA process and the potential

[4] Communities of Practice are what Leave and Wenger [77] describe as a group driven by common interest that collectively improves and gains knowledge

of a single video being translated into 115 languages this creates a massive workload. This workload could be assisted by Machine Translation (MT), which has the potential to reduce it by over 40% [31] through providing base translations. Yet, the use of MT increases workload in terms of the necessary QA. Therefore, there is room to involve crowdsourcing and citizen science in the process of creating quality subtitles, which in itself can be quite complex [108], with the help of Machine Learning (ML) and Natural Language Processing (NLP), through supervised ML and QA tasks. Both quantity and quality of subtitles for educational content is important, as they provide access to ideas to non-native speakers and to the hard of hearing. Older adults were a good target for being introduced into this workflow as QA contributors, as they are experts at their native language which makes them able to find crucial mistakes – especially that transcription, that is same-language subtitles, is the most important step, as it is the source for all translations. Moreover, they could help with QA of translation into their native languages, where they can easily ascertain if utterances sound natural.

2.2.2 Wiki Engine for Wikipedia Wikipedia is a massive crowdsourcing initiative, with 40 million registered users on the English page alone [2]. Even though all of them have the capacity to provide content to Wikipedia, actually, a small minority of them are active regularly – these are the ones who decide on take of each entry [78]. However, the disproportion between Wikipedia's active editors and viewers is increasing each year [84] making the voice of the active few ever more prominent. Among the factors that drive Wikipedia editors to contribute is their need to maintain a positive image of themselves [141]. Yet, some senior contributors may discourage novice ones from participating [44], which, on top of Wikipedia's problems with accessibility and the confusing GUI [96], complexity of procedures and multitude of guidelines drives new participants away from the project. These problems have also come to the attention of the Wikipedia crowdsourcing community [5]. Apart from contributing to the content of Wikipedia through following up-to-date research [129] or localizing articles, which is still a massive task [140], there are other possible contribution venues, such as validating editors' credibility [128], or verifying information [107]. Moreover, there were attempts to use Machine Translation to lessen the load on editors [144], especially that it could ensure that a similar format and take exist in each language, thus increasing its internal consistency.

Overall, there are multiple challenges that Wikipedia is facing, including the confusing interface which has a poor visual editor, the necessity to learn the markup language as well

[5] This is evidenced, for example, by the analysis found here: https://en.wikipedia.org/wiki/Wikipedia:Why_is_Wikipedia_losing_contributors__Thinking_about_remedies

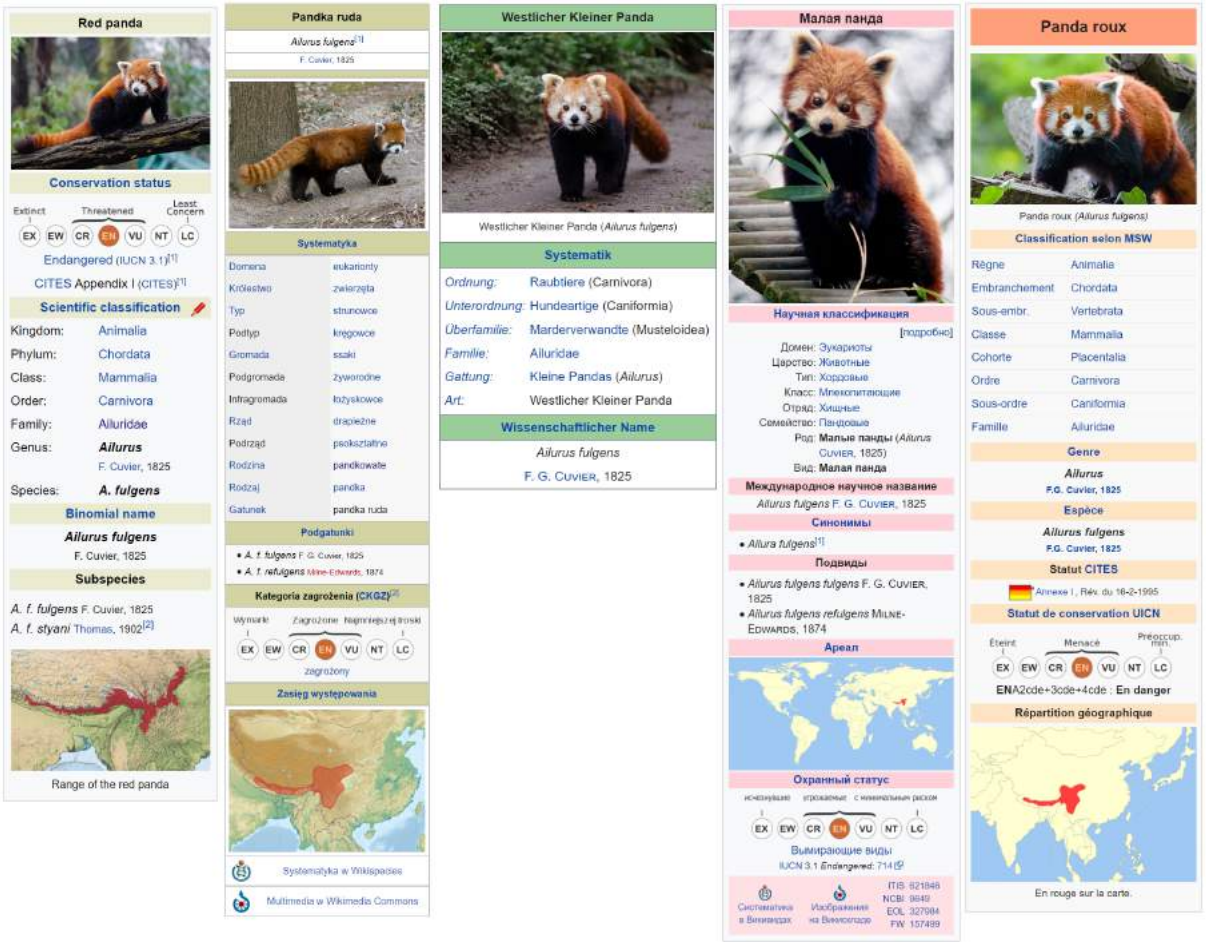


Fig. 2. Example of different infobox design in different language versions (from left to right: English, Polish, German, Russian and French) of the same article on Red pandas.

as lack of automation and confusing policies, requiring new contributors to read a lot [6]. Systems that could mitigate these problems could give voice to groups excluded from the main technological discourse, and allow Wikipedia to further expand. There are multiple opportunities in terms of content that could be broken into smaller crowdsourcing tasks that would make use of older adults' strong suits and that includes tagging articles for clean-up, classifying articles, verifying sentiment (guidelines for the unbiased treatment) as well as assessing information credibility and verifying machine translated parts of articles, such as infoboxes. Those are a required element in some types of articles, moreover, they are often inconsistent across languages (as seen in Figure 2) and they do not update automatically when article texts are changed. Therefore, they are a good first step in translation automation and adding Quality Assurance (QA) to ML workflows, especially for older adults, who are experts in their native language.

[6] New volunteers are expected to read pages such as: https://en.wikipedia.org/wiki/Wikipedia:Guidance_for_younger_editors (accessed on 28.01.2020, revision ID=937682704) or https://en.wikipedia.org/wiki/Wikipedia:A_primer_for_newcomers (accessed on 28.01.2020, revision ID=933431927)

2.2.3 Zooniverse for Citizen Science Projects Zooniverse (www.zooniverse.org) is a platform supporting citizen science, where contributors can choose tasks they want to solve from among many different projects [121] with diverse task types and related tutorials of various complexity [110]. The Zooniverse platform allows researchers to tap into the impressive potential of volunteers, and in this way process a massive amount of data [15,35] in various projects, for example through annotation [118] or transcription [46], especially supported by ML. Volunteer activity on this platform follows different patterns, which are also influenced by the design of the platform and the crowdsourcing projects. Task division within crowdsourcing projects may have an impact on the quality of contributions, as quality may improve when tasks get bigger [17]. Volunteers range from ones who contribute a lot in many projects or in a single project to those who contribute once, and then never again [58], depending on their motivations [40] and their approach to making contributions [56], for example treating them like a leisure activity. There is some research about the participation of teenagers and younger adults in citizen science activities [50], but very little on the participation patterns of older adults, apart from a self-reported blog survey [131].

Fig. 3. Cards with descriptions and depictions of the Zooniverse projects from the "Into the Zooniverse Vol. II" publication. These cards were used to select the projects to represent in my research.

As there are multiple Zooniverse projects active on the platform at all times, the projects represented in my work were selected from among 40 Zooniverse community-chosen research projects. These projects were active on the platform in the 2019-20 academic year and spotlit in the publication "Into the Zooniverse Vol II" [7] published on the 17th of November 2020 (see Figure 3).

[7] The book is available for download here: https://blog.zooniverse.org/2020/11/17/into-the-zooniverse-vol-ii-now-available/

2.3 Older Adults as Potential Volunteers and Contributors

Older adults are a very non-homogeneous group, the most commonly accepted definition places them at 65+ years of age, yet the ageing process is unique for every individual and chronological age may not reflect biological age [41]. In general, older adults' successful ageing is related to their mental well-being, which acts as a protective factor in the onset of neurodegenerative disorders and social engagement (both direct and online) [87]. Encouraging older adults to engage in online crowdsourcing is a worthwhile research topic and practice goal. It not only gives them the potential to form social bonds, provides of a form of cognitive training, but also may protect their mental well-being thanks to creating a positive image of themselves as successful contributors.

Moreover, older adults have proven to be valuable, dedicated and diligent contributors [119,90] especially in the projects built specifically with their participation in mind, despite still being discouraged by various ICT-based and other barriers [126,68].

While some older adults lack confidence in their ICT skills [114] there are multiple ways to mitigate this problem. For example, support and positive social reinforcement [100], and intergenerational interaction [70,71] can encourage older adults even to become co-designers, who directly contribute to technological solutions [100]. Addressing held stereotypes is also important, as older adults hold beliefs about their age groups which may affect their performance and motivation - so they ought to be encouraged to trust their experience, abilities and skills [66].

Older adults' volunteering benefits the society [92] in general, but also has marked benefits for different communities [106], but also the population, as it increases their well-being [93,42] as well as health [82], both mental and physical. Such active ageing activities can be seen as a protective factor for their psychological wellness [42,47] helping address depression [82] and delaying the onset of age-related issues, such as mild cognitive decline [75]. Older adults' chance of developing such age-related issues as mild cognitive impairment and Alzheimer's disease seems to be connected to their motivational reserve [34], which moderates their engagement in other activities. Older adults eagerly volunteer to engage in the roles of keepers of history, or "organizational firekeepers" [23] and multiple other social projects offline, which shows this groups' great potential to engage in similar online activities, especially if the offline ones become less accessible to them, for example due to declining health or the epidemic situation.

2.4 Designing Interfaces for Older Adults

There are many general works on designing services and interfaces for older adults [27] yet, in the context of novel interfaces, they may offer insights rather than guidelines. General insights are related to such features of ageing as lower dexterity, hearing, vision and changes in cognitive functioning [95]. As a broad starting point, Pan et al. [103,102] offered a design framework to facilitate older adults' technology use which consisted of "symbolic familiarity, cultural familiarity, and actionable familiarity". The design framework pointed to the need to explore non-traditional, that is not PC-based, computer scenarios for crowdsourcing, based instead on familiar interaction paradigms, such as TV-based and chat-based crowdsourcing, which was done as part of this work.

Pak's "Designing Displays for Older Adults" is a premier source on the subject and offers insights related to the need to provide broad context and consistency to offset older adults' low tolerance of cluttered designs [101] due to slower visual search. The need to simplify visual design is also echoed by Silva et al. in a study on a TV dance application, where older adults using it may have experienced an information overload due to visual clutter [120]. Older adults also seem to be quite comfortable using tablets [70] which may be because of their larger screens, when compared with phones [130]. Smart TVs then are an interesting area of inquiry in terms of application design, as they offer a comfortable and familiar experience of interacting with the use of a remote control, which may solve some of the problems explored by Fisk et al. and related to "mapping actions to devices" [32]. There were early studies on Smart TVs done with older adults in Living Labs [9] which explored their feasibility for everyday use. More explicitly, a study by Nunes et al. recommends limiting steps in navigation, clearly marking choices, using the center of the screen for key information as well as simple language with long display times of text [98]. Moreover, there are general guidelines coming from older research on interactive TV services [22], but these need to be updated to allow for interaction with more advanced apps, such as those built for Android-powered Smart TVs.

Voice interfaces also are a promising avenue of research in the context of older adults [76] as they lessen the cognitive burden of interacting with computers [91] due to the natural speech-mediated communication. Older adults value conversational interaction due to its social aspect and naturalness [67], which can still be challenging [89]. General design guidelines show that chatbots ought to be made as "tools, toys and friends" [18] and evaluated on four dimensions [91] of "conscientiousness, originality, manners and thoroughness" while minding the key elements of successful human-human conversations [24], such as their social dimensions [8]. Increasing users' acceptance of chatbots can be done, for example, by using proper personalized repair strategies [10] lowering users'

expectations [21]. Studies also indicate, that chatbots may have the potential to engage users in relationships that may affect their well-being [79], which is an important aspect in building sustainable crowdsourcing systems for older adults. Overall, conversational crowdsourcing could constitute a prominent contribution in the trend of "chatbots for social good" [145].

2.5 Designing Experiences for Older Adults

Highlighting older adults' strong suits, such as their lifelong experience [13,14], and knowledge of history, as in this example of a historical location-based game [70], can greatly empower older adults to engage with technology. Crystallized intelligence, which is based on experience is primarily measured by language skills and general world knowledge [12], shows evidence of not only staying constant despite age [136], but even increasing with age [86]. Another method is related to making older adults' aware of the benefits that technology can bring [11]. Proper empowerment can even encourage older adults to actively engage in software development process as experts [72] who can teach the researchers, and learn from them [7] at the same time. Empowerment like this may also allow older adults to become more successful crowdsourcing contributors.

When designing experiences for older adults it is necessary to keep in mind the effects that the natural ageing process can have on the brain, especially the decreased ability to acquire new information related to the deterioration of working memory. Even though working memory may deteriorate with age, older adults excel at building situation models, which lets them focus on meaning and aids in general comprehension of utterances [105]. In old age, together with the expected declines in cognitive resources related to the healthy ageing process comes the tendency to become more selective in choosing their activities [52]. According to Hess's "selective engagement theory" the increased fatigue connected to expending cognitive resources causes a decline in intrinsic motivation to engage in cognitively demanding activities [51]. Yet, these activities could have a positive effect on cognitive health [51]. Hence, experiences designed for older adults ought to provide a low barrier to entry. For example, these challenges, can be overcome by designing instructions in a way that target deficiencies in working memory [139]. Yet, many projects requiring the engagement of significant cognitive resources do not have support mechanisms [26] which would enable participants to approach them with confidence, engage in practice sessions without the risk of being wrong, or if needed, receive help from the community.

2.6 Older Adults and Crowdsourcing Systems

Research indicates that the average age of people engaging in crowdsourcing is between 20-30 years [68,109]. Consequently, there is not a lot of research regarding older adults' participation in crowdsourcing activities. They are a population that may be difficult to reach when it comes to conducting online research [134]. Among the most influential studies are Brewer's study about Mechanical Turk [19], even though the study was declarative and survey-based, without the actual experience of crowdsourcing. Some other studies offer activities that are stereotypically associated with older adults, such as historical photo tagging [142], proofreading scanned books [57,69] or sharing life experience via answering questions [54], but they do not comprehensively address or directly study other needs and opportunities touched upon in this literature review.

Older adults are more aware of their needs and skills [29] and they choose their activities more carefully; they also differ from younger adults in decision-making [49]. Both of these factors may contribute to older adults' generally lower participation in crowdsourcing activities. There are few studies considering older adults' motivation in crowdsourcing, but many study it in other contexts, for example of fitness [115,94], work [59] or learning [60]. Gamification is often considered together with crowdsourcing and linked with motivational paradigms [147] rooted in psychology, for example the self-determination theory [111], also verified with older adults [132]. Even though some gamification elements were tested in the context of crowdsourcing and older adults in the past [57] and recently [119], there are doubts whether such elements can foster the creation of a sustainable crowdsourcing experience [19]. Overall, the motivations offered by many crowdsourcing projects, especially related to small financial incentives [19], are incompatible with older adults' preferences, their need for intrinsic motivation [61] and individual aspirations [119]. The effectiveness of crowdsourcing systems and engagement with crowdsourcing projects relies, among other things, on the successful use of technology to break down crowdsourcing barriers, which are outlined in Table 1 in detail, to encourage sustainable contributions.

	WORKERS	PROJECTS	TASKS	SYSTEM
MOTIVATION	No explicit information on personal benefits. Task topic not relevant to own interest. [20,123]	Lack of clear target group benefiting from the work. [123] No feeling of community.	Tasks are too short for proper immersion. [123] Challenge level mismatched to skill. [124]	Little chance to find the platform. Lack of transparency of data storage and use.
COMMUNICATION	Lack of awareness of crowdsourcing opportunities. [123]	Unclear relevance of the project to own experience. [20]	Tasks don't showcase what can be learned from them.	Focus on the requesters and not on workers.
FEATURES	Uncomfortable setup for prolonged work.	Not sure who to contact. [26]	Quick expiration of tasks. Lack of ability to pause and easily resume.	Unsure where to start. [26] Unfamiliar interface. [96]
LEARNING	Self-stereotypes about own ICT-skills and knowledge. [72]	Insufficient information on broader context of the project.	No sandbox mode. [126] Lack of documentation and tutorials.	Inaccessible. [96] Unclear where and how to ask for help. [26]
FOLLOW-UP	No feedback after completing work.	Unclear communication of importance and usefulness to society. [19]	Lack of ability to review previous work.	No historical feedback on the changes of quality of one's work.

Table 1. Barriers to crowdsourcing, especially prominent for older adults [68,48,80,63,36,148,25,146,143]. Own elaboration; extended from the article on the AFFORCE framework [125].

2.7 Conclusions from State of the Art Overview

Existing crowdsourcing systems failed to engage and motivate older adults, as they did not provide a convenient, interesting and accessible crowdsourcing experience. Moreover, such systems are not designed in a way that would support ICT learning, community building or contributors' personal development. So far, the problem of designing crowdsourcing systems that would be well-aligned with the wishes and needs of older adults remained largely unexplored and this conclusion hinges on the following points:

1. There is a lack of hands-on studies of how older adults use crowdsourcing systems.
2. There is a lack of crowdsourcing systems that address barriers to older adults' engagement and crowdsourcing systems that make use of older adults' strengths.
3. There are no actionable guidelines for the design of crowdsourcing systems for older adults.

Hence, my work focuses on exploring these research opportunities and formulating actionable guidelines for researchers and practitioners designing crowdsourcing systems for older adults, as shown in the following chapter.

3 Contributions

In my thesis I explored and addressed the problem of designing crowdsourcing systems for older adults. I evaluated older adults' interaction with diverse existing crowdsourcing projects and tasks in a standardized remote study. I also proposed and verified novel crowdsourcing systems, relying on unique modes of interaction in non-traditional computer use scenarios. The cases for using Smart TV-based system and a chatbot system are presented; this choice of platforms and interfaces alone constitutes a significant change to existing crowdsourcing paradigms. Finally, I gathered the resulting insights as well as considerations from related studies at the intersection of older adults' use of ICT tools and crowdsourcing into an actionable crowdsourcing system design framework: AFFORCE.

3.1 Overview of Key Contributions

The main contribution is in the domain of **Informatics** is in the areas of **Information Systems for Crowdsourcing** and **Human-Centered Computing**.

The contributions to **Human-Centered Computing** in the area of interaction and user experience design are related to the verification of barriers and new interaction paradigms with these diverse crowdsourcing systems (see Sections 3.2.1 and 3.2.3). The insights from these aided in the formulation of the crowdsourcing system design framework: AFFORCE (see Section 3.2.5).

The contributions to the area of **Information Systems for Crowdsourcing** constitute the proposal and verification of novel interfaces, such as a Smart TV interface (see Sections 3.2.1 and 3.2.3), conversational crowdsourcing (see Section 3.2.2) and traditional web-based systems (see Section 3.2.4).

All of the aforementioned contributions relate to the problem of designing ICT solutions for the current and future older adults and serve to formulate the AFFORCE framework. Both the exploratory work leading up to the formulation of the AFFORCE framework, and the key elements of the framework are summarized in the following sections of this chapter.

The published articles describing these contributions in detail [123,124,125,126,127] are attached in full to this dissertation and follow in the final chapter (see Section 5).

3.2 Contributions to Informatics

Due to the iterative nature of research in the area of HCI the work on the first two research objectives listed below was conducted simultaneously:

Objective 1: ***Analysed older adults' interaction with diverse crowdsourcing systems, including ICT-related barriers to participation and motivational paradigms especially prominent for older adults***

Objective 2: ***Developed, designed and verified novel crowdsourcing interfaces for older adults***

3.2.1 Smart TV Interface for Detecting Errors in Video Subtitles: First, based on the needs of the TEDx crowdsourcing community (see Section 2.2.1) and the possible match of older adult's skills and preferences (see Section 2.3) a working crowdsourcing interface for Android TV was designed and developed. It made use of familiar patterns of interaction with a TV set via a remote control. The crowdsourcing content benefited from older adults' language-related contributions strengthened by their extensive linguistic and life experience.

The resulting "Dream TV" application allowed users to detect errors in same-language (transcription) and foreign-language (translation) subtitles for TED, TEDx and TED-Ed videos. The application allowed us to keep track of contributors' actions, filter videos by language combinations, choose the interface type (either regular or advanced) and choose test videos. The users could watch the videos, detect errors, read the full script as well as keep track of the videos they have watched and their progress (see Figure 4).

These subtitles were created by the volunteer TED Translators' community via the Amara subtitling platform and synchronized via Amara API. The subtitles to selected videos, as they were created by non-professionals, contained multiple errors with a different saturation depending on the case. They were either organic, introduced for the sake of this research based on previously identified common error types in subtitles or introduced by Machine Translation systems. The prepared videos varied in terms of topic, length, source language (spoken) as well as jargon, reading speed and having the speaker on/off the frame.

The aim of this study was to evaluate the potential of Smart TVs as crowdsourcing platforms and verify and test the design of this specific solution with older adults.

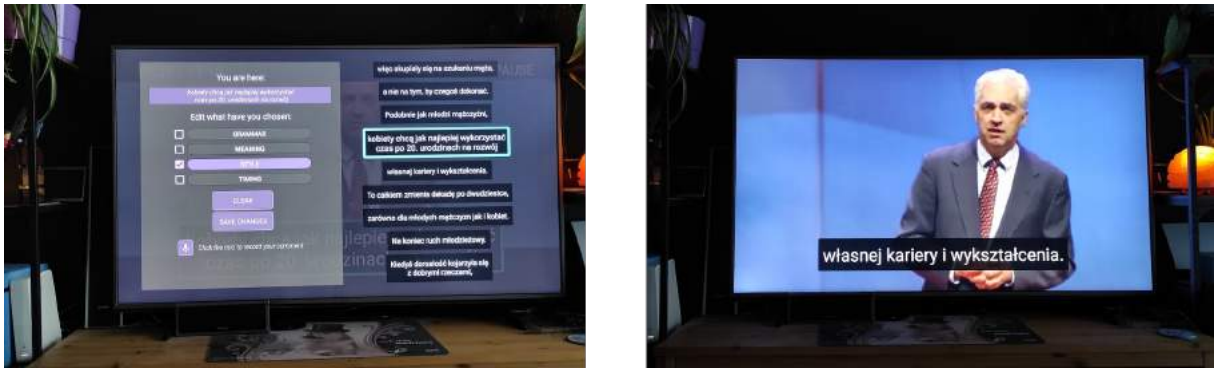


Fig. 4. The unique setup of TV-enabled crowdsourcing system "DreamTV" allowed the participants to engage with crowdsourcing tasks from the comfort of their sofa, with a remote control. They could mark an error in subtitles (left) and then resume watching the video pursuing their interests (right).

Based on the outcome of this exploratory research I formulated a set of preliminary considerations for designing crowdsourcing platforms, with a special focus on TV-based systems, related to the following areas:

1. Autonomy and freedom of choice (e.g. selecting own interests among diverse topics, ability to choose task duration, self-chosen accuracy and intensity of contribution as it may be time-consuming to do well);
2. Physical comfort and familiarity (e.g. being able to easily take breaks, physically comfortable setup designed for longer engagement, using the remote is a familiar experience although buttons ought to be labelled and colour coded);
3. Cognitive comfort to aid comprehension (e.g. adjustable reading speed - especially as Machine-translated subtitles are not compressed, ability to view the context at any time, ability to go back easily to compensate for slower reflexes when clicking pause, adding a dictionary to help with new jargon to offset the changes due to the evolution of language and unfamiliarity with some knowledge domains);
4. Building confidence (e.g. highlighting older adults' experience and skills, providing a tutorial as well as a practice mode);
5. Introducing the edutainment value [8] (using wide-domain educational content of value to contributors, alignment with interests).

These considerations are elaborated on in the article "**Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality**" [126]. Moreover, it describes the application and the verification of the interface with older adults following a rigorous study protocol. The article was presented at **CSCW 2018 (CORE A)** and published in an ACM journal series "Proceedings of the ACM on Human-Computer Interaction" dedicated to research on the multiple aspects of the intersection between human factors and computing systems. The conclusion of the research was as follows:

[8] Edutainment is a contraction, which stands for educational entertainment.

"The Smart TV setup and the interface are promising in the field of research with older adults, especially due to the familiarity, large screen size, general accessibility and comfortable at-home setup. Crowdsourcing subtitle errors is also perceived as an interesting task, which can be very engrossing and motivating depending on how it is framed." [126]

3.2.2 Chatbot Interface for Editing Wikipedia Infobox Translations: Next, based on the needs of the Wikipedia community (see Section 2.2.2) and the skills of older adults (see Section 2.3) a micro-task system to verify machine translations in infobox migration from the English to the Polish version of Wikipedia was proposed. For this purpose, a conversational crowdsourcing interface for editing Wikipedia, allowing for verification of machine-translated article infobox content was designed. It is discussed at length in the article "**Conversational Crowdsourcing for Older Adults: a Wikipedia Chatbot Concept**" [124] presented at **ECSCW 2020 (CORE B)** and published in "Proceedings of the 18th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies – Exploratory Papers, Reports of the European Society for Socially Embedded Technologies".

The conversational interface was deemed as having great potential for simplifying interaction with complex systems. It mitigated barriers to entry, by mapping crowdsourcing interaction into natural conversations, without artificial conventions. The novel design of the chatbot inverted the human-chatbot interaction, by making the user the one who assisted the chatbot. The case study presented the solution not as a universal one, but a viable chat-based alternative to Wikipedia's graphical user interface, especially addressing the consistency of infoboxes, which are difficult to maintain.

In this study further key considerations, related to the design of voice-enabled crowdsourcing systems for older adults, on the example of "Gizmo" chatbot, were formulated and they concerned:

1. Possible crowdsourcing tasks fit for chat-based interaction (e.g. verification of information, classification of articles, tagging articles for clean-up)
2. Modes and patterns of initiating interaction with the users (e.g. timing, triggers, self-initiated interaction)
3. Recovery techniques (e.g. suggesting a break, assisting in task completion to provide cognitive closure, rephrasing requests, switching tasks)
4. Evaluation of successful engagement (e.g. engagement and disaffection indicators, sustainability)

5. Motivation and well-being of the users (e.g competition, sense of being useful, providing cognitive training)

The chatbot is an early voice in the "chatbots for social good" [33] movement as well as the emerging trend of emotive web, making interactions with diverse online services more accessible to users.

3.2.3 Smart TV Interface: Age-comparative Study: To explore the considerations related to this research objective, the Smart-TV interface was further verified in an age-comparative study with younger adults [127], who are heavy video users, more familiar with Android applications and subtitles and could contribute different insights. The results of this study were presented at the **INTERACT 2019 (CORE A)** conference under the title of "**A Comparative Study of Younger and Older Adults' Interaction with a Crowdsourcing Android TV App for Detecting Errors in TEDx Video Subtitles**" and published in the conference proceedings.

The study listed key differences in what motivates older adults to engage with crowdsourcing in relation to the view of younger adults. While the task was fun for both younger and older adults, the former treated it more like work and expected payment. On the other hand, older adults were shown to be a promising target for this type of crowdsourcing, as it not only provided them with new and interesting content, but also allowed them to stay mentally active and keep learning. They treated it as entertainment and cognitive practice, while marking fewer mistakes, but still spotting key ones, especially related to the meaning of phrases. Moreover, it seemed that older adults' approach to the task, especially as a volunteer activity, was more sustainable in the long-run. Both age groups had trouble finding synchronization errors, which indicates that this ought to be a separate task. Moreover, both groups seemed to find fewer errors the more they enjoyed the video. Both groups also paused the videos too late and had to go back to mark the errors.

Overall, this study highlighted opportunities and issues related to the interaction with the Android-TV interface. Interesting areas of differences in motivational paradigms and preferences regarding crowdsourcing interaction and supportive features were discovered between the participant groups of younger and older adults.

3.2.4 Diverse Web-based Citizen Science Tasks: To further evaluate older adults' prowess, motivation and engagement with different task types the interaction with various citizen-science tasks in the categories of image, audio, text and pattern recognition was examined in a remote study with older adults [123] entitled "**Older Adults' Motivation**

and Engagement with Diverse Crowdsourcing Citizen Science Tasks" which was presented at **INTERACT 2021**.

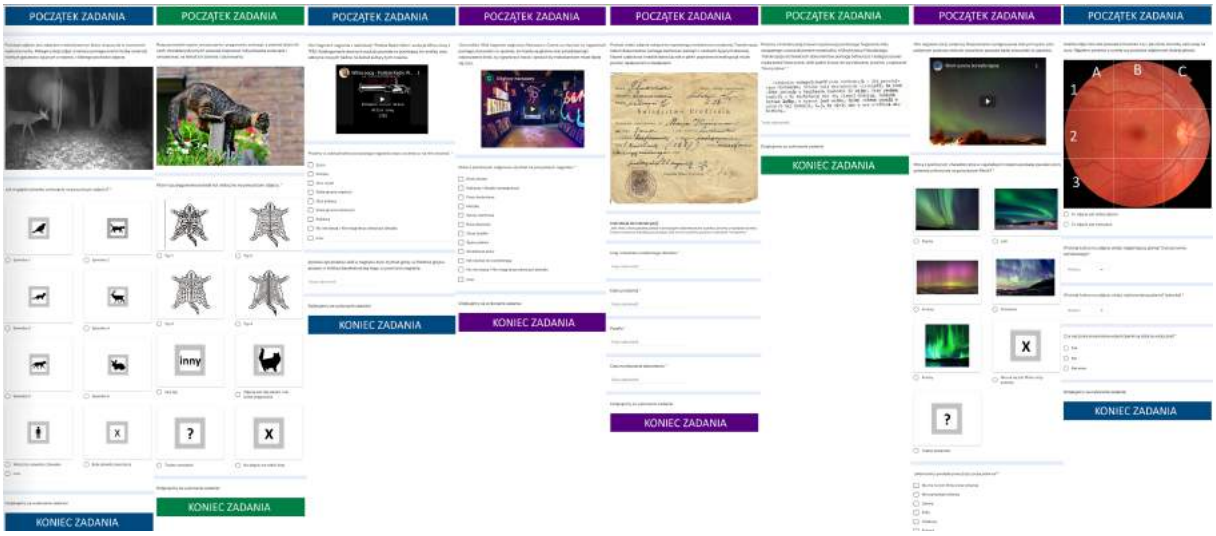


Fig. 5. The study participants solved these 8 diverse, yet standardized, microtasks, 2 in each category of image, audio, text and pattern recognition and after having solved each one evaluated it and suggested what could make it more engaging. This study was unique, as it touched on many diverse types of crowdsourcing citizen science tasks. Moreover, previous studies on the topic on older adults' engagement in crowdsourcing tasks in general did not require older adults to solve them, so they surveyed attitudes towards the idea, not practice of crowdsourcing.

In this remote study participants provided their socioeconomic information, explained their motivation to volunteer in general. Next, as part of the study, they completed, evaluated and provided suggestions regarding each crowdsourcing task, one by one as shown in Figure 5. Finally, they shared information about any previous crowdsourcing experience they may have had and disclosed what could encourage them to engage with similar tasks in the future.

Older adults in the study often wished for more context, higher difficulty and longer task duration - both because they were curious about the context, and because they wanted more additional data to verify their answers. It seems that micro-tasks, which are broken down for efficiency and to decrease the cost of task abandonment ought to be elaborated to become more satisfying, especially if they involve video, audio or image data. Also transcription tasks ought to be longer and involve the aspect of learning or the difficulty of guessing exact words because of low legibility to become engaging.

Based on the results of this study I formulated some considerations for the design of crowdsourcing tasks related to the preferences of older adults. First of all, tasks should be presented in a way that targets intrinsic motivation of learning something interesting. Next, they ought to explain the greater purpose they serve, so that contributors are aware

of how their work was useful. Finally, the tasks themselves could also be more complex, varied and elaborate, to pose an interesting challenge and provide learning opportunities.

Overall, despite lack of training and no financial incentive, older adults in the study provided very high quality contributions, even with tasks requiring longer transcription or listening to audio files. They also expressed willingness to do crowdsourcing tasks in their free time, so it seems that the first large barrier to their involvement in crowdsourcing is their lack of awareness of such opportunities. Thus, older adults were confirmed to be a very good target group of diverse crowdsourcing projects, provided that crowdsourcing systems meet their expectations, preferences and needs.

Objective 3: ***Formulated a framework for designing crowdsourcing systems for older adults***

3.2.5 AFFORCE Framework: Third, and key, contribution builds on all of my prior work in exploratory studies related to crowdsourcing [126,127,124,123], gained insights and explored considerations, both on crowdsourcing and in the area of older adults' ICT use [73,100,13,72,96] as well as a keyword-driven related works overview, and constitutes the proposal of the AFFORCE framework [9] (Fig 6). The article describing the framework was presented at **WI-IAT 2021 (CORE B)** under the title "**AFFORCE: Actionable Framework for Designing Crowdsourcing Experiences for Older Adults**". The framework is informed by identified barriers and dropout points for older adults using crowdsourcing systems and addresses them through mapping an actionable set of guidelines for designing crowdsourcing systems that could engage some groups of older adults in a sustainable way.

This contribution [125] incorporates key aspects of the crowdsourcing experience, such as: the design of the platform, tasks, user-empowerment, training, profiling and motivation as well as project discovery and communication, and gathers them in a comprehensive framework, as seen in Figure 6.

In short, crowdsourcing systems which aim to increase the number of contributions from older adults should allow easy **promotion (10)** of the projects, sharing links and clear calls to action, referring to appropriate motivations, on social media. The **platforms (1)** ought to be accessible, perhaps based on the principles of familiar interaction, without using unintuitive interaction paradigms. The tasks presented there need to be longer,

[9] The AFFORCE acronym was chosen to represent the framework as it stands for Actionable Framework FOR Crowdsourcing Experiences, while the word "afforce" itself according to the Collins dictionary means "to make stronger; consolidate; reinforce."

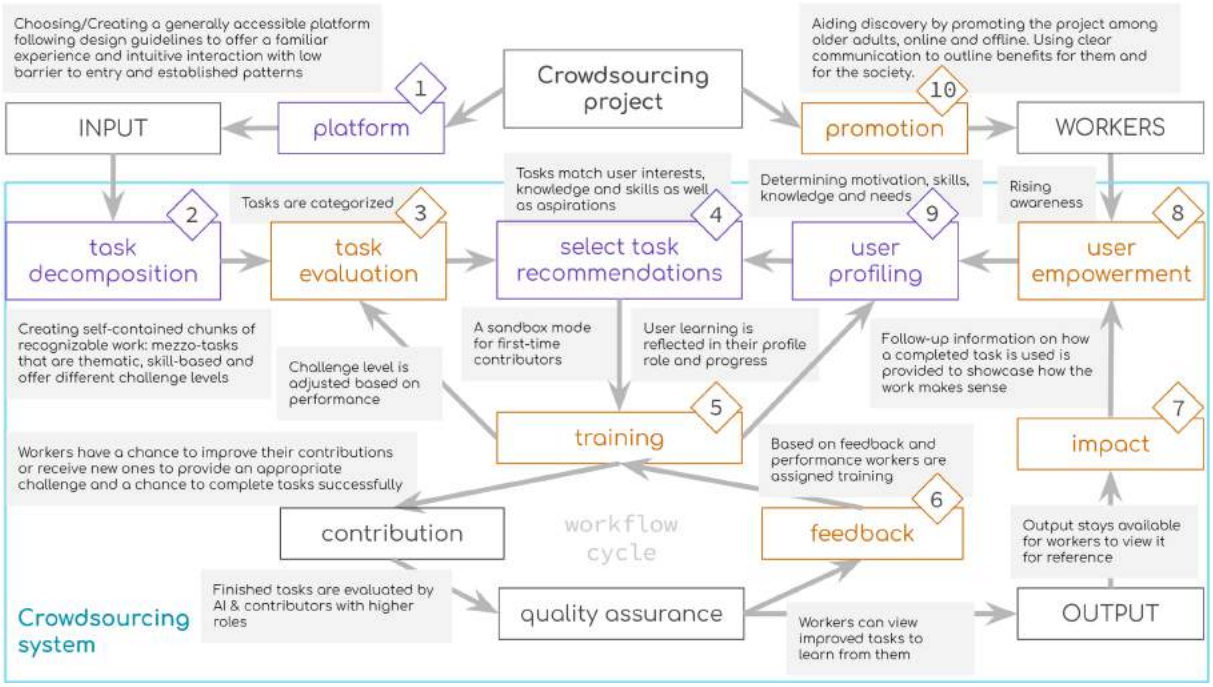


Fig. 6. This overview of framework elements presents key functionalities and the benefits of incorporating each of these modules into sustainable crowdsourcing systems. In relation to Kittur et al. [61] the elements in violet are changed, and in gold added. Own elaboration; appearing in the article on the AFFORCE framework [125].

challenging and self-contained so that they are understandable on their own. They should also provide optional context, more data and some educational content. Instead of micro-tasks they should be converted into **mezzo-tasks (2)**. These tasks then could be selected according to one's free time, skills, preferences regarding the topic, task intensity and manner of contributing. For this reason, they have to be **evaluated and appropriately categorized (3)**. Users, who are already **empowered (8)**, treated like experts, and **profiled (9)** should be able to select tasks from a shorter list of **task recommendations (4)** based on their preferences and skills. Once they select a new type of task they ought to receive **training (5)** and have access to a sandbox mode, where they can make mistakes without any consequences. That will allow them to also receive **feedback (6)** early, both automatic from the system and from their crowdsourcing peers. This should strengthen their sense of belonging to the online community – satisfying their need for social engagement. Such platform could spotlight the **impact (7)** their contributions have on the society.

Overall, the AFFORCE framework provides a set of actionable guidelines for practitioners designing crowdsourcing systems and researchers in the area of ICT for the aging population, on how to create more engaging crowdsourcing experiences for older adults. It is a step towards building crowdsourcing solutions more focused on the experiences of the crowd workers. AFFORCE aims to encourage more contributions from older adults to tap

into their massive potential and help them stay active for longer. This could pave way for more of them to use other ICT services, which may bridge the gap of digital divide.

3.3 Contributions to Other Disciplines

In the field of Psychology, my main contribution is to **Cognitive and Behavioural Psychology**, as the researched projects touch the problem of motivation and cognitive engagement in crowdsourcing tasks. The proposed crowdsourcing systems were used to evaluate various motivational schemes and incentives. Moreover, I explored and compared older and younger adults' motivation for engaging with crowdsourcing. In general, while younger adults saw solving crowdsourcing tasks as an opportunity for additional income, older adults thought of it mostly in terms of learning something interesting.

My contribution is also related to the area of **Psychology and Aging** as my research aims to explore opportunities for older adults' involvement in crowdsourcing tasks based on their strong points, such as their lifelong experience, knowledge and linguistic skills, especially in their native language. At the same time, the proposed AFFORCE framework aims to mitigate some of the barriers to crowdsourcing that may be especially prominent for older adults, as explored in my research and listed in Table 1 in Section 2.6. This approach to designing crowdsourcing solutions for older adults transforms the perception of the ageing process into a great opportunity instead of a challenge.

My work also involves **Human Factors Psychology** in the area of Human-Computer Interaction. The proposed AFFORCE framework explores and proposes actionable guidelines on how to enable and improve interaction with crowdsourcing interfaces. The framework takes into account both the crowdsourcing task workflow and the optimized user flow diagrams within the crowdsourcing platforms.

3.4 Limitations

One of the common arguments against extensive research on designing IT systems for older adults is that the people who will enter old age will have increasingly higher ICT skills, and therefore will not need any special design considerations. This argument, however, rests on the assumptions that technologies and ways of interacting with them will remain the same. New interaction metaphors are being introduced all the time. Recent examples include touchscreen gestures and controls, multi-touch, swiping or air-gestures. Even using new software effectively often requires one to learn new keyboard shortcuts. Thus, it is not a given that people who are "digital natives" now, will not feel that future IT systems are alien to them [95], especially if they continue to be designed by younger people without

older adults' participation [72]. Moreover, even despite advancements in medical sciences people's skills and cognitive abilities, such as working memory, will decline with age so the related considerations presented here as well as research methods will be, still, at least partly applicable and relevant.

Other limitations are related the nature of exploratory studies, that is involving fewer participants and being based on in-depth qualitative data. Also, the population of older adults is quite difficult to reach, especially online, while the older adults who volunteer to take part in research are a specific group, often more open to new experiences. What follows is that they may have different characteristics than the older adults from the general population. However, such crowdsourcing systems ought to be designed to cater to the needs of the groups, who are open to these kinds of activities – and may not be for everybody. Limitations specific to each study are described in detail in relevant articles [123,124,125,126,127], which constitute this thesis.

Overall, there is great potential for further research in the area of building crowdsourcing systems for older adults – especially when it comes to longitudinal studies of the proposed interfaces as well as systems which introduce the elements proposed in the AFFORCE framework. It will be interesting to explore how different combinations of the AFFORCE framework elements affect user experience for different age groups. It is also important to evaluate whether crowdsourcing systems designed with older adults in mind with the use of the AFFORCE framework, benefit from the curb-cut effect, that is increase usability for all age groups. Yet, this limitation also constitutes a reason why the AFFORCE framework was designed: to encourage and guide the creation of sustainable crowdsourcing systems that would provide a more positive crowdsourcing experience and encourage older adults to join crowdsourcing platforms in greater numbers.

3.5 Future Work

Nowadays there are multiple technological trends which open novel development directions for the area of crowdsourcing, and especially designing crowdsourcing systems for older adults. On the one hand, there is a trend for technology to mimic real life, with immersive virtual environments (IVR) copying or augmenting real reality (RR), and what follows, attempting to provide the most familiar mode of interaction, at the same time operating in the conditions of ecological validity. On the other hand, technology is seamlessly integrating into our daily lives and spaces we live in; to the point where interactions with it "in the wild" become almost second nature as smart homes and cities emerge, geolocation services are always on and smartphones enabling constant access to the Internet dominate the mobile market. These two trends, which exist on opposite ends of the

nature/reality-technology/virtuality spectrum, offer exciting crowdsourcing opportunities and promising ground for implementing diverse elements of the AFFORCE framework in VR and AR.

3.5.1 Crowdsourcing in VR Virtual reality constitutes one of the most promising interfaces for older adults [74] and a promising and growing area of academic research making use of IVR environments, as I argued in the article "All Factors Should Matter! Reference Checklist for Describing Research Conditions in Pursuit of Comparable IVR Experiments" [122]. The journey into virtual reality can quickly become both intuitive and immersive, as it often requires the understanding of very few metaphors to easily navigate. There are early attempts to introduce crowdsourcing in VR, especially in terms of finding the participants for scientific experiments. As a proof of concept, the VRChat application [112] and React VR framework [83] were used to run crowdsourced experiments on diverse populations. Games such as "Virtual Virtual Reality" [10] in a humorous way depict the potential of human contributors performing tasks in VR. Yet, the potential of VR for crowdsourcing is much greater and it includes not only novel ways of visualizing data and tasks or showing contextual information, but also completely new patterns of interaction with the tasks themselves. This could be especially beneficial for projects attempting to map the brain such as EyeWire [11] or Mozak [12], which ask users to build models of neurons in 3D and could benefit from the massive potential for manipulating 3D visualisations in IVR or AR environments. Better yet, for older adults specifically, the benefits of engaging with crowdsourcing tasks could go beyond socializing and staying cognitively active but could include physical fitness as certain physical movements could be required to choose or work out correct solutions.

3.5.2 Crowdsourcing "in the wild" Crowdsourcing can also be done in the wild, right there in the real world. Whenever there are potential users who have some time to spare, either taking a walk, waiting for public transportation, at a doctor's office or at an airport, there is the opportunity to use their time for good. Crowdsourcing systems could be integrated into advertisement "kiosks" or present in the AR Smart City landscape, thanks to the use of geolocation and AR-enabled smartphones. Crowdsourcing is increasingly used for gathering feedback on urban architecture such as its accessibility [104,113], also with the help of gamification techniques [88]. It has also been used in crisis mapping based on data obtained from social media, for example in the case of wildfires. The view of crowdsourcing activities, as something one can do in their leisure time for fun

[10] The page for the game in VR: https://www.oculus.com/experiences/quest/2111191602278229/
[11] The website of the EyeWire project: https://eyewire.org/explore
[12] The site explaining the Mozak project: https://www.mozak.science/landing

is gaining traction. There seems to be great potential for further research in combining crowdsourcing with mixed reality in general, for example via AR pass-through or smart glasses, in connection with real-life locations. By observing the evolution of Smart City landscapes and the development of personal IoT devices we may find new ways to engage with our environments to gather, process and use crowdsourced data for social good.

4 References

[1] World population aging 2015. Tech. rep., United Nations Department of Economic and Social Affairs (2015)

[2] Wikipedia statistics (2018), https://en.wikipedia.org/wiki/Wikipedia:Statistics

[3] Ageing Europe - statistics on working and moving into retirement (2019)

[4] World population ageing 2019. Tech. rep., United Nations (2019)

[5] Population structure and ageing (2020)

[6] Aassve, A., Cavalli, N., Mencarini, L., Plach, S., Bacci, M.L.: The covid-19 pandemic and human fertility. Science **369**(6502), 370–371 (2020). https://doi.org/10.1126/science.abc9520, https://www.science.org/doi/abs/10.1126/science.abc9520

[7] Abdullah Sani, Z., Mohd Nizam, D., Baharum, A.: Participatory design: Apps from the older adults to the older adults. Pertanika Journal of Science and Technology **29** (10 2021). https://doi.org/10.47836/pjst.29.4.35

[8] Aiello, L.M.: Ten social dimensions of conversations. Medium (Apr 2020), https://medium.com/socialdynamics/ten-social-dimensions-of-conversations-155200fa64f6

[9] Alaoui, M., Lewkowicz, M.: Practical issues related to the implication of elderlies in the design process – the case of a living lab approach for designing and evaluating social tv services. IRBM **36**(5), 259–265 (2015). https://doi.org/10.1016/j.irbm.2015.06.002, http://www.sciencedirect.com/science/article/pii/S195903181500069X

[10] Ashktorab, Z., Jain, M., Liao, Q.V., Weisz, J.D.: Resilient chatbots: Repair strategy preferences for conversational breakdowns. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 254:1–254:12. CHI '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300484, http://doi.acm.org/10.1145/3290605.3300484

[11] Aula, A.: Learning to use computers at a later age. In: HCI and the Older Population, pp. 3–5. University of Glasgow, Leeds, UK (Sep 2004)

[12] Baghaei, P., Tabatabaee, M.: The c-test: An integrative measure of crystallized intelligence. Journal of Intelligence **3**(2), 46–58 (2015)

[13] Balcerzak, B., Kopeć, W., Nielek, R., Kruk, S., Warpechowski, K., Wasik, M., Węgrzyn, M.: Press F1 for help: participatory design for dealing with on-line and real life security of older adults. In: Computer Sciences and Information Technologies (CSIT), 2017 12th International Scientific and Technical Conference on. vol. 1, pp. 240–243. IEEE (2017)

[14] Balcerzak, B., Kopec, W., Nielek, R., Warpechowski, K., Czajka, A.: From close the door to do not click and back. security by design for older adults. In: Conference on Computer Science and Information Technologies. pp. 40–53. Springer (2017)

[15] Barber, S.T.: The Zooniverse is expanding: Crowdsourced solutions to the hidden collections problem and the rise of the revolutionary cataloging interface. Journal of Library Metadata **18**(2), 85–111 (2018). https://doi.org/10.1080/19386389.2018.1489449

[16] Barber, S.J.: The applied implications of age-based stereotype threat for older adults. Journal of Applied Research in Memory and Cognition **9**(3), 274–285 (2020). https://doi.org/10.1016/j.jarmac.2020.05.002, https://doi.org/10.1016/j.jarmac.2020.05.002

[17] Barbosu, S., Gans, J.: Storm crowds: Evidence from Zooniverse on crowd contribution design. Research Policy **51**, 104414 (01 2022). https://doi.org/10.1016/j.respol.2021.104414

[18] Brandtzaeg, P.B., Følstad, A.: Why people use chatbots. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2017). https://doi.org/10.1007/978-3-319-70284-1_30

[19] Brewer, R., Morris, M.R., Piper, A.M.: "why would anybody do this?": Understanding older adults' motivations and challenges in crowd work. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. p. 2246–2257. CHI '16, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2858036.2858198

[20] Brewer, R., Piper, A.M.: "tell it like it really is": A case of online content creation and sharing among older adult bloggers. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 5529–5542. CHI '16, ACM, New York, NY, USA (2016). https://doi.org/10.1145/2858036.2858379, http://doi.acm.org/10.1145/2858036.2858379

[21] Broadbent, E., Stafford, R., MacDonald, B.: Acceptance of Healthcare Robots for the Older Population: Review and Future Directions. International Journal of Social Robotics **1**(4), 319 (Oct 2009). https://doi.org/10.1007/s12369-009-0030-6, https://doi.org/10.1007/s12369-009-0030-6

[22] Carmichael, A.: Style guide for the design of interactive television services for elderly viewers (1999)

[23] Carroll, J., Convertino, G., Farooq, U., Rosson, M.B.: The firekeepers: Aging considered as a resource. Universal Access in the Information Society **11**(1), 7–15 (03 2012). https://doi.org/10.1007/s10209-011-0229-9, https://doi.org/10.1007/s10209-011-0229-9

[24] Clark, L., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., Murad, C., Munteanu, C., Wade, V., Cowan, B.R.: What makes a good conversation?: Challenges in designing truly conversational agents. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 475:1–475:12. CHI '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300705, http://doi.acm.org/10.1145/3290605.3300705

[25] Culbertson, G., Shen, S., Andersen, E., Jung, M.: Have your cake and eat it too: Foreign language learning with a crowdsourced video captioning system. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. pp. 286–296. ACM (2017)

[26] Davidson, J.L., Mannan, U.A., Naik, R., Dua, I., Jensen, C.: Older adults and free/open source software: A diary study of first-time contributors. In: Proceedings of The International Symposium on Open Collaboration. p. 1–10. OpenSym '14, Association for Computing Machinery, New York, NY, USA (2014). https://doi.org/10.1145/2641580.2641589

[27] Dickinson, A., Arnott, J., Prior, S.: Methods for human-computer interaction research with older people. Behaviour & Information Technology **26**(4), 343–352 (2007)

[28] Dickinson, H., Fortson, L., Scarlata, C., Beck, M., Walmsley, M.: Modeling with the crowd: Optimizing the human-machine partnership with Zooniverse (03 2019)

[29] Djoub, Z.: ICT education and motivating elderly people. In: Ariadna; cultura, educación y tecnología. vol. 1, pp. 88–92 (Jul 2013). https://doi.org/http://dx.doi.org/10.6035/Ariadna.2013.1.15

[30] Estellés-Arolas, E., L. Guevara, F.G.: Towards an integrated crowdsourcing definition. Journal of Information Science **38** (04 2012). https://doi.org/10.1177/0165551512437638

[31] Etchegoyhen, T., Bywood, L., Fishel, M., Georgakopoulou, P., Jiang, J., Van Loenhout, G., Del Pozo, A., Maucec, M.S., Turner, A., Volk, M.: Machine translation for subtitling: A large-scale evaluation. In: LREC. pp. 46–53 (2014)

[32] Fisk, A., Czaja, S., Rogers, W., Charness, N., Sharit, J.: Designing for Older Adults: Principles and Creative Human Factors Approaches, Second Edition. Human Fac-

tors and Aging Series, CRC Press (2009)

[33] Følstad, A., Brandtzaeg, P.B., Feltwell, T., Law, E.L.C., Tscheligi, M., Luger, E.A.: Sig: Chatbots for social good. In: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. CHI EA '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3170427.3185372, https://doi.org/10.1145/3170427.3185372

[34] Forstmeier, S., Verena Thoma, M., Maercker, A.: Motivational Reserve: The Role of Motivational Processes in Cognitive Impairment and Alzheimer's Disease. Multiple Pathways of Cognitive Aging, Oxford University Press, New York (2021). https://doi.org/10.1093/oso/9780197528976.003.0007

[35] Fortson, L., Lynn, S.: Talking in the Zooniverse: A collaborative tool for citizen scientists. pp. 1–2 (05 2014). https://doi.org/10.1109/CTS.2014.6867533

[36] Franklin, A., Myneni, S.: Engagement and design barriers of mhealth applications for older adults. In: Proceedings of the Technology, Mind, and Society. TechMindSociety '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3183654.3183695

[37] Cámara de la Fuente, L.: Multilingual crowdsourcing motivation on global social media, case study: Ted otp. Sendebar **25**, 197–218 (2014)

[38] Gadiraju, U., Fetahu, B., Kawase, R., Siehndel, P., Dietze, S.: Using worker self-assessments for competence-based pre-selection in crowdsourcing microtasks. ACM Transactions on Computer-Human Interaction **24**, 1–26 (08 2017). https://doi.org/10.1145/3119930

[39] Gao, G., Pieczkiewicz, D., Kerr, M., Lindquist, R., Chi, C.L., Maganti, S., Austin, R., Kreitzer, M.J., Todd, K., Monsen, K.A.: Exploring older adults' strengths, problems, and wellbeing using de-identified electronic health record data. AMIA ... Annual Symposium proceedings. AMIA Symposium **2018**, 1263–1272 (Dec 2018), https://pubmed.ncbi.nlm.nih.gov/30815168, 30815168[pmid]

[40] Gay, P., Brown, S., Huang, A., Lehan, C.: Behaviors and motivations observed in the Zooniverse (01 2022)

[41] Gialluisi, A., Di Castelnuovo, A., Donati, M.B., de Gaetano, G., Iacoviello, L., , t.M.s.S.I.: Machine learning approaches for the estimation of biological aging: The road ahead for population studies. Frontiers in Medicine **6**, 146 (2019). https://doi.org/10.3389/fmed.2019.00146, https://www.frontiersin.org/article/10.3389/fmed.2019.00146

[42] Greenfield, E.A., Marks, N.F.: Formal volunteering as a protective factor for older adults' psychological well-being. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **59**(5), S258–S264 (2004)

[43] Halaweh, H., Dahlin-Ivanoff, S., Svantesson, U., Willén, C.: Perspectives of older adults on aging well: A focus group study. Journal of Aging Research **2018**, 9858252 (Nov 2018). https://doi.org/10.1155/2018/9858252, https://doi.org/10.1155/2018/9858252

[44] Halfaker, A., Geiger, R.S., Morgan, J.T., Riedl, J.: The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline. American Behavioral Scientist **57**(5), 664–688 (2013). https://doi.org/10.1177/0002764212469365, https://doi.org/10.1177/0002764212469365

[45] Han, L., Roitero, K., Gadiraju, U., Sarasua, C., Checco, A., Maddalena, E., Demartini, G.: The impact of task abandonment in crowdsourcing. IEEE Transactions on Knowledge and Data Engineering **33**(5), 2266–2279 (2021). https://doi.org/10.1109/TKDE.2019.2948168

[46] Hanson, D., Simenstad, A.: Combining human and machine transcriptions on the Zooniverse platform. In: Proceedings of the 2018 EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-generated Text. pp. 215–216. Association for Computational Linguistics, Brussels, Belgium (Nov 2018). https://doi.org/10.18653/v1/W18-6129, https://aclanthology.org/W18-6129

[47] Hao, Y.: Productive activities and psychological well-being among older adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **63**(2), S64–S72 (2008)

[48] Harley, D., Howland, K., Harris, E., Redlich, C.: Online communities for older users: What can we learn from local community interactions to create social sites that work for older people. In: Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014 - Sand, Sea and Sky - Holiday HCI. p. 42–51. BCS-HCI '14, BCS, Swindon, GBR (2014)

[49] von Helversen, B., Abramczuk, K., Kopeć, W., Nielek, R.: Influence of consumer reviews on online purchasing decisions in older and younger adults. Decision Support Systems (2018)

[50] Herodotou, C., Aristeidou, M., Miller, G., Ballard, H., Robinson, L.: What do we know about young volunteers? an exploratory study of participation in Zooniverse. Citizen Science: Theory and Practice **5** (01 2020). https://doi.org/10.5334/cstp.248

[51] Hess, T.M.: Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. Perspectives on psychological science : a journal of the Association for Psychological Science **9**(4), 388–407 (Jul 2014). https://doi.org/10.1177/1745691614527465, https://pubmed.ncbi.nlm.nih.gov/26173272, 26173272[pmid]

[52] Hess, T.M.: Subjective Perceptions of Cognitive Costs: Determinants and Impact on Motivation and Engagement. Multiple Pathways of Cognitive Aging, Oxford University Press, New York (2021). https://doi.org/10.1093/oso/9780197528976.003.0002

[53] Hesse, B.: Curb cuts in the virtual community: telework and persons with disabilities. In: Proceedings of the Twenty-Eighth Annual Hawaii International Conference on System Sciences. vol. 4, pp. 418–425 vol.4 (1995). https://doi.org/10.1109/HICSS.1995.375707

[54] Hiyama, A., Nagai, Y., Hirose, M., Kobayashi, M., Takagi, H.: Question first: Passive interaction model for gathering experience and knowledge from the elderly. In: Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on. pp. 151–156. IEEE (2013)

[55] Horn, J.L., Cattell, R.B.: Age differences in fluid and crystallized intelligence. Acta Psychologica **26**, 107–129 (1967). https://doi.org/https://doi.org/10.1016/0001-6918(67)90011-X, https://www.sciencedirect.com/science/article/pii/000169186790011X

[56] Ibrahim, K., Khodursky, S., Yasseri, T.: Gender imbalance and spatiotemporal patterns of contributions to citizen science projects: the case of Zooniverse (01 2021)

[57] Itoko, T., Arita, S., Kobayashi, M., Takagi, H.: Involving senior workers in crowdsourced proofreading. In: International Conference on Universal Access in Human-Computer Interaction. pp. 106–117. Springer (2014)

[58] Jackson, C., Østerlund, C., Maidel, V., Crowston, K., Mugar, G.: Which way did they go? newcomer movement through the Zooniverse. pp. 623–634 (02 2016). https://doi.org/10.1145/2818048.2835197

[59] Kanfer, R., Beier, M.E., Ackerman, P.L.: Goals and motivation related to work in later adulthood: An organizing framework. European Journal of Work and Organizational Psychology **22**(3), 253–264 (2013)

[60] Kim, A., Merriam, S.B.: Motivations for learning among older adults in a learning in retirement institute. Educational gerontology **30**(6), 441–455 (2004)

[61] Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., Horton, J.: The future of crowd work. In: Proceedings of the 2013 conference on Computer supported cooperative work. pp. 1301–1318. ACM (2013)

[62] Kittur, A., Smus, B., Khamkar, S., Kraut, R.E.: Crowdforge: Crowdsourcing complex work. In: Proceedings of the 24th annual ACM symposium on User interface software and technology. pp. 43–52 (2011)

[63] Knearem, T., Wang, X., Carroll, J.M.: Sustaining engagement in volunteer activities for older adults. In: Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing. p. 299–303. CSCW '20

Companion, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3406865.3418337

[64] Knowles, B., Hanson, V.L.: Older adults' deployment of 'distrust'. ACM Trans. Comput.-Hum. Interact. **25**(4), 21:1–21:25 (Aug 2018). https://doi.org/10.1145/3196490, http://doi.acm.org/10.1145/3196490

[65] Knowles, B., Hanson, V.L.: The wisdom of older technology (non)users. Commun. ACM **61**(3), 72–77 (Feb 2018). https://doi.org/10.1145/3179995, http://doi.acm.org/10.1145/3179995

[66] Knowles, B., Hanson, V.L., Rogers, Y., Piper, A.M., Waycott, J., Davies, N., Ambe, A., Brewer, R.N., Chattopadhyay, D., Dee, M., Frohlich, D., Gutierrez-Lopez, M., Jelen, B., Lazar, A., Nielek, R., Pena, B.B., Roper, A., Schlager, M., Schulte, B., Yuan, I.: Tthe harm in conflating aging with accessibility. In: Communications of the ACM. ACM (2020)

[67] Atay (née Knuepffer, C., Ireland, D., Liddle, J., Wiles, J., Vogel, A., Angus, D., Bradford, D., Campbell, A., Rushin, O., Chenery, H.: Can a smartphone-based chatbot engage older community group members? the impact of specialized content. Alzheimer's and Dementia **12**, P1005–P1006 (07 2016). https://doi.org/10.1016/j.jalz.2016.06.2070

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

[69] Kobayashi, M., Ishihara, T., Itoko, T., Takagi, H., Asakawa, C.: Age-based task specialization for crowdsourced proofreading. In: International Conference on Universal Access in Human-Computer Interaction. pp. 104–112. Springer (2013)

[70] Kopeć, W., Abramczuk, K., Balcerzak, B., Juźwin, M., Gniadzik, K., Kowalik, G., Nielek, R.: A location-based game for two generations: Teaching mobile technology to the elderly with the support of young volunteers. In: eHealth 360, pp. 84–91. Springer (2017)

[71] Kopeć, W., Balcerzak, B., Nielek, R., Kowalik, G., Wierzbicki, A., Casati, F.: Older adults and hackathons: A qualitative study. Empirical Software Engineering **23**(4), 1895–1930 (2018). https://doi.org/10.1007/s10664-017-9565-6, https://doi.org/10.1007/s10664-017-9565-6

[72] Kopeć, W., Nielek, R., Wierzbicki, A.: Guidelines towards better participation of older adults in software development processes using a new SPIRAL method and participatory approach. In: Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering. p. 49–56.

CHASE '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3195836.3195840

[73] Kopeć, W., Skorupska, K., Jaskulska, A., Abramczuk, K., Nielek, R., Wierzbicki, A.: LivingLab PJAIT: Towards better urban participation of seniors. In: Proceedings of the International Conference on Web Intelligence. pp. 1085–1092. WI '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3106426.3109040

[74] Kopeć, W., Wichrowski, M., Kalinowski, K., Jaskulska, A., Skorupska, K., Cnotkowski, D., Tyszka, J., Popieluch, A., Voitenkova, A., Masłyk, R., Gago, P., Krzywicki, M., Kornacka, M., Biele, C., Kobyliński, P., Kowalski, J., Abramczuk, K., Zdrodowska, A., Pochwatko, G., Możaryn, J., Marasek, K.: VR with older adults: Participatory design of a virtual ATM training simulation. IFAC-PapersOnLine **52**(19), 277–281 (2019). https://doi.org/https://doi.org/10.1016/j.ifacol.2019.12.110, https://www.sciencedirect.com/science/article/pii/S2405896319319457, 14th IFAC Symposium on Analysis, Design, and Evaluation of Human Machine Systems HMS 2019

[75] Kötteritzsch, A., Koch, M., Lemân, F.: Adaptive training for older adults based on dynamic diagnosis of mild cognitive impairments and dementia. In: International Workshop on Ambient Assisted Living. pp. 364–368. Springer (2014)

[76] Kowalski, J., Jaskulska, A., Skorupska, K., Abramczuk, K., Biele, C., Kopeć, W., Marasek, K.: Older adults and voice interaction: A pilot study with google home. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. pp. LBW0187:1–LBW0187:6. CHI EA '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3290607.3312973

[77] Lave, J., Wenger, E.: Situated Learning: Legitimate Peripheral Participation. Learning in Doing: Social, Cognitive and Computational Perspectives, Cambridge University Press (1991)

[78] Lee, J., Seo, D.: Crowdsourcing not all sourced by the crowd: An observation on the behavior of Wikipedia participants. Technovation (2016)

[79] Lee, M., Ackermans, S., van As, N., Chang, H., Lucas, E., IJsselsteijn, W.: Caring for vincent: A chatbot for self-compassion. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 702:1–702:13. CHI '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300932, http://doi.acm.org/10.1145/3290605.3300932

[80] Li, Q., Gajos, K.Z., Reinecke, K.: Volunteer-based online studies with older adults and people with disabilities. In: Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility. p. 229–241. AS-

SETS '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3234695.3236360

[81] Lockwood, P.L., Abdurahman, A., Gabay, A.S., Drew, D., Tamm, M., Husain, M., Apps, M.A.J.: Aging increases prosocial motivation for effort. Psychological Science **32**(5), 668–681 (2021). https://doi.org/10.1177/0956797620975781, https://doi.org/10.1177/0956797620975781, pMID: 33860711

[82] Lum, T.Y., Lightfoot, E.: The effects of volunteering on the physical and mental health of older people. Research on aging **27**(1), 31–55 (2005)

[83] Ma, X., Cackett, M., Park, L., Chien, E., Naaman, M.: Web-based VR experiments powered by the crowd. In: Proceedings of the 2018 World Wide Web Conference. p. 33–43. WWW '18, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE (2018). https://doi.org/10.1145/3178876.3186034, https://doi.org/10.1145/3178876.3186034

[84] Mandiberg, M.: Mapping Wikipedia - an unprecedented data set shows where the encyclopedia's editors are, where they aren't, and why. The Atlantic (Feb 2020), https://www.theatlantic.com/technology/archive/2020/02/where-Wikipedias-editors-are-where-they-arent-and-why/605023/

[85] Martin, W.F., Welsch, H.: Wasted talent: Battling exclusionary forces among senior entrepreneurs. The International Journal of Entrepreneurship and Innovation **20**(2), 130–143 (2019). https://doi.org/10.1177/1465750318809448, https://doi.org/10.1177/1465750318809448

[86] McArdle, J.J., Hamagami, F., Meredith, W., Bradway, K.P.: Modeling the dynamic hypotheses of Gf–Gc theory using longitudinal life-span data. Learning and Individual Differences **12**(1), 53–79 (2000)

[87] Mintzer, J., Donovan, K., Kindy, A., Lock, S., Chura, L., Barracca, N.: Lifestyle choices and brain health. Frontiers in Medicine **6** (10 2019). https://doi.org/10.3389/fmed.2019.00204

[88] Miyata, A., Okugawa, K., Yamato, Y., Maeda, T., Murayama, Y., Aibara, M., Furuichi, M., Murayama, Y.: A crowdsourcing platform for constructing accessibility maps supporting multiple participation modes. pp. 1–6 (05 2021). https://doi.org/10.1145/3411763.3451688

[89] Moore, R.J., Arar, R., Ren, G.J., Szymanski, M.H.: Conversational UX design. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. pp. 492–497 (2017)

[90] Moore, R., Hancock, J.: Older adults, social technologies, and the coronavirus pandemic: Challenges, strengths, and strategies for support. Social Media + Society **6**, 205630512094816 (07 2020). https://doi.org/10.1177/2056305120948162

[91] Morrissey, K., Kirakowski, J.: 'Realness' in chatbots: Establishing quantifiable criteria. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2013). https://doi.org/10.1007/978-3-642-39330-3_10

[92] Morrow-Howell, N.: Volunteering in later life: Research frontiers. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **65**(4), 461–469 (2010)

[93] Morrow-Howell, N., Hinterlong, J., Rozario, P.A., Tang, F.: Effects of volunteering on the well-being of older adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **58**(3), S137–S145 (2003)

[94] Navarro, J.E.J.B., Sanz, J.L.G., Del Castillo, J.M., Izquierdo, A.C., Rodríguez, M.M.: Motivational factors and physician advice for physical activity in older urban adults. Journal of aging and physical activity **15**(3), 241–256 (2007)

[95] Newell, A.F.: Design and the digital divide: Insights from 40 years in computer support for older and disabled people (2011), https://ieeexplore.ieee.org/document/6812886

[96] Nielek, R., Lutostańska, M., Kopeć, W., Wierzbicki, A.: Turned 70?: It is time to start editing Wikipedia. In: Proceedings of the International Conference on Web Intelligence. pp. 899–906. WI '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3106426.3106539, http://doi.acm.org/10.1145/3106426.3106539

[97] Nielek, R., Pawlowska, J., Rydzewska, K., Wierzbicki, A.: Adapting Algorithms on the Web to Deal With Cognitive Aging. Multiple Pathways of Cognitive Aging, Oxford University Press, New York (2021). https://doi.org/10.1093/oso/9780197528976.003.0016

[98] Nunes, F., Kerwin, M., Silva, P.A.: Design recommendations for TV user interfaces for older adults: Findings from the eCAALYX project. In: Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility. pp. 41–48. ASSETS '12, ACM, New York, NY, USA (2012). https://doi.org/10.1145/2384916.2384924, http://doi.acm.org/10.1145/2384916.2384924

[99] Ortman, J.M., Velkoff, V.A., Hogan, H., et al.: An aging nation: the older population in the United States. Washington, DC: US Census Bureau pp. 25–1140 (2014)

[100] Orzeszek, D., Kopec, W., Wichrowski, M., Nielek, R., Balcerzak, B., Kowalik, G., Puchalska-Kaminska, M.: Beyond participatory design: Towards a model for teaching seniors application design. CEUR Workshop Proceed (2017)

[101] Pak, R., McLaughlin, A.: Designing Displays for Older Adults. Human Factors and Aging Series, CRC Press (2010)

[102] Pan, Z.: The roles of familiarity design in active ageing. ArXiv **abs/1601.06869** (2016)

[103] Pan, Z., Miao, C., Yu, H., Leung, C., Chin, J.J.: The effects of familiarity design on the adoption of wellness games by the elderly. In: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). vol. 2, pp. 387–390. IEEE (2015)

[104] Prandi, C., Mirri, S., Ferretti, S., Salomoni, P.: On the need of trustworthy sensing and crowdsourcing for urban accessibility in Smart City. ACM Transactions on Internet Technology **18**, 1–21 (10 2017). https://doi.org/10.1145/3133327

[105] Radvansky, G.: Aging, memory, and comprehension. Current Directions in Psychological Science **8**, 49–53 (04 1999). https://doi.org/10.1111/1467-8721.00012

[106] Ranzijn, R.: The potential of older adults to enhance community quality of life: Links between positive psychology and productive aging. Ageing International **27**(2), 30–55 (Mar 2002). https://doi.org/10.1007/s12126-002-1001-5, https://doi.org/10.1007/s12126-002-1001-5

[107] Redi, M., Fetahu, B., Morgan, J., Taraborelli, D.: Citation needed: A taxonomy and algorithmic assessment of Wikipedia's verifiability. In: The World Wide Web Conference. pp. 1567–1578 (2019)

[108] Romero-Fresco, P., Pöchhacker, F.: Quality assessment in interlingual live subtitling: The NTR model **16**, 149–167 (01 2017)

[109] Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., Tomlinson, B.: Who are the crowdworkers? Shifting demographics in Mechanical Turk. In: CHI '10 Extended Abstracts on Human Factors in Computing Systems. p. 2863–2872. CHI EA '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1753846.1753873

[110] Rosser, H., Wiggins, A.: Tutorial designs and task types in Zooniverse. pp. 177–180 (10 2018). https://doi.org/10.1145/3272973.3274049

[111] Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American psychologist **55**(1), 68 (2000)

[112] Saffo, D., Di Bartolomeo, S., Yildirim, C., Dunne, C.: Crowdsourcing virtual reality experiments using VRChat (05 2020). https://doi.org/10.1145/3334480.3382829

[113] Saha, M., Saugstad, M., Maddali, H., Zeng, A., Holland, R., Bower, S., Dash, A., Chen, S., Li, A., Hara, K., Froehlich, J.: Project sidewalk: A web-based crowdsourcing tool for collecting sidewalk accessibility data at scale (05 2019). https://doi.org/10.1145/3290605.3300292

[114] Sandhu, J., Damodaran, L., Ramondt, L.: ICT skills acquisition by older people: Motivations for learning and barriers to progression. International Journal of Education and Ageing **3**(1), 25–42 (2013)

[115] Schutzer, K.A., Graves, B.S.: Barriers and motivations to exercise in older adults. Preventive medicine **39**(5), 1056–1061 (2004)

[116] Sedek, G., Hess, T., Touron, D. (eds.): Multiple Pathways of Cognitive Aging: Motivational and Contextual Influences. Oxford University Press, New York (2021). https://doi.org/10.1093/oso/9780197528976.001.0001, https://doi.org/10.1093/oso/9780197528976.001.0001

[117] Sedek, G., Verhaeghen, P., Lengsfeld, K., Rydzewska, K.: Using stories to assess linear reasoning abolishes the age-related differences found in formal tests. Thinking & Reasoning **27**(4), 623–633 (2021). https://doi.org/10.1080/13546783.2021.1911847

[118] Semenzin, C., Hamrick, L., Seidl, A., Kelleher, B., Cristia, A.: Describing vocalizations in young children: A big data approach through citizen science annotation. Journal of Speech, Language, and Hearing Research **64**, 1–16 (06 2021)

[119] Seong, E., Kim, S.: Designing a crowdsourcing system for the elderly: A gamified approach to speech collection. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–9. CHI EA '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3334480.3382999

[120] Silva, P.A., Nunes, F., Vasconcelos, A., Kerwin, M., Moutinho, R., Teixeira, P.: Using the smartphone accelerometer to monitor fall risk while playing a game: The design and usability evaluation of Dance! Don't fall **8027**, 754–763 (07 2013)

[121] Simpson, R., Page, K.R., De Roure, D.: Zooniverse: Observing the world's largest citizen science platform. In: Proceedings of the 23rd International Conference on World Wide Web. p. 1049–1054. WWW '14 Companion, Association for Computing Machinery, New York, NY, USA (2014). https://doi.org/10.1145/2567948.2579215

[122] Skorupska, K., Cnotkowski, D., Paluch, J., Masłyk, R., Jaskulska, A., Kornacka, M., Kopeć, W.: All factors should matter! Reference checklist for describing research conditions in pursuit of comparable IVR experiments (2021)

[123] Skorupska, K., Jaskulska, A., Masłyk, R., Paluch, J., Nielek, R., Kopeć, W.: Older adults' motivation and engagement with diverse crowdsourcing citizen science tasks. In: Ardito, C., Lanzilotti, R., Malizia, A., Petrie, H., Piccinno, A., Desolda, G., Inkpen, K. (eds.) Human-Computer Interaction – INTERACT 2021. pp. 93–103. Springer International Publishing, Cham (2021)

[124] Skorupska, K., Kamil, Warpechowski, Nielek, R., Kopeć, W.: Conversational crowdsourcing for older adults: a Wikipedia chatbot concept (2020)

[125] Skorupska, K., Nielek, R., Kopec, W.: AFFORCE: Actionable Framework for Designing Crowdsourcing Experiences for Older Adults. IEEE/WIC/ACM International Conference on Web Intelligence (WIIAT '21) (Dec 2021). https://doi.org/10.1145/3486622.3494026, https://doi.org/10.1145/3486622.3494026

[126] Skorupska, K., Núñez, M., Kopeć, W., Nielek, R.: Older adults and crowdsourcing: Android TV app for evaluating TEDx subtitle quality. Proc. ACM Hum.-Comput. Interact. **2**(CSCW), 159:1–159:23 (Nov 2018). https://doi.org/10.1145/3274428

[127] Skorupska, K., Núñez, M., Kopeć, W., Nielek, R.: A comparative study of younger and older adults' interaction with a crowdsourcing android TV app for detecting errors in TEDx video subtitles. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (eds.) Human-Computer Interaction – INTERACT 2019. pp. 455–464. Springer International Publishing, Cham (2019)

[128] Suzuki, Y., Nakamura, S.: Assessing the quality of Wikipedia editors through crowdsourcing. In: Proceedings of the 25th International Conference Companion on World Wide Web. pp. 1001–1006 (2016)

[129] Thelwall, M.: Does Astronomy research become too dated for the public? Wikipedia citations to Astronomy and Astrophysics journal articles 1996-2014. El Profesional de la Información **25**(6), 893 (Nov 2016). https://doi.org/10.3145/epi.2016.nov.06, https://recyt.fecyt.es/index.php/EPI/article/view/epi.2016.nov.06

[130] Tsai, H.y.S., Shillair, R., Cotten, S.R., Winstead, V., Yost, E.: Getting grandma online: Are tablets the answer for increasing digital inclusion for older adults in the US? Educational gerontology **41**(10), 695–709 (2015)

[131] Ttfnrob: Who are the Zooniverse community? We asked them... (Mar 2015), https://blog.zooniverse.org/2015/03/05/who-are-the-zooniverse-community-we-asked-them/

[132] Vallerand, R.J., O'Connor, B.P.: Motivation in the elderly: A theoretical framework and some promising findings. Canadian Psychology/Psychologie canadienne **30**(3), 538 (1989)

[133] Verhaeghen, P., Martin, M., Sędek, G.: Reconnecting cognition in the lab and cognition in real life: The role of compensatory social and motivational factors in explaining how cognition ages in the wild. Neuropsychology, development, and cognition. Section B, Aging, neuropsychology and cognition **19**(1-2), 1–12 (2012). https://doi.org/10.1080/13825585.2011.645009, https://pubmed.ncbi.nlm.nih.gov/22313173, 22313173[pmid]

[134] Verma, N., Shiroma, K., Rich, K., Fleischmann, K.R., Xie, B., Lee, M.K.: Conducting quantitative research with hard-to-reach-online populations: Using prime panels to rapidly survey older adults during a pandemic. In: Toeppe, K., Yan, H., Chu,

S.K.W. (eds.) Diversity, Divergence, Dialogue, vol. 12646, pp. 384–393. Springer International Publishing (2021)

[135] Vines, J., Pritchard, G., Wright, P., Olivier, P., Brittain, K.: An age-old problem: Examining the discourses of ageing in HCI and strategies for future research. ACM Trans. Comput.-Hum. Interact. **22**(1), 2:1–2:27 (Feb 2015). https://doi.org/10.1145/2696867, http://doi.acm.org/10.1145/2696867

[136] Wang, J.J., Kaufman, A.S.: Changes in fluid and crystallized intelligence across the 20-to 90-year age range on the K-BIT. Journal of Psychoeducational Assessment **11**(1), 29–37 (1993)

[137] Wang, X., Knearem, T., Carroll, J.M.: Ordinary innovations of older adults: An interview study to understand interest development in aging. In: Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing. p. 253–256. CSCW '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3272973.3274068

[138] Williams, D.: On and off the 'Net: Scales for Social Capital in an Online Era. Journal of Computer-Mediated Communication **11**(2), 593–628 (07 2017). https://doi.org/10.1111/j.1083-6101.2006.00029.x, https://doi.org/10.1111/j.1083-6101.2006.00029.x

[139] Wolfson, N.E., Cavanagh, T.M., Kraiger, K.: Older adults and technology-based instruction: Optimizing learning outcomes and transfer. Academy of Management Learning & Education **13**(1), 26–44 (2014)

[140] Wulczyn, E., West, R., Zia, L., Leskovec, J.: Growing Wikipedia Across Languages via Recommendation. Proceedings of the ... International World-Wide Web Conference. International WWW Conference **2016**, 975–985 (Apr 2016). https://doi.org/10.1145/2872427.2883077, https://pubmed.ncbi.nlm.nih.gov/27819073

[141] Yang, H.L., Lai, C.Y.: Motivations of Wikipedia content contributors. Computers in Human Behavior **26**(6), 1377–1383 (2010). https://doi.org/https://doi.org/10.1016/j.chb.2010.04.011

[142] Yu, H., Miao, C., Liu, S., Pan, Z., Khalid, N.S.B., Shen, Z., Leung, C.: Productive aging through intelligent personalized crowdsourcing. In: 30th AAAI Conference on Artificial Intelligence (AAAI-16) (2016)

[143] Yu, L., André, P., Kittur, A., Kraut, R.: A comparison of social, learning, and financial strategies on crowd engagement and output quality. In: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. pp. 967–978. ACM (2014)

[144] Zhang, Y., Paradis, T., Hou, L., Li, J., Zhang, J., Zheng, H.: Cross-lingual infobox alignment in Wikipedia using entity-attribute factor graph. In: International Se-

mantic Web Conference. pp. 745–760. Springer (2017)

[145] Zheng, Y., Zhao, K., Stylianou, A.C.: Building social capital in online communities: a perspective of information and system quality. In: Santana, M., Luftman, J.N., Vinze, A.S. (eds.) Sustainable IT Collaboration Around the Globe. 16th Americas Conference on Information Systems, AMCIS 2010, Lima, Peru, August 12-15, 2010. p. 481. Association for Information Systems (2010), http://aisel.aisnet.org/amcis2010/481

[146] Zhu, H., Dow, S.P., Kraut, R.E., Kittur, A.: Reviewing versus doing: Learning and performance in crowd assessment. In: Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. pp. 1445–1455. ACM (2014)

[147] Zichermann, G., Cunningham, C.: Gamification by design: Implementing game mechanics in web and mobile apps. " O'Reilly Media, Inc.", 1st edn. (2011)

[148] Zyskowski, K., Morris, M.R., Bigham, J.P., Gray, M.L., Kane, S.K.: Accessible crowdwork? Understanding the value in and challenge of microtask employment for people with disabilities. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing. p. 1682–1693. CSCW '15, Association for Computing Machinery, New York, NY, USA (2015). https://doi.org/10.1145/2675133.2675158

5 Articles comprising the thesis

5.1 Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality (CSCW 2018, 140 pts)

Title	Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality
Authors	Kinga Skorupska, Manuel Nunez, Wieslaw Kopec, and Radoslaw Nielek
Conference	ACM Conference on Computer Supported Cooperative Work [CSCW 2018] (CORE A, 140 points on the list of the Polish Ministry of Science and Higher Education, as accessed on 5.04.2022)
Published in	"The Proceedings of the ACM on Human Computer Interaction (HCI)" - a journal series for research relevant to multiple aspects of the intersection between human factors and computing systems.
Author Credit	I was the lead author of this article and my key contributions constituted the research design, application concept, study protocol design, UI and UX testing, analysis of common error types, preparation of research content, data gathering, data analysis, literature review, conclusions, manuscript drafting and revision. I have also presented the research outcomes described in this paper at the conference (in person).
Description	This paper describes the insights from an exploratory qualitative pilot study testing the feasibility of a solution that would encourage older adults to participate in online crowdsourcing tasks in a non-traditional computer scenario. Therefore, we developed an Android TV application using Amara API to retrieve subtitles for TEDx talks which allows the participants to detect and categorize errors to support the quality of the translation and transcription processes. It relies on the older adults' innate skills as long-time native language users and the motivating factors of this socially and personally beneficial task. The study allowed us to verify the underlying concept of using Smart TVs as interfaces for crowdsourcing, as well as possible barriers, including the interface, configuration issues, topics and the process itself. We have also assessed the older adults' interaction and engagement with this TV-enabled online crowdsourcing task and we are convinced that the design of our setup addresses some key barriers to crowdsourcing by older adults. It also validates avenues for further research in this area focused on such considerations as autonomy and freedom of choice, familiarity, physical and cognitive comfort as well as building confidence and the edutainment value.
Contribution and Impact	This study comprises the first implementation and verification of a crowdsourcing system for Smart TVs. As such, it has already made an impact in the area of crowdsourcing, especially related to crowdsourcing in alternative, non-traditional computer use scenarios. This contribution is, so far, cited in the following studies by researchers from other institutions: – Hettiachchi, Danula, et al. "" Hi! I am the Crowd Tasker" Crowdsourcing through Digital Voice Assistants." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020. – Postal, Manoela Rogofski Brum, and Rafael Rieder. "User Interface Evaluation Methods for Elderly: A Systematic Review." 2019 21st Symposium on Virtual and Augmented Reality (SVR). IEEE, 2019. – Seong, Eunjin, and Seungjun Kim. "Designing a Crowdsourcing System for the Elderly: A Gamified Approach to Speech Collection." Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. 2020. – M. Storer, Kevin, and Stacy M. Branham. "Deinstitutionalizing Independence: Discourses of Disability and Housing in Accessible Computing." The 23rd International ACM SIGACCESS Conference on Computers and Accessibility. 2021. – Verma, Nitin, et al. "Conducting Quantitative Research with Hard-To-Reach-Online Populations: Using Prime Panels to Rapidly Survey Older Adults During a Pandemic." Diversity, Divergence, Dialogue 12646 (2021): 384. – Antoine, Axel, et al. "Interaction Illustration Taxonomy: Classification of Styles and Techniques for Visually Representing Interaction Scenarios." Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021. – M. Storer, Kevin, and Stacy M. Branham. "Deinstitutionalizing Independence: Discourses of Disability and Housing in Accessible Computing." The 23rd International ACM SIGACCESS Conference on Computers and Accessibility. 2021. – Biele, Cezary. "Human Movements in Human-Computer Interaction (HCI)." (2022). – Kou, Ziyi, et al. "HC-COVID: A Hierarchical Crowdsource Knowledge Graph Approach to Explainable COVID-19 Misinformation Detection." Proceedings of the ACM on Human-Computer Interaction 6.GROUP (2022): 1-25.

Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality

KINGA SKORUPSKA, Polish-Japanese Academy of Information Technology, Poland
MANUEL NÚÑEZ, Polish-Japanese Academy of Information Technology, Poland
WIESŁAW KOPEĆ, Polish-Japanese Academy of Information Technology, Poland
RADOSLAW NIELEK, Polish-Japanese Academy of Information Technology, Poland

In this paper we describe the insights from an exploratory qualitative pilot study testing the feasibility of a solution that would encourage older adults to participate in online crowdsourcing tasks in a non-computer scenario. Therefore, we developed an Android TV application using Amara API to retrieve subtitles for TEDx talks which allows the participants to detect and categorize errors to support the quality of the translation and transcription processes. It relies on the older adults' innate skills as long-time native language users and the motivating factors of this socially and personally beneficial task. The study allowed us to verify the underlying concept of using Smart TVs as interfaces for crowdsourcing, as well as possible barriers, including the interface, configuration issues, topics and the process itself. We have also assessed the older adults' interaction and engagement with this TV-enabled online crowdsourcing task and we are convinced that the design of our setup addresses some key barriers to crowdsourcing by older adults. It also validates avenues for further research in this area focused on such considerations as autonomy and freedom of choice, familiarity, physical and cognitive comfort as well as building confidence and the edutainment value.

CCS Concepts: • **Information systems → Crowdsourcing**; • **Human-centered computing → Displays and imagers**; **Pointing devices**; **Empirical studies in HCI**; **Collaborative and social computing**; **Collaborative and social computing systems and tools**; *Graphical user interfaces*; *Empirical studies in accessibility*; • **Social and professional topics → Seniors**; • **Software and its engineering → Software prototyping**; *Requirements analysis*; Collaboration in software development;

Additional Key Words and Phrases: crowdsourcing, Smart TV, Android TV, older adults, application development, software engineering, social inclusion, volunteering, subtitling, edutainment

ACM Reference Format:
Kinga Skorupska, Manuel Núñez, Wiesław Kopeć, and Radoslaw Nielek. 2018. Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 159 (November 2018), 23 pages. https://doi.org/10.1145/3274428

1 INTRODUCTION

According to the latest population data published by Eurostat the share of older adults, defined as people aged 65+, is increasing in every member and candidate state. In European Union member states (EU-28) this share has risen by 2.4% between 2006 and 2016. Moreover, the long term 2015

Authors' addresses: Kinga Skorupska, Polish-Japanese Academy of Information Technology, 86 Koszykowa str. 02-008, Warsaw, Poland, kinga.skorupska@pja.edu.pl; Manuel Núñez, Polish-Japanese Academy of Information Technology, 86 Koszykowa str. 02-008, Warsaw, Poland, manuel.nunez@pja.edu.pl; Wiesław Kopeć, Polish-Japanese Academy of Information Technology, 86 Koszykowa str. 02-008, Warsaw, Poland, kopec@pja.edu.pl; Radoslaw Nielek, Polish-Japanese Academy of Information Technology, 86 Koszykowa str. 02-008, Warsaw, Poland, nielek@pja.edu.pl.

2573-0142/2018/11-ART159 $15.00
https://doi.org/10.1145/3274428

Table 1. Four groups of stakeholders

Group	Their stake
Older adults	as the target group for our solution and the crowdsourcers motivated by the edutainment aspect of the task and its positive impact
Translation communities	as direct beneficiaries of the insights from the errors detected and problems signaled by the crowdsourcers
General public	as the ones benefiting from the improved quality of subtitles
Developers	as the ones involved in addressing the barriers to crowdsourcing and benefiting from insights from our development process and tests

EUROPOP projection covering the time up to 2080 shows that this trend will continue, and people aged 65+ are expected to comprise 29.1% of the total EU-28 population by 2080, while latest data shows they already made up 19.2% of the population in 2016 [3]. The same is increasingly true of Western societies all over the world [2]. For example, according to the U.S. Census Bureau by the year 2050, the population aged 65+ in the United States will almost double, reaching over 20% of the society [44]. While the number of older adults is already significant, their potential remains largely untapped because of a shortage of adequate research-informed activities and programs allowing them to contribute to the society. Moreover, numerous studies confirm that developing sustainable solutions for older adults is still challenging. In particular, Knowles and Hanson underscore that there is still room for further research and establishing a more holistic approach to the problem, despite extensive progress in that field over the last 20 years [29, 30]. Access to a lot of available activities relies heavily on ICT skills and familiarity with computers and technology, as well as the accessibility of the solutions. Moreover, older adults are often not provided with adequate motivation to take part in these activities, e.g. crowdsourcing [11], as they are not challenging, fun or do not hold an explicit connection to real life.

It is our intention to address some of these concerns by designing and testing the feasibility of an engaging crowdsourcing solution for older adults that would encourage them to participate in crowdsourcing tasks. Therefore, we decided to assess older adults' interaction and engagement with an online crowdsourcing task relying on their innate language skills, at home in a non-computer scenario. For this purpose we launched a qualitative pilot study to build and test a language-focused solution and verify the underlying concept and barriers, including the interface, configuration issues, interaction, topics and the process itself.

In the course of the study, we developed and tested an Android TV application for crowdsourcing TED and TEDx subtitle errors that enables the participants to detect and categorize them. In this study we operate at the intersection of interests of four groups of potential stakeholders, depicted in Table 1.

Our assumptions, based on our previous work with older adults in a similar field [34], were that they enjoy using Android tablets [31] and therefore can learn to proficiently use Smart TVs, which have the benefits of relying on a familiar setting (TV set) and interfaces (e.g. Teletext), a simple remote, as well as sport a larger screen size. Furthermore, older adults are experienced users of their native language and can ascertain if utterances seem natural. It was our expectation that they would see improving the quality of subtitles for TED and TEDx talks as an interesting and motivating task due to its educational value.

Therefore, our study aims to explore a few research aspects. Our prime objective was to determine the potential of Smart TVs as crowdsourcing platforms for older adults. This necessitated to first answer the question of whether older adults interact with Smart TVs in a more familiar fashion

than with computers. Secondly, we wanted to ascertain if they consider themselves competent enough in their native language to detect errors in subtitles. And lastly, we examined if such a task is engaging for them, and what are the emergent barriers to their involvement, and possible ways to overcome them. We hope that researching these questions will help us draw conclusions on how to better include older adults in ICT-dependent crowdsourcing tasks, and in online activities with the use of some new, but familiar interfaces and to find promising avenues for further study. At the same time, we want to propose and verify the design and setup of online crowdsourcing translation process support based on commodity goods, in this case the Smart TV platform, and to mitigate the barriers that may exist for this type of activity, also related to vital challenges of crowd work systems raised at CSCW in recent years [15, 28, 61, 62].

The paper is organized as follows: first we present the related works on topics, such as older adults' empowerment, crowdsourcing and volunteering, subtitling TED and TEDx, designing for older adults and Android TV in particular; second, we describe our methods, including the design of the DreamTV crowdsourcing tool and the testing protocol; third the results section follows, providing a summary of our collected qualitative data; finally, we interpret our results in the discussion section, which is followed by our conclusions and future research plans.

2 RELATED WORK

2.1 Challenges of empowering older adults

There exist increasingly more studies at the intersection of HCI and aging, but as Vines et al. concluded in their discourse analysis, many focus on stereotypes related to health, socialization and technology [56] instead of exploring the aging process and looking for opportunities. One positive example is Carrol et al.'s study of older adults as "organizational firekeepers" [14] in roles as keepers of history, co-designers and members of intergenerational teams contributing valuable complementary skills. So, while it is true that some older adults do not have faith and confidence in their skills, especially if they are ICT related [51], such barriers can be partially mitigated to enable their valuable contributions. One way to do this is by introducing a positive social context and support [45]. While in our previous research this context was strongly based on direct social interaction between generations [31, 32], in our current study it is built on the clear social benefit of the opportunity to crowdsource edutainment content [48]: having a lasting positive effect on a large number of viewers using subtitles e.g. for learning languages and educating the general population. This translates into a strong positive motivator, as one of the most common needs expressed by older adults is the need to feel useful, help others and contribute to the common good in a social setting [11]. Additionally, it also uses one of the advantages that older adults hold over the younger generation: their experience, explored for example by Balcerzak et al. [9, 10]. Just as their vast knowledge of the cultural and historical context had a positive impact on their performance in a historical location based game [31], so can their long time experience with their native language act as an empowering factor, allowing them to feel confident, and competent enough to climb the ladder of ICT proficiency (a benefit they are aware of [7]); eventually even making them ready to join the ICT solution development process, as in the SPIRAL method [33]. On top of the question of feeling competent, there are specific physiological changes in the brain of older adults, which comprise of deterioration of working memory, and consequently, the ability to acquire new information [59]. But, this effect can be mitigated by creating step-by-step instructions for doing an ICT task in a workshop format [34], and specific instruction design elements targeting these problems [59]. At the same time, there is evidence that crystallized intelligence, which consists of general knowledge derived from experience, not only does not decline [58], but may benefit from aging [38]. This

may prove to be an asset in tasks related to the native language, e.g. subtitling, as crystallized intelligence is primarily measured through general world knowledge and language tasks [8].

2.2 Volunteering and crowdsourcing

Encouraging older adults' engagement in volunteering and crowdsourcing activities in general [39] benefits the society as a whole. Such activities have positive effects on older adults' well-being [40], as well as their mental and physical health [37]. Thus, they can be regarded as a protective factor for their psychological well-being [24, 25]. In particular, they can also pose certain health advantages, as staying active and learning new things can delay the onset of age-related issues, e.g. mild cognitive decline and some related aspects of dementia [35]. These benefits were confirmed in a study of Lum and Lightfoot, who showed that volunteering slows the negative effects of aging and helps to combat depression [37]. They may extend to online crowdsourcing tasks, as the mental benefits may be connected to the results of the study by Yang and Cheng-Yu [60], who found that Wikipedia editors are internally driven by self-concept motivation: their need to maintain a consistent and positive image of themselves, which they thus satisfy. This issue was also explored in previous studies on Wikipedia editing by older adults [42]. As such, engaging older adults in interesting tasks which have a high social value may be beneficial on all fronts.

Older adults have proven to be more aware of their needs and abilities than the younger generation [20]; therefore, they are more selective in choosing their activities. Moreover, they differ from the younger generation in their online behavior and decision-making [57], which, alongside their generally lower ICT skills, may explain how little interest they expressed in the Mechanical Turk platform populated by tedious and repetitive crowdsourcing tasks [11].

Volunteering and user engagement are often considered in terms of using gamification techniques, i.e. the use of game design elements in non-game contexts [18]. Some researchers claim that gamification is a more psychological than technological issue [63]. Many methods and gamification tools are based on sound psychological foundations, like the well-established self-determination theory [50], which was verified in the context of older adults [54]. Although there are some promising reports on using gamification elements in older adults crowdsourcing tasks [26], there are also more recent reports that show difficulties in this field [11].

2.3 Subtitling TED and TEDx talks

TED[1] is a nonprofit organization that organizes conferences on the topics of technology, entertainment and design. Its mission is to share "ideas worth spreading" with the wider audience through the talks recorded during its conferences. To popularize the format across the world, TED launched the TEDx[2] initiative of volunteer organized TED-like conferences around the globe. Since 2009, when this initiative was launched, the number of published TEDx talks in over 100 different languages has reached 100,000. These talks totaled a billion views in large thanks to the TED Open Translation Program, now rebranded as TED Translators[3]. Within the TED Translators initiative volunteers contribute subtitles to TED, TEDx and TED-Ed videos in 100+ languages, making them accessible to everyone.

[1]When TED first launched it was an invite-only conference combining Technology, Entertainment and Design (TED) which eventually became an annual event in California. In 2001 the conference was acquired by Chris Anderson, who in 2006 shared the recorded talks online, paving the way for the current popularity of TED as an edutainment platform. More information about TED is available online at: https://www.ted.com/about/our-organization/history-of-ted

[2]The TEDx programme allows volunteers to apply for a TEDx license to organize a TED-like conference in their local area. More information about TEDx is available online at: https://www.ted.com/about/programs-initiatives/tedx-program

[3]The TED Translators programme allows volunteers to create subtitles for TED, TEDx and TED-Ed talks on Amara. More information about this programme is available online at: https://www.ted.com/about/programs-initiatives/ted-translators

The TED Translators initiative is a true example of a *community of practice* which Leave and Wenger [36] describe as a group driven by common interest that collectively improves and gains knowledge. In 2014, Cámara, in a survey-based study [12], identified the need to contribute to the TED mission of "spreading ideas" as a key volunteering motivation of the TED Translators. As such, improving the quality of these subtitles is a promising example of a potentially engaging task, which is beneficial both socially (granting access to all) and individually (learning about interesting topics). As the subtitles are sourced in a large part from non-professional volunteers, who make mistakes that can be detected by experienced native speakers of the language, this makes the task doubly important and rewarding.

Additionally, within the TED Translators project, there exist various challenges, largely dependent on the language community. For example, not enough volunteers to satisfy the ever-growing demand for translations; as of April 17, 2018 there are 56,617 TEDx videos added to the TED team in Amara[4] [27], an in-browser subtitling platform used by TED for its TED Translators project. Moreover, as TED Translators employ a 3-step quality assurance process, there are not enough experienced reviewers (step 2) and Language Coordinators (step 3) who can correct, and approve the subtitles for the talks so that they can be published.

Thus, not only is there potential to introduce a fourth Quality Assurance step, or to facilitate steps 2 and 3, but also to employ Machine Transcription and Translation to grant the international audience access to the content of the videos that were not yet translated. The inclusion of Machine Translation (MT) in the human translation process, especially where quantity is important, is a clear path ahead as it can shorten the process by up to 40% [21]. This trend is also evident in the development avenues of market-leading Computer Assisted Translation (CAT) tools, such as MemoQ or Trados. There is clearly room for involving crowdsourcing and citizen science in the context of Machine Learning and Natural Language Processing of the subtitles e.g. through supervised learning processes. Increasing both the quantity and the quality of subtitles is important, because they provide access to a wide range of ideas, also scientific, to the hard of hearing, non-native speakers and the general public in an easy to digest, edutainment format. So, making them more accessible can be a positive motivator for the volunteers.

2.4 Designing for older adults and Android TV

Although older adults are a very non-uniform group in terms of their ICT-skills and needs, there are some general guidelines, also relevant in Smart TV solution design. In "Designing Displays for Older Adults", Pak noted the older adults' need for context, consistency and low working memory burden, and their low tolerance for design clutter [46]. A similar focus on the need to simplify is visible in the findings of the study by Silva et al. [52] concerning a TV dance application, where users likely experienced an information overload due too much information being visible on the screen when trying to repeat dance moves. More generally, Pan et al. [47] proposed a framework which consists of "symbolic familiarity, cultural familiarity, and actionable familiarity", all of which have positive effects on the adoption of technology by older adults. Also, as shown by our research involving an Android location-based game, [31] older adults are quite comfortable with tablets. According to Tsai et al. [53], this is due to the larger screen size than phones. Thus, as large screen size and design familiarity, consistency and simplicity are key considerations, for our exploratory study we decided to make use of the increasingly popular Smart TVs. Moreover, such TV sets can be controlled by a remote, which may mitigate older adults' possible problems with "mapping actions to devices" signaled by Fisk et al. [23]. So through focusing on Smart TVs, we not only deliver a

[4]Amara is a subtitling platform where people can volunteer to create subtitles for videos with the use of a streamlined online editor. More on Amara is available online at: https://amara.org

Fig. 1. The Xiaomi MiBox, a TV Set-top box (STB) with a simple remote can turn any TV with an HDMI socket into a Smart TV with the Android TV OS. (Source: Wikipedia, CC BY-SA 3.0 license)

large screen size, but also a familiar and simple experience of using a remote control and interacting with text on screen, not unlike Teletext. Some preliminary studies on using Smart TVs for older adults in the context of Living Labs had been initiated [5], however, this is still a largely unexplored area of research with few general studies on interfaces for TV apps for older adults, such as the one by Nunes et al., who among other things, recommended minimizing navigation steps, clearly marking selections, using the middle of the screen, simple language and allowing for enough time to read [43]. This encouraged us, based on our findings concerning older adults and Living Labs [34], to explore the new technological, but familiar, opportunities that Smart TVs present in the home environment, especially in the context of crowdsourcing. According to numerous analyses, there is still tremendous dominance of using TV sets over computer and mobile equipment, especially among older adults (defined as 65+). For example, Nielsen research on the US market shows that older adults spend over 50 hours a week in front of the TV, while fewer than 5 hours in front of the PC, including using video on PC (less than an hour a week).[5] Although there are numerous works on various aspects of designing services and interfaces for older adults [19], in particular for interactive television use-case scenario [13], the development of new trends of Smart TV, VOD (Video On Demand) and OTT (Over the Top) internet-based streaming services and applications, which all largely allow users to freely choose content to watch, led us to thoroughly rethink the service we planned in this crowdsourcing use case with older adults.

3 METHODS

Based on our research and experience, we developed a pilot of an exploratory study concept for using crowdsourcing without the computer in a TV-enabled setup, as well as the tool to support the subtitle translation and review processes: the DreamTV application. The findings from our previous work in this field in the Living Lab context [34, 41, 45] informed the design of our testing protocol, which presents subtitling as a task useful for a wide group of viewers (motivation), and introduces the users to its basic concepts, as well as verifies their understanding with an on-paper exercise (empowerment). We believe that the design of our study addresses key barriers to information technology and crowdsourcing by older adults, as presented in Table 2.

Here, it is important to note that subtitles have only started to become more common in Poland in the 1990s, mostly in cinemas and later online. The preferred TV format is using a voice-over translation, which is also a popular practice in Cambodia, Mongolia, Vietnam and some other East

[5]More information on TV viewing habits by age can be found online at: https://www.recode.net/2016/6/27/12041028/tv-hours-per-week-nielsen

Table 2. Overview of signaled barriers to crowdsourcing by older adults with proposed solutions inspired by literature.

Barriers	**Solutions**
Uncomfortable and costly setup [31, 51]	TV sets at home [5, 34]
Unfamiliar interfaces [7, 16, 17, 45]	Text and remote control [23]
Repetitive tasks [11]	Educational wide-domain TEDx videos [4]
Unclear social benefit [11]	Improving subtitles for all [12]
Unclear personal benefit [11, 56]	Learning, positive self-image [42, 60]
Unsocial nature of the task [55]	Part of the community [14, 41]

Fig. 2. The main screen of DreamTV allows the users to select new videos, or continue watching their previous ones.

European countries.[6] Unlike dubbing, which employs a cast of actors to recreate the soundtrack in the target language, voice-over translation usually employs the technique of using a single voice actor talking over a quieter original film soundtrack. This fact may form an additional barrier for the older adults' engagement, as our tests were conducted with Polish participants using Polish language subtitles.

3.1 DreamTV application

Therefore, based on our analysis of possible solutions to the aforementioned barriers and taking into account the needs of various translation communities, for the purpose of our exploratory study we developed the DreamTV application, which allows users to detect errors in subtitles. Its main features are described below.

3.1.1 Video selection. The application allows the users to choose the videos to play (Fig. 2). It remembers the previous choices and the progress of the user on each video, making it possible to resume the tasks. It keeps track of the users' contributions and generates statistics. The identified errors are written in a database, which is accessible online.

3.1.2 Playing the video mode. The application works as a regular player, so between choosing error categories the users can enjoy watching videos with subtitles (Fig. 3), as an edutainment activity [48].

[6]Voice-over translation is common in part as it is cheaper than dubbing, as it usually involves only one voice actor, and there is no need to mix the original soundtrack which just plays quiet in the background, but also because of force of habit.

Fig. 3. The video player screen with subtitles

3.1.3 Error detection and category choice dialog. Once the users spot a mistake in the subtitles appearing in the video mode they can click the middle round button on the remote (Fig. 1) to pause the video. The application then overlays the error detection and category selection dialog over the video; on the right the subtitle is visible in the context of the subtitles surrounding it (Fig. 4). The user can then select if this is the subtitle they meant to pause at, and then they can choose and save the appropriate error category for the selected subtitle, at the same time restarting the video playback from a little before it was paused.

3.1.4 Error categories. The error categories whose feasibility we decided to explore are based on key categories of errors in subtitles as extrapolated from OTPedia[7] (TED Translators' wiki with resources used as training and reference materials for translators). They are selected with the assumption of being easy to detect and interpret by viewers without much prior training.

- Timing: if the subtitle starts too early, or too late, or finishes too early in relation to the audio track, or if it disappears too quickly and it is impossible to read it - an important consideration when audio is present according to Armstrong [6].
- Grammar: if there is a grammatical mistake: either with the tense, form, ending, punctuation or spelling.
- Meaning: if there is difficulty understanding the meaning of the utterance or a suspected difference in meaning, especially in translation.
- Style: if the wrong register is used in relation to the topic of the talk, or if the wording of a phrase is awkward, a potential calque, or could be corrected.

These four main error categories are also more intuitive, unlike existing models of quality assessment of subtitles by professionals [49], and they focus on information which could be useful for reviewers to pinpoint the errors later. This is unlike existing studies of subtitle quality assessment that focus on error detection without the intent to make later corrections of the same texts easier, as they largely relate to live TV subtitling, e.g. by Ofcom[8]. It is worth noting, that the Android TV remote used in the study was equipped with a microphone that could be used to provide additional comments or hints for translators through an option available in the basic interface (called advanced) visible in Fig. 4. This interface also allows the users to choose either one main, or a few error types in each subtitle, and display the reading speed required for it.

[7]The wiki site of TED Translators features subtitling guides to style, workflow and most common considerations for subtitles, including errors and formatting issues. The wiki can be found online at: https://translations.ted.com/Portal:Main
[8]Ofcom is the communications regulator of the UK, operating under Communications Act 2003 to ensure great quality of communications services in the UK. More information on Ofcom can be found online at: https://www.ofcom.org.uk/

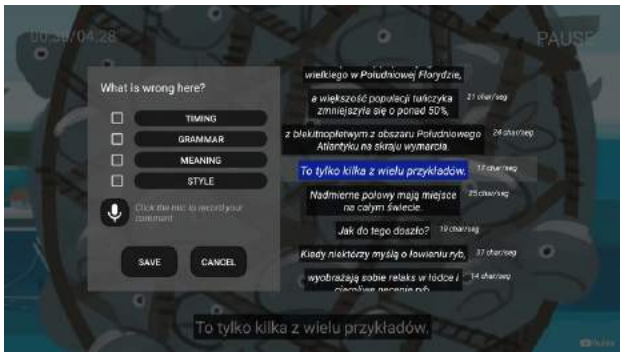


Fig. 4. The advanced dialog allows users not only to select error categories, but also to record voice comments and see the subtitle reading speed.

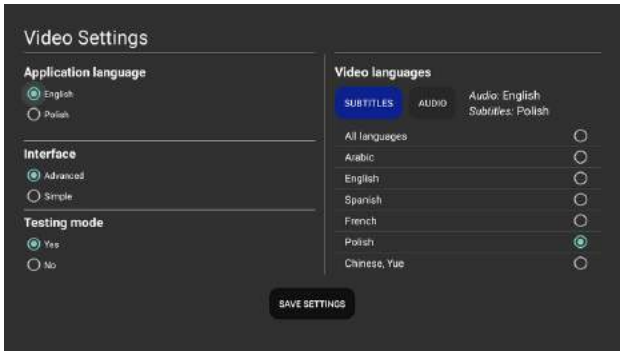


Fig. 5. The settings allow the user to switch from the simple to the advanced mode, as well as limit the choice to specific videos in the testing mode.

To carry out our tests, however, we have developed a simplified interface (called simple), which allows the users to mark only one error category. This was done to estimate if the participants would request more, or less complexity in the error detection dialog, and to keep the task middle-weight as a fusion of edutainment (postulated error detection only, without categorizations) and full correction crowdsourcing (existing basic interface allowing the users to record short messages, justifying their choices and offering corrections).

3.1.5 Settings. Apart from the choice between the simple and advanced interface mentioned above, it is possible to choose to work on videos based on the audio and subtitle language combinations, where choosing the same language for both would limit the videos to transcription tasks only. To aid in further study iterations we introduced the "testing mode" to limit the TEDx video selection to our predefined videos.

3.2 The journey of application development

The application development process was based on early usability tests, co-design and consultations with the TED Translators community members. It was also informed by insights concerning empowerment and engagement from our aforementioned Living Lab's activities centered around older adults. Based on these we scaled down our initial idea of editing faulty subtitles to error detection, which is more aligned with the need for simplicity, the edutainment aspect, and more suitable for control via the remote. Since then, the development of the application has focused on the unique way of interaction with the app using the simplest remote, as well as creating a nonintrusive interface to promote the fun aspect of the task. In line with Android TV guidelines,

Fig. 6. The at-home setup with a TV set facilitates the edutainment value of the task.

we designed a pop-up dialog system that appears on screen only when the users press the middle PLAY/PAUSE button, as this way they get the full screen experience without any clutter elements that could distract them. Some of the other early major changes included:

- Testing multiple combinations of the remote buttons.
- UX color coding, sizing (readability).
- UX flow and default behaviors (settings, auto-selection).
- Interface labeling and action defining for clarity.
- Changes to initial error categories (made the same for translation and transcription).
- Overlay view for subtitle context visibility.
- Subtitle skipping forward/backwards functionality (timing adjustment for the needs of older adults and consistency).

While the application was received favorably in our early tests, we decided to field test it using a structured protocol with older adults in the context that it was meant to be used (Living Lab): their own TV sets at home.

3.3 Study design, participants and testing protocol

As it was mentioned above, we decided to employ the distributed Living Lab approach. In particular, we invited older adults to participate in the pilot study supervised by the researchers. The research protocol, involving individual testing at home (Fig. 6), consists of the following main elements: an introduction to the study, the DigComp survey testing familiarity with computers and variety of ICT tasks, a semi-structured introductory interview that tests openness to new experiences as well as familiarity and preferences regarding subtitles, the explanation of the main elements of the project, that is the study and its benefits, an introduction to subtitles, and an on-paper exercise training error category detection skills, a demonstration and a hands-on test and the free interaction with the application and our five pre-selected and redacted test videos. The whole protocol takes about two hours to complete, and the reported preliminary tests were conducted in the first months of 2018.

3.3.1 Introduction. As the first step, we thanked the recruited participants for wanting to take part in this research and informed them that they can stop at any time and for any reason. We read the declaration of consent to them and asked them if they would like to sign it, and only then continued the research session.

3.3.2 DigComp based survey. The survey is a tool for measuring indicators of digital competence, developed by IPTS, and funded by the European Commission. It uses questions related to the Digital Skills Indicator[1] and broad ICT competence areas related to information (e.g. reading online news), communication (e.g. sending e-mails or making video calls), content creation (e.g. creating content, programming), safety (e.g. using anti-virus or a firewall) and problem solving (e.g. installing new devices, using Internet banking), based on the Digital Competence Framework [22].

3.3.3 Introductory interview. The semi-structured in depth introductory interview aimed to gauge two main issues. First, familiarity and attitude towards subtitles with questions related to cinema and viewing habits. Second, openness to new experiences, including questions about habits related to new experiences, especially educational, and key interests and topic preferences.

3.3.4 On-paper exercises. The paper exercises are a key component of our step-by-step introduction method for user empowerment as they allow the older adults to become familiar with key notions in subtitling. This includes an explanation of what subtitles are for, and who can benefit from them, as well as some technical information such as reading speed, maximum line length and best places to introduce line breaks, and finally the specific task of subtitle error categorization. The exercises consist of a selection of text snippets from our chosen test videos, each containing either one, or two errors. Towards the end of the task, the results are discussed with the participant to ensure that they understand the key principles of subtitling, as well as the error categories in the application. The test also allowed us to estimate the confidence of older adults when detecting language errors.

3.3.5 Using the DreamTV application with test videos. For our study we have selected five videos to represent different types of challenges. For this reason, these videos were controlled for the following characteristics:

General features:

- a) Topic.
- b) Video length.
- c) Source language (spoken).

Ease of comprehension:

- d) Jargon saturation.
- e) Required reading speed expressed in characters per second (ch/s).
- f) Presence of the speaker on-screen.

Types of errors:

- g) Error source (machine or human).
- h) Error saturation.
- i) Error category.

The videos selected for the preliminary study, visible in Table 3, allowed us to observe a variety of factors at play, in order to determine the most interesting areas of further inquiry. The machine translations were generated using the original human-made transcripts, which were then imported to SubtitleEdit[9]. This software has the option to use the Google Translate MT engine to generate subtitle translations; the generated lines then had their alignment fixed to reflect the transcript. The human natural errors are a combination of organic errors from the community, as well as errors introduced by researchers based on the most common translation/transcription errors lists on

[9]SubtitleEdit is a free open-source subtitle editor with a wealth of useful features which available online at: http://www.nikse.dk/subtitleedit/

Table 3. Characteristics of the videos used in the study. (Explanations: (a) Topic. b) Video length. c) Source language (spoken). d) Jargon saturation. e) Required reading speed expressed in characters per second (ch/s). f) Presence of the speaker on-screen. g) Error source (machine or human). h) Error saturation. i) Error category.)

No	Video	Characteristics
1	*Rebuilding Dreams One Bedroom at a time* by Joanna McCoy at TEDxKraków	a) Activism b) 10 min. c) **Polish** d) moderate modern e) medium (around 10–18 ch/s) f) on-screen g) human natural and simulated h) high-moderate (around 1 error /40 sec of video) i) mostly grammar with some timing
2	*General Einstein Theory of Relativity* by Krzysztof Meissner at TEDxMarszałkowska	a) Physics b) 11 min. c) **Polish** d) moderate science e) medium (around 11–18 ch/s) f) on-screen g) human natural and simulated h) moderate (around 1 error /1 minute of video) i) predominantly grammar
3	*Will the Ocean ever run out of Fish* by Ayana Elizabeth-Johnson and Jeniffer Jacquet	a) Environment b) 10 min. c) **English** d) high e) very fast (around 16–21 ch/s) f) off-screen narration over a cartoon g) human natural h) very low (around 1 error /2 minutes of video) i) mixed
4	*The hidden ways stairs shape your life* by David Rockwell	a) Culture b) 3 min. c) **English** d) low e) fast (around 13–21 ch/s) f) on-screen and off-screen g) Machine Translation with no corrections other than line alignment h) very high (around 1 error /10 seconds of video) i) mostly style and grammar
5	*Inventing is the easy part. Marketing takes work* by Daniel Schnitzer at TEDxPittsburgh	a) Travel and technology b) 5 min. 30 s. c) **English** d) low e) medium (between 11–18 ch/s) f) on-screen g) MT with human corrections for obvious mistakes h) high (around 1 error /30 sec. of video) i) mostly grammar, style and meaning

OTPedia[10] to introduce varying error saturation levels. The tests were done with Polish language subtitles, as they were conducted with older adults in Poland.

3.3.6 Exit IDIs. In the exit interviews, we asked, among other things, about the feelings towards the general task of finding mistakes in subtitles, and in particular about their attitudes towards different videos we have tested. We also asked for any suggestions for increasing the attractiveness of this task in general, as well as concerning the topics of the videos, the reading speed, the preparatory exercises and the controls (remote) and setup. Additionally, we asked the participants if they would like to continue this task on their own, if they enjoyed the language part of the task and if they would like to try more/different types of language-related tasks in the future.

4 RESULTS

Below we present the characteristics of the participants with the study summary, followed by some more detailed results.

4.1 Study group and research highlights

We invited seven older adults to our preliminary exploratory study: three female participants and four male participants. The study group was thoroughly selected from our Living Lab in Warsaw

[10]The TED Translators wiki features common error lists for many of the project's languages. To see the list for Polish refer to the materials available online at: https://translations.ted.com/Polish

to cover several conditions: different age, occupation (one active person before retirement, two retired but still professionally active, two retired for several years and two retired for more than fifteen years). All participants live in Warsaw, the capital city of Poland, and they are native Polish speakers with limited or very limited English skills. There was a 20 year age span: the youngest participant was 60 years old and the oldest one was 79, with a mean age of 70,85 (SD=6,87).

Based on introductory interviews and DigComp surveys we can describe this group as active users, above basic ICT skills. Everybody, except one, has smartphones, and the majority have Smart TVs. They are rather intensive and frequent Internet users, more than once a week.

Older adults in our study, even ones who report using the Internet only occasionally, and owning no smartphone, are very comfortable using Smart TVs at home. They do not see them as complex, but rather "interesting" (P4) and "useful" (P3), and "quite easy after some practicing" (P1, P2), some picked up on how to navigate them with a remote control within minutes (P1, P2, P3, P5), and all of them (P1–7) were able to learn how to perform the task in the DreamTV application in just one session. In case of P6 and P7, the learning curve was lower because their remote was similar to the one used in the study.

While all of the participants report that they rarely watch movies with subtitles, and prefer the voice over (P1–7) or dubbed content (P6), they see subtitles as "useful" (P5, P6, P7), as "cheaper and faster to make than voice over" (P2), and the task of improving them as "pleasurable and fun" (P3), "quite pleasant" (P1, P2) and "manageable" (P5). Most appreciate how "interesting" the videos are (P1, P3, P4, P6, P7), but one participant wished the topics were more "practical and useful for me" and said it was "not their thing" (P5). While the older adults were either good or very good at finding stylistic mistakes related to their natural language experience, especially in MT texts, they had trouble with technicalities such as timing, and small grammatical mistakes, like punctuation. Additionally, all of them complained about the reading speed required to understand the subtitles and sometimes paused the video to read the previous lines (P3, P4, P5, P6). While very few of the lines were above the 21 ch/s, which according to TED Translators' guidelines is the maximum reading speed, even the common industry standard of 17ch/s maximum was an issue.

4.2 Detailed results

4.2.1 On-paper exercises. The participants were very sure of their answers when taking the tests, and most of them were able to explain their reasoning for choosing one error category over another, or marking a few at once. However, there appeared to be a dilemma relating to ambiguity and an overlap in the categories perceived by the participants, which is typical, as one type of error can be easily related to others. For this reason, we provided the option to mark multiple errors in the basic interface of the application.

The biggest issue was with the subtitle stylistic conventions, such as correct subtitle division (e.g. not leaving linking words at the end of the line and breaking up syntactic wholes), or the written format of sound information.

4.2.2 Test videos. The participants were mostly very confident when marking mistakes, and in the videos which were more saturated with grammatical and stylistic mistakes (videos no. 2, 5), they detected many of them. The stylistic aspects connected to language conventions were easy to pick up for older adults, especially in video no. 5. However, in video no. 4, which had machine translated subtitles, the stylistic problems were so saturated, that the meaning of large parts was unclear (P6, P7) or the whole video was unclear to some study participants (P1, P2, P4), even when reading the dialog list (Fig. 4). Moreover, video no. 4, which was Machine Translated, did not have any compression done to the subtitles, which pushed the reading speed up. The dialog list visible on the subtitle error choice screen (Fig. 4) proved to be very useful for two reasons. The main

reason is that it was common that all of the study participants (P1–7) paused the video too late, and they had to use the dialog list to go back to the subtitle with the error, even reading it out loud (P3, P1, P6). Additionally, it helped comprehension as when the study participants felt they were getting lost, because of high reading speed required, or the jargon, they would pause the video and read the subtitles to make sure they understood the text (P3, P4). One surprise was that none of the study participants detected any timing errors in the videos. When asked about this, one participant replied that "they had to read" (P5).

While some study participants were indecisive about choosing the main error category (P4, P5) others (P3) were so eager to continue that they kept forgetting to mark it, naturally wanting to mark an error only. This correlated to how engaged they were in the topic of the video. The more interesting they found the video, the fewer mistakes they detected (P1–P4, P6). Some participants requested to introduce an in-app tutorial, which would allow them to mark errors "without the worry that they will make a mistake" (P1, P4). Some study participants found the reading aspect to be difficult because of impaired vision. However, they just sat closer to the TV set to mitigate the effect and continue the task (P1, P3), but for most this was not an issue.

4.2.3 Exit IDIs. The participants felt they were capable of detecting errors in Polish subtitles for the videos (P1–7), and one of them even said: "There should be more subtitle testers like me, but not young people because they have little experience" (P3). Most of them claimed that they would be very happy to be able to do the task for a longer time by themselves with new videos like these (P1–4), but all of them would welcome the incentive of choosing topics interesting for them (P1–P7), as well as controlling for time ("The movies should be shorter, then I could watch anything! Just give me ten 5 minute films and I can do that for an hour" P3). Also, some of them (P2) are willing to do this activity in intergenerational context, i.e. with their grandchildren, who live abroad, when they come home to visit them during the holidays. Moreover, the videos in the study were deemed as quite interesting, and most of the study participants said that they "learned a lot" (P1–4). However, the participants had some issues with the specialized jargon used by the speakers (e.g. "it is not explained what is this photon" or "Spiderman, this is not Polish" by P2 and "kryptonite, must be a mistake" by P5, P6) and with the subtitles disappearing "too fast" (P1–7). While they enjoy finding errors in subtitles, they would not be interested to create their own, as it is "too much typing and time" (P1, P2, P4), apart from one who said that they "would have to try" (P3). They also said they enjoyed the language aspect, and they would be willing to try other language-related tasks in the future (P1–3, P6, P7).

4.3 Notes on errors marked

None of the synchronization errors were found, but also none were incorrectly marked. Stylistic errors were marked as such if the problem was with language style, while largely ignored when it was connected to subtitle conventions and formatting. The most common errors found were thus stylistic and grammatical errors related to spelling and general language correctness. Capitalization and punctuation errors, belonging to the grammar category, were difficult to spot for the older adults, we expect because some of the errors required them to remember the punctuation from the previous lines. On top of this, many study participants interpreted the "meaning" category, as something to mark when they encountered an unfamiliar term, or a divergence between spoken phrase and its transcription for the native language, which made for a few false positives to mark more specialized words and subtitles with higher compression rates.

5 DISCUSSION

Based on the exploratory research and pilot evaluation of our setup of online crowdsourcing translation process support via Smart TV platform we identified five promising areas of consideration when designing similar systems. These include (1) autonomy and freedom of choice, (2) physical comfort and (3) cognitive comfort to aid comprehension as well as (4) building confidence and (5) providing the edutainment value. While these five considerations form a starting point as preliminary guidelines for further inquiries, it is important to remember the limitations of this exploratory study, as it was conducted with older adults in Poland, among the more active and ICT-aware well-educated individuals belonging to the economic lower and middle-class[11]. The following discussion is aimed to inform the design of systems which could better include such older adults in ICT-dependent crowdsourcing tasks. For each of the signaled general considerations there are preliminary proposals of best practices to mitigate or overcome the observed barriers and accommodate preferences which are discussed below.

Promising considerations for designing crowdsourcing systems for older adults:

5.1 Autonomy and freedom of choice

The question of the ability to choose the topic, duration and engagement level with the content at any time appeared to be of crucial importance for the engagement of our group of older adults.

5.1.1 Freedom to choose the topic. The older adults wanted to **freely choose the content to watch**, based on their preferences. This is of key importance, as in our use case their primary motivation comes from the alignment with their interests providing the edutainment value. Thus, one crucial insight is that it is necessary to introduce older adults' profiling, perhaps in the form of a few questions during the initial setup of the app to allow the participants to be served videos based on their interests, as well as topic categorizations within the app. Moreover, for longer studies it is important to introduce the function to search by the topic of the video, as there are clearly some topics favored by older adults and because in this case their interests act as a motivational element. This could be supported by a user profiling system coupled with an adaptive mechanism, that monitors and follows user preferences and their changes over time.

5.1.2 Freedom to control task duration. The older adults wished to be able to choose how long they want the videos to be, and how much time they would like to devote to this task overall, which is another potential adaptive service.

5.1.3 Freedom to contribute as little or as much as one wishes. The older adults were pleased about the learning potential of the videos, but they were also eager to finish watching the videos early if they were not as interested in the topic (P1, P5). This suggests that the task is just as engaging as the video, which corresponds to the edutainment concept. However, the more engrossed older adults were the fewer errors they seemed to detect, which suggests prompts or some other technique could be introduced to remind the viewers of the task. This, combined with their unwillingness to create subtitles themselves, as it is too much time and effort, shows that **this type of crowdsourcing should rely on quantity of contributors, rather than their accuracy** because participants are unlikely to go back to the videos to watch them again and verify their own accuracy when detecting errors.

[11]The best status class to describe this group would be Intelligentsia, which is the class of educated, but not necessarily well-off people in Poland.

5.2 Physical comfort and familiarity

It indeed seems that our group of older adults interacted with Smart TVs in a more familiar, and definitely more comfortable fashion than with computers due to their long-time familiarity with the setup and controls, as well as the physical comfort granted by their leisure home environment.

5.2.1 Physically comfortable setup (TV). Regarding the Smart TV interface, **people who watch TV at home already have a very comfortable setup for spending time in front of the TV Set**, unlike their computer setup at home, which, as we observed, is mostly designed for short sessions. This corresponds with numerous market research done e.g. by Nielsen. Despite the task appearing in the same setup as one of their favorite leisure time activities, that is watching TV, our group of older adults viewed using the DreamTV application as more demanding, which raises the questions, how much time would they be willing to devote to it, and if the solution is sustainable in the long run. These ought to be further researched.

5.2.2 Ease of navigation with clearly labeled controls (remote). The **Android TV interface proved to be very easy to navigate**, as the older adults participating in our study are quite used to navigating the menus and teletext by clicking arrow buttons on the remote. We observed that this was a largely familiar and comfortable experience for them, despite the remote being much different from theirs for P1–P5. One suggestion we received was to mark the remote control with buttons of different colors, as now "they look the same" (P3, P4) or separate the direction keys (P6, P7). The colors could then correspond to the colors of actions which can be taken on the screen to assure cohesion of design between the physical controller (the remote) and the application design on screen.

5.2.3 Ability to take breaks. The older adults valued the task also for the ability to take breaks, pause, make tea (P3, P5) and readjust without losing sight of the context and feeling pressured into hurrying. This suggests that there ought to be no timeouts or time limits imposed in the design of other crowdsourcing tasks.

5.2.4 Adjustable text size and colors. There ought to be the possibility to adjust the size of the letters and introduce high contrast selection to accommodate vision impairments. The text size should be adjustable based on the possible needs of older adults, as not all of them have their TVs at a comfortable distance to read the fonts, and have varying degrees of light in the living room. While some of them compensated by moving the chairs closer to the screens, for long time use it would be better if they could stay on the couch where they are comfortable. We expect this to be a lesser issue in societies where subtitle use rates are higher.

5.3 Cognitive comfort to aid comprehension

5.3.1 Adjustable reading speed. **The required reading speed ought to be in the low range (14 ch/s at most)** or possibly the speed of the video playback could be adjustable, or even adaptive. What also follows, is that in order to allow older adults to be more effective at finding mistakes in Machine Translated subtitles the software has to use language compression algorithms, to cap the required reading speed at 14 ch/s for the ease of comprehension. This also holds true for human organic errors, as some human-translated videos, especially by untrained volunteers, use close to literal translation with no compression[12], which tends to make subtitles longer.

[12]More information on problems and techniques of subtitle compression can be found on TED Translators' wiki in the following location:https://translations.ted.com on the "How to Compress Subtitles" wiki page.

5.3.2 Ability to pause, go back and view the context for reference. It should be possible to **view the video transcript** and to go back to any previous point, as older adults have slower reflexes and often clicked pause too late. They also welcomed the ability to pause and check the context of the dialog list[13] for reference, in case they felt they may have missed something.

5.4 Building confidence

Based on our participatory observation and exit interviews, the older adults in our group consider themselves competent enough in their native language to detect errors, and with a well structured step-by-step tutorial, also errors in subtitles. With a more extensive tutorial on types of errors common in subtitles, combined with a quick refresher of grammar rules, especially focused on punctuation, the older adults will be a good group for this type of crowdsourcing.

5.4.1 Making use of older adults' language skills. Without extensive prior training and with some limited empowerment provided by the researchers, the older adults can **use their natural language experience to find mistakes**, especially with grammar and style, which works well with MT errors, but excludes technical problems such as formatting and synchronization. While making use of older adults' expert native user proficiency in language enabled crowdsourcing is very empowering, the question of the evolution of language also deserves attention. Some older adults were not familiar with technical or pop culture references (P2, P5 and P6) and they assumed these may have been mistakes. To mitigate this barrier we propose that a reference dictionary could be available to the participants via the app, in which they could verify their linguistic hunches and at the same time learn more modern terms for the contexts they encounter.

5.4.2 Providing a theoretical and practical tutorial. While grammar and style seem to be the best category for this type of task, the timing issues were not detected by anybody, neither in transcription nor in translation tasks. It is possible the difficulty with finding timing (sync) issues was due to not including this category in the exercises before the task, as they were made on-paper and were limited to the choice between "grammar", "style" and "meaning". Moreover, the technical aspects of the subtitling process, despite appearing in the protocol on-paper exercise, such as subtitle stylistic conventions, were largely overlooked, and would require more training, or an introduction of the subtitle-specific tutorial to the application. In general, it seems that tasks where the subtitles are a translation (videos no. 3, 4, 5) are a better target for language correction than transcription tasks (videos no. 1, 2) where older adults tended to try to focus if the words written are exactly as said when finding mistakes. (especially P5). We conclude that **it would be beneficial to introduce a tutorial within the app** that would prime the older adults to focus on the most common error types.

5.5 Providing the edutainment value

The content we used for this activity, that is the TED and TEDx talks is quite engaging and broad in terms of domain. One challenge that exists then, is to **frame other crowdsourcing tasks in such way as to provide edutainment value** for the participants, which would allow them to consider this task to be closer to their leisure time than to work.

5.5.1 Using wide-domain educational content. As we observed the older adults from our group find improving the quality of subtitles for TED and TEDx talks to be an interesting and motivating task because of its educational value, and what has to be stressed, personally for them with the topics of their choice. Therefore it seemed that easily providing older adults with the choice of topic they are interested in, and at the same time ensuring that they are learning is a necessity

[13] More on this feature was explained in section 3.1.3, while the dialog list providing context is visible in Fig. 4.

when designing edutainment-like crowdsourcing. The DreamTV application facilitates this, as TED and TEDx talks are characterized by a very large domain, and multiple videos have a wider focus than just technology, entertainment and design.

Additional note

As more ICT-proficient generations enter retirement age in Western societies, such studies may soon prove to be more generalizable, both because of rising levels of ICT literacy, and the introduction of affordable solutions such as the Android TV STB, like Xiaomi MiBox, which can turn any TV set into a Smart TV. As such, it is a promising setup for an increasingly larger group of older adults, who are more ICT-savvy and have access to cheap technology able to turn their TV sets into Smart TVs.

6 CONCLUSIONS AND FURTHER WORK

6.1 Overview

The Smart TV setup and the interface are promising in the field of research with older adults, especially due to the familiarity, large screen size, general accessibility and comfortable at-home setup. Crowdsourcing subtitle errors is also perceived as an interesting task, which can be very engrossing and motivating depending on how it is framed. This study allowed us to evaluate the feasibility and direction of further inquiries in the area of TV-enabled crowdsourcing tasks, relying on native language proficiency of older adults. The key areas of interest when designing such systems based on our exploratory study seem to be (1) autonomy and freedom of choice, (2) physical comfort and familiarity, (3) cognitive comfort to aid comprehension as well as (4) building confidence and (5) introducing the edutainment value. Therefore, we consider our study an important voice, and a contribution into the discussion of vital research areas of human aspects of collaborative systems development and crowd work, raised in particular at CSCW in the recent years by numerous researchers, e.g. from Carnegie Mellon, Northwestern and Stanford Universities [15, 28, 61, 62]. In particular we are convinced that the Smart TV interface is promising for developing sustainable crowdsourcing solutions, and platforms that could truly involve older adults in crowd tasks by lowering ICT barriers and motivating them in various ways, from providing engaging content and edutainment, to ensuring social inclusion and enabling contribution to the society at large.

6.2 A follow-up long term study

However, further work is necessary to verify our presented preliminary considerations in a follow-up long-term study. This is feasible, as we have promising findings that some participants would like to continue, and some (P2 and P3) in fact have already subscribed, with one participant (P2) suggesting to do the task with grandchildren, thus implying that the intergenerational dimension could be an additional motivator. Below we present a brief overview of five main areas of future research.

6.2.1 Sustainability of the solution and verification of preliminary findings. Given our observations and feedback, in the long-term study we would like to verify the five above-mentioned design considerations as well as the sustainability of this solution in terms of: continued interest, likelihood of stopping playback to mark an error, detecting false error positives, as well as changing motivation depending on the context of the task (e.g. intergenerational). As observed in our preliminary study, older adults wish to engage in crowdsourcing if it resembles leisure activities they can do for fun. In general, we believe that many older adults will be able to sustain interest in the task, once it has a wider range of content available that may be better aligned with their interests.

6.2.2 Error detection rates. Another interesting research path is the examination of the inverse relationship between the study participants' interest in a specific video, and the number of errors they detect. In order to test this, we intend to allow older adults to add an option to rate each video they have watched. To further examine the differences between detection rates the long term study would feature videos with both errors naturally occurring after human translation, as well as originating from machine translation-aided tasks. As our results indicate in both of these scenarios, the quality of translations can benefit from the involvement of older adults as proofreaders marking style and comprehension problems as well as grammar issues. Such data could be later used to facilitate Machine Learning processes in the Machine Translation field, but also help improve the quality of human translations.

6.2.3 Comprehension. Yet another aspect of the further study could focus on ways of addressing the comprehension difficulties we encountered, caused by the need to focus on many aspects of the subtitles and making choices regarding the error categories, in particular, by removing error categories altogether, and just enabling an error to be reported. There exists one related ground of research: crowdsourcing the errors that are perceived as errors by the older adults, and comparing them to mistakes which are generally regarded by the translation communities and edutainment scene as a whole. Older adults often have signaled problems with the speed and comprehension of the subtitles themselves. Freely gathered data on perceived errors could better inform ways of making these forms more accessible to older adults.

6.2.4 Ease of setup. At the same time, although the Smart TV interface and our application proved to be enjoyable and easy to interact with, it is important to verify if older adults would be able to set up the system on their own, and if not, how to make it easier for them. This alone would pave way for further inquiries into the feasibility of using Smart TV services for older adults.

6.2.5 Motivation. As is often the case with crowdsourcing, this task also depends on the scale of contributions. Even if some older adults, according to their psychological profile and motivation, may choose to just watch the videos for fun, the others may use the opportunity to detect errors, and with the effect of scale working the true errors would be detected more reliably. We would also like to test the motivational component involving statistics of the errors found, and possibly gamifying them. Once the DreamTV application has been more extensively tested and developed to include the suggestions, and to resolve the difficulties occurring, we would like to allow as many users as possible free access to the application, and to collect longitudinal quantitative data on the errors reported. At the same time, we would like to test the personal (increasing the positive self-image) and social (engaging in a community) potential of this type of task we postulated when designing the DreamTV crowdsourcing solution. We could test it by encouraging older adults to engage with the TED Translators community, members of which could request their help in proofreading their own translations with the help of the application.

6.3 Closing remarks

Although some of our findings can be generalized to inform future design further work needs to be conducted with larger and more diverse groups alongside follow-up long-term studies. For example, in terms of wealth, education, language abilities and proficiency with technology, as well as with younger adults or speakers of other languages (e.g. migrants). This will help to explore which of our preliminary findings may be universal, and which ones rather depend on factors specific to different groups.

Fig. 7. As the application allows users to select their favorite videos and save their progress it has potential for sustaining engagement, which is also helpful for conducting long term studies.

7 ACKNOWLEDGMENTS

We would like to thank the TED Translators community for their insights and work on making interesting ideas accessible to more people through their subtitles, as well as the older adults who participated in this study. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 690962.

REFERENCES

[1] 2014. Measuring Digital Skills across the EU: EU wide indicators of Digital Competence. https://ec.europa.eu/digital-single-market/en/news/measuring-digital-skills-across-eu-eu-wide-indicators-digital-competence. (2014).

[2] 2015. *World Population Aging 2015.* Technical Report. United Nations Department of Economic and Social Affairs. 9–39 pages.

[3] 2017. Population structure and ageing. http://ec.europa.eu/eurostat/statistics-explained/index.php/Population_structure_and_ageing. (2017).

[4] 2017. Turning to Education for Fun. (Dec 2017). Retrieved August 20, 2018 from https://www.nytimes.com/2015/03/20/education/turning-to-education-for-fun.html

[5] M. Alaoui and M. Lewkowicz. 2015. Practical issues related to the implication of elderlies in the design process – The case of a Living Lab approach for designing and evaluating social TV services. *IRBM* 36, 5 (2015), 259–265. https://doi.org/10.1016/j.irbm.2015.06.002

[6] Mike Armstrong. 2013. *The Development of a Methodology to Evaluate the Perceived Quality of Live TV Subtitles.* Technical Report. British Broadcasting Company.

[7] Anne Aula. 2004. Learning to use computers at a later age. In *HCI and the Older Population.* University of Glasgow, Leeds, UK, 3–5.

[8] Purya Baghaei and Mona Tabatabaee. 2015. The C-Test: An integrative measure of crystallized intelligence. *Journal of Intelligence* 3, 2 (2015), 46–58.

[9] Bartłomiej Balcerzak, Wiesław Kopeć, Radosław Nielek, Sebastian Kruk, Kamil Warpechowski, Mateusz Wasik, and Marek Węgrzyn. 2017. Press F1 for Help: Participatory Design for Dealing with On-line and Real Life Security of Older Adults. In *Proceedings of the 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2017*, Vol. 1. IEEE, 240–243. https://doi.org/10.1109/STC-CSIT.2017.8098778

[10] Bartlomiej Balcerzak, Wieslaw Kopec, Radoslaw Nielek, Kamil Warpechowski, and Agnieszka Czajka. 2018. From Close the Door to Do not Click and Back. Security by Design for Older Adults. In *Advances in Intelligent Systems and Computing II.* Springer, 40–53. https://doi.org/10.1007/978-3-319-70581-1_3

[11] Robin Brewer, Meredith Ringel Morris, and Anne Marie Piper. 2016. Why would anybody do this?: Understanding Older Adults' Motivations and Challenges in Crowd Work. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM, 2246–2257.

[12] L Cámara de la Fuente. 2014. Multilingual Crowdsourcing Motivation on Global Social Media, Case Study: TED OTP. *Sendebar* 25 (2014), 197–218.

[13] Alex Carmichael. 1999. Style guide for the design of interactive television services for elderly viewers. (1999).

[14] John Carroll, Gregorio Convertino, Umer Farooq, and Mary Beth Rosson. 2012. The firekeepers: Aging considered as a resource. *Universal Access in the Information Society* 11, 1 (03 2012), 7–15. https://doi.org/10.1007/s10209-011-0229-9

[15] Gabriel Culbertson, Solace Shen, Erik Andersen, and Malte Jung. 2017. Have your cake and eat it too: Foreign language learning with a crowdsourced video captioning system. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 286–296.

[16] Sara J Czaja, Neil Charness, Arthur D Fisk, Christopher Hertzog, Sankaran N Nair, Wendy A Rogers, and Joseph Sharit. 2006. Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (CREATE). *Psychology and aging* 21, 2 (2006), 333.

[17] Sara J Czaja and Chin Chin Lee. 2009. Information technology and older adults. *Human-computer interaction: Designing for diverse users and domains* (2009), 18–30.

[18] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. ACM, 9–15.

[19] Anna Dickinson, John Arnott, and Suzanne Prior. 2007. Methods for human-computer interaction research with older people. *Behaviour & Information Technology* 26, 4 (2007), 343–352.

[20] Zineb Djoub. 2013. ICT education and motivating elderly people. In *Ariadna; cultura, educación y tecnología*, Vol. 1. 88–92. https://doi.org/10.6035/Ariadna.2013.1.15

[21] Thierry Etchegoyhen, Lindsay Bywood, Mark Fishel, Panayota Georgakopoulou, Jie Jiang, Gerard Van Loenhout, Arantza Del Pozo, Mirjam Sepesy Maucec, Anja Turner, and Martin Volk. 2014. Machine Translation for Subtitling: A Large-Scale Evaluation.. In *LREC*. 46–53.

[22] Anusca Ferrari. 2013. *DIGCOMP: A framework for developing and understanding digital competence in Europe*. Technical Report.

[23] A.D. Fisk, S.J. Czaja, W.A. Rogers, N. Charness, and J. Sharit. 2009. *Designing for Older Adults: Principles and Creative Human Factors Approaches, Second Edition*. CRC Press.

[24] Emily A Greenfield and Nadine F Marks. 2004. Formal volunteering as a protective factor for older adults' psychological well-being. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 59, 5 (2004), S258–S264.

[25] Yanni Hao. 2008. Productive activities and psychological well-being among older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 63, 2 (2008), S64–S72.

[26] Toshinari Itoko, Shoma Arita, Masatomo Kobayashi, and Hironobu Takagi. 2014. Involving senior workers in crowdsourced proofreading. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 106–117.

[27] Dean Jansen, Aleli Alcala, and Francisco Guzman. 2014. Amara: A sustainable, global solution for accessibility, powered by communities of volunteers. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 401–411.

[28] Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 1301–1318.

[29] Bran Knowles and Vicki L. Hanson. 2018. Older Adults' Deployment of 'Distrust'. *ACM Trans. Comput.-Hum. Interact.* 25, 4, Article 21 (Aug. 2018), 25 pages. https://doi.org/10.1145/3196490

[30] Bran Knowles and Vicki L. Hanson. 2018. The Wisdom of Older Technology (Non)Users. *Commun. ACM* 61, 3 (Feb. 2018), 72–77. https://doi.org/10.1145/3179995

[31] Wiesław Kopeć, Katarzyna Abramczuk, Bartłomiej Balcerzak, Marta Juźwin, Katarzyna Gniadzik, Grzegorz Kowalik, and Radosław Nielek. 2017. A Location-Based Game for Two Generations: Teaching Mobile Technology to the Elderly with the Support of Young Volunteers. In *eHealth 360*. Springer, 84–91. https://doi.org/10.1007/978-3-319-49655-9_12

[32] Wiesław Kopeć, Bartłomiej Balcerzak, Radosław Nielek, Grzegorz Kowalik, Adam Wierzbicki, and Fabio Casati. 2018. Older Adults and Hackathons: A Qualitative Study. *Empirical Software Engineering* 23, 4 (2018), 1895–1930. https://doi.org/10.1007/s10664-017-9565-6

[33] Wiesław Kopeć, Radosław Nielek, and Adam Wierzbicki. 2018. Guidelines Towards Better Participation of Older Adults in Software Development Processes using a new SPIRAL Method and Participatory Approach. In *Proceedings of the CHASE'18: International Workshop on Cooperative and Human Aspects of Software (ICSE '18)*. ACM, New York, NY, USA. https://doi.org/10.1145/3195836.3195840

[34] Wiesław Kopeć, Kinga Skorupska, Anna Jaskulska, Katarzyna Abramczuk, Radoslaw Nielek, and Adam Wierzbicki. 2017. LivingLab PJAIT: Towards Better Urban Participation of Seniors. In *Proceedings of the International Conference on Web Intelligence (WI '17)*. ACM, New York, NY, USA, 1085–1092. https://doi.org/10.1145/3106426.3109040

[35] Anna Kötteritzsch, Michael Koch, and Fritjof Lemân. 2014. Adaptive training for older adults based on dynamic diagnosis of mild cognitive impairments and dementia. In *International Workshop on Ambient Assisted Living*. Springer, 364–368.

[36] J. Lave and E. Wenger. 1991. *Situated Learning: Legitimate Peripheral Participation.* Cambridge University Press.

[37] Terry Y Lum and Elizabeth Lightfoot. 2005. The effects of volunteering on the physical and mental health of older people. *Research on aging* 27, 1 (2005), 31–55.

[38] John J McArdle, Fumiaki Hamagami, William Meredith, and Katherine P Bradway. 2000. Modeling the dynamic hypotheses of Gf–Gc theory using longitudinal life-span data. *Learning and Individual Differences* 12, 1 (2000), 53–79.

[39] Nancy Morrow-Howell. 2010. Volunteering in later life: Research frontiers. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 65, 4 (2010), 461–469.

[40] Nancy Morrow-Howell, Jim Hinterlong, Philip A Rozario, and Fengyan Tang. 2003. Effects of volunteering on the well-being of older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 58, 3 (2003), S137–S145.

[41] Radoslaw Nielek, Miroslaw Ciastek, and Wiesław Kopeć. 2017. Emotions Make Cities Live: Towards Mapping Emotions of Older Adults on Urban Space. In *Proceedings of the International Conference on Web Intelligence (WI '17)*. ACM, New York, NY, USA, 1076–1079. https://doi.org/10.1145/3106426.3109041

[42] Radoslaw Nielek, Marta Lutostańska, Wiesław Kopeć, and Adam Wierzbicki. 2017. Turned 70?: It is Time to Start Editing Wikipedia. In *Proceedings of the International Conference on Web Intelligence (WI '17)*. ACM, New York, NY, USA, 899–906. https://doi.org/10.1145/3106426.3106539

[43] Francisco Nunes, Maureen Kerwin, and Paula Alexandra Silva. 2012. Design Recommendations for TV User Interfaces for Older Adults: Findings from the eCAALYX Project. In *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '12)*. ACM, New York, NY, USA, 41–48. https://doi.org/10.1145/2384916.2384924

[44] Jennifer M Ortman, Victoria A Velkoff, Howard Hogan, et al. 2014. An aging nation: the older population in the United States. *Washington, DC: US Census Bureau* (2014), 25–1140.

[45] Dorota Orzeszek, Wieslaw Kopec, Marcin Wichrowski, Radoslaw Nielek, Bartlomiej Balcerzak, Grzegorz Kowalik, and Malwina Puchalska-Kaminska. 2017. Beyond Participatory Design: Towards a Model for Teaching Seniors Application Design. In *Proceedings of the ER Forum 2017 co-located with the 36th International Conference on Conceptual Modelling (ER 2017), Valencia, Spain, - November 6-9, 2017. (CEUR Workshop Proceedings)*, Vol. 1979. 29–43.

[46] R. Pak and A. McLaughlin. 2010. *Designing Displays for Older Adults.* CRC Press.

[47] Zhengxiang Pan, Chunyan Miao, Han Yu, Cyril Leung, and Jing Jih Chin. 2015. The effects of familiarity design on the adoption of wellness games by the elderly. In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, Vol. 2. IEEE, 387–390.

[48] Kowit Rapeepisarn, Kok Wai Wong, Chun Che Fung, and Arnold Depickere. 2006. Similarities and differences between learn through play and edutainment. In *Proceedings of the 3rd Australasian conference on Interactive entertainment.* Murdoch University, 28–32.

[49] Pablo Romero-Fresco and Franz Pöchhacker. 2017. Quality assessment in interlingual live subtitling: The NTR Model. 16 (01 2017), 149–167.

[50] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[51] Jatinder Sandhu, Leela Damodaran, and Leonie Ramondt. 2013. ICT Skills Acquisition by Older People: Motivations for learning and barriers to progression. *International Journal of Education and Ageing* 3, 1 (2013), 25–42.

[52] Paula Alexandra Silva, Francisco Nunes, Ana Vasconcelos, Maureen Kerwin, Ricardo Moutinho, and Pedro Teixeira. 2013. Using the Smartphone Accelerometer to Monitor Fall Risk while Playing a Game: The Design and Usability Evaluation of Dance! Don't Fall. 8027 (07 2013), 754–763.

[53] Hsin-yi Sandy Tsai, Ruth Shillair, Shelia R Cotten, Vicki Winstead, and Elizabeth Yost. 2015. Getting grandma online: Are tablets the answer for increasing digital inclusion for Older Adults in the US? *Educational gerontology* 41, 10 (2015), 695–709.

[54] Robert J Vallerand and Brian P O'Connor. 1989. Motivation in the elderly: A theoretical framework and some promising findings. *Canadian Psychology/Psychologie canadienne* 30, 3 (1989), 538.

[55] John Vines, Mark Blythe, Paul Dunphy, and Andrew Monk. 2011. Eighty something: banking for the older old. In *Proceedings of the 25th BCS Conference on Human-Computer Interaction.* British Computer Society, 64–73.

[56] John Vines, Gary Pritchard, Peter Wright, Patrick Olivier, and Katie Brittain. 2015. An Age-Old Problem: Examining the Discourses of Ageing in HCI and Strategies for Future Research. *ACM Trans. Comput.-Hum. Interact.* 22, 1, Article 2 (Feb. 2015), 27 pages. https://doi.org/10.1145/2696867

[57] Bettina von Helversen, Katarzyna Abramczuk, Wiesław Kopeć, and Radoslaw Nielek. 2018. Influence of Consumer Reviews on Online Purchasing Decisions in Older and Younger Adults. *Decision Support Systems* 113 (2018), 1–10. https://doi.org/10.1016/j.dss.2018.05.006

[58] Jing-Jen Wang and Alan S Kaufman. 1993. Changes in fluid and crystallized intelligence across the 20-to 90-year age range on the K-BIT. *Journal of Psychoeducational Assessment* 11, 1 (1993), 29–37.

[59] Natalie E Wolfson, Thomas M Cavanagh, and Kurt Kraiger. 2014. Older adults and technology-based instruction: Optimizing learning outcomes and transfer. *Academy of Management Learning & Education* 13, 1 (2014), 26–44.

[60] Heng-Li Yang and Cheng-Yu Lai. 2010. Motivations of Wikipedia content contributors. *Computers in Human Behavior* 26, 6 (2010), 1377–1383. https://doi.org/10.1016/j.chb.2010.04.011

[61] Lixiu Yu, Paul André, Aniket Kittur, and Robert Kraut. 2014. A comparison of social, learning, and financial strategies on crowd engagement and output quality. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 967–978.

[62] Haiyi Zhu, Steven P Dow, Robert E Kraut, and Aniket Kittur. 2014. Reviewing versus doing: Learning and performance in crowd assessment. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. ACM, 1445–1455.

[63] Gabe Zichermann and Christopher Cunningham. 2011. *Gamification by design: Implementing game mechanics in web and mobile apps* (1st ed.). " O'Reilly Media, Inc.".

Received April 2018; revised July 2018; accepted September 2018

5.2 Conversational Crowdsourcing for Older Adults: a Wikipedia Chatbot Concept (ECSCW 2020, 70 pts)

Title	Conversational Crowdsourcing for Older Adults: a Wikipedia Chatbot Concept
Authors	Kinga Skorupska, Kamil Warpechowski, Radosław Nielek, Wiesław Kopeć
Conference	European Conference on Computer Supported Cooperative Work [ECSCW 2020] (CORE B, 70 points on the list of the Polish Ministry of Science and Higher Education, as accessed on 5.04.2022)
Published in	Proceedings of the 18th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies.
Author Credit	I was the lead author of this article and my key contributions constituted the general chatbot concept, chatbot design analysis and text interaction flow, literature review, conclusions and future work potential, manuscript drafting and revision. I have also presented the paper at the conference (remotely).
Description	Based on our research on the Wikipedia interface and crowdsourcing with older adults, we propose a conversational interface to streamline Wikipedia editing, engage new contributors and increase their well-being. The use of a conversational interface may mitigate the problem of a steep learning curve for new contributors to encourage more people to contribute to Wikipedia, and thus, little by little, make it more accurate, consistent and democratic. This solution can also negotiate some of the barriers apparent in older adults' interaction with the Wikipedia interface. To achieve these goals, we conceptualized a friendly chatbot called "Gizmo" which inverts the human-chatbot interaction paradigm by making the user be the one to aid the chatbot. In doing so, we explored some of the requirements and challenges associated with the design of a conversational interface to enable Wiki contributions. These include the choice of the appropriate task to crowdsource, in our case the infobox translation verification, the initiation of the conversation as well as the motivational component with key disaffection indicators. At the same time, we discuss some opportunities within the domain of CSCW related to the design and applications of novel conversational crowdsourcing interfaces.
Contribution and Impact	This contribution proposes one way to address the issue of uneven participation of different geographical and age groups in Wikipedia, and crowdsourcing in general. The proposed conversational interface for microtasking is a voice in the debate on how to use chatbots for social good, especially initiated by Asbjørn Følstad, a senior researcher at SINTEF, Norway. This contribution is, so far, cited in the following studies by researchers from other institutions: – Veglis, Andreas, and Efthimis Kotenidis. "Employing chatbots for data collection in participatory journalism and crisis situations." Journal of Applied Journalism & Media Studies (2021). – Gluza, Wioletta, Izabella Turaj, and Florian Meier. "Wikipedia Edit-a-thons and Editor Experience: Lessons from a Participatory Observation." 17th International Symposium on Open Collaboration. 2021. – Geszten, Dalma. "Teammunkát támogató szoftverek team-szintű használhatósági vizsgálati módszerének fejlesztése." (2021).

Kinga Skorupska, Kamil Warpechowski, Radosław Nielek, Wiesław Kopeć (2020): Conversational Crowdsourcing for Older Adults: a Wikipedia Chatbot Concept. In: Proceedings of the 18th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2020_ep05

Conversational Crowdsourcing for Older Adults: a Wikipedia Chatbot Concept

Kinga Skorupska, Kamil Warpechowski, Radosław Nielek, Wiesław Kopeć
Polish-Japanese Academy of Information Technology, Warsaw, Poland
kinga.skorupska@pja.edu.pl

Abstract. Based on our research on the Wikipedia interface and crowdsourcing with older adults, we propose a conversational interface to streamline Wikipedia editing, engage new contributors and increase their well-being. The use of a conversational interface may mitigate the problem of a steep learning curve for new contributors to encourage more people to contribute to Wikipedia, and thus, little by little, make it more accurate, consistent and democratic. This solution can also negotiate some of the barriers apparent in older adults' interaction with the Wikipedia interface. To achieve these goals, we conceptualized a friendly chatbot called "Gizmo" which inverts the human-chatbot interaction paradigm by making the user be the one to aid the chatbot. In doing so, we explored some of the requirements and challenges associated with the design of a conversational interface to enable Wiki contributions. These include the choice of the appropriate task to crowdsource, in our case the infobox translation verification, the initiation of the conversation as well as the motivational component with key disaffection indicators. At the same time, we discuss some opportunities within the domain of CSCW related to the design and applications of novel conversational crowdsourcing interfaces.

Introduction

Demographic trends show a constant increase in the share of older adults in the western world, and the long term 2015 EUROPOP projection predicts this pattern will continue. Thus, it is increasingly important to encourage more older adults to remain active, both mentally and physically to foster their well-being, as in Morrow-Howell et al. (2003), and at the same time to contribute to the society and feel needed, even while staying at home, for example through engaging with crowdsourcing platforms. Crowdsourcing projects may benefit from older adults' experience, including their native language proficiency, as shown in Skorupska et al. (2018). Another such opportunity is engaging with Wikipedia, which is the largest crowdsourcing project, averaging 591 new English articles per day, with, according to data from Wikipedia.org (2020), over 39 million registered users on the English site. While all of these users have the technical possibility to contribute only a minority of them do so regularly, and those, as proven by Lee and Seo (2016), set the tone and focus of the content.

This is due to multiple reasons, including the complexity of the process, accessibility issues and motivation. As shown in our previous research by Nielek et al. (2017) on editing Wikipedia with older adults, these barriers are especially prominent for them, as they may have lower ICT skills and do not enjoy regular crowdsourcing tasks, as visible in a study by Brewer et al. (2016). Moreover, some older adults lack confidence in their ability to contribute valuable knowledge and insights in technology-mediated contexts as shown in Kopeć et al. (2018), which may also be amplified by a rich set of Wikipedia community guidelines[1] and senior editors who sometimes may stop new editors from becoming regulars, as noticed by Halfaker et al. (2013). To overcome these challenges we propose a chatbot for crowdsourcing Wikipedia edits, that would negotiate the editing complexity for the user, break the editing tasks into smaller chunks which are more likely to be accepted, and encourage a stream of new contributors thanks to inhabiting IM platforms.

Khan and Das (2018) defined chatbots as computer programs that "process natural-language input from a user and generate smart and relative responses that are then sent back to the users", thus, there is a focus on assisting the users. The proposed chatbot inverts the human-chatbot interaction paradigm, by being the one to gather information from the users. Our chatbot seeks assistance from the users, and acts as an intermediary for crowdsourcing Wikipedia contributions. Its dialogue structure hides the steps of content creation behind a simple conversational interface for new editors - one that can be accessed with many devices, such as a smartphone, a Smart TV or even a smart speaker via voice interaction, which is a promising venue as shown in research by Kowalski et al.

[1] New editors are encouraged to read pages such as: https://en.wikipedia.org/wiki/Wikipedia:Guidance_for_younger_editors (accessed on 28.01.2020, revision ID=937682704) or https://en.wikipedia.org/wiki/Wikipedia:A_primer_for_newcomers (accessed on 28.01.2020, revision ID=933431927) which may discourage potential contributors, especially if they are technology-shy.

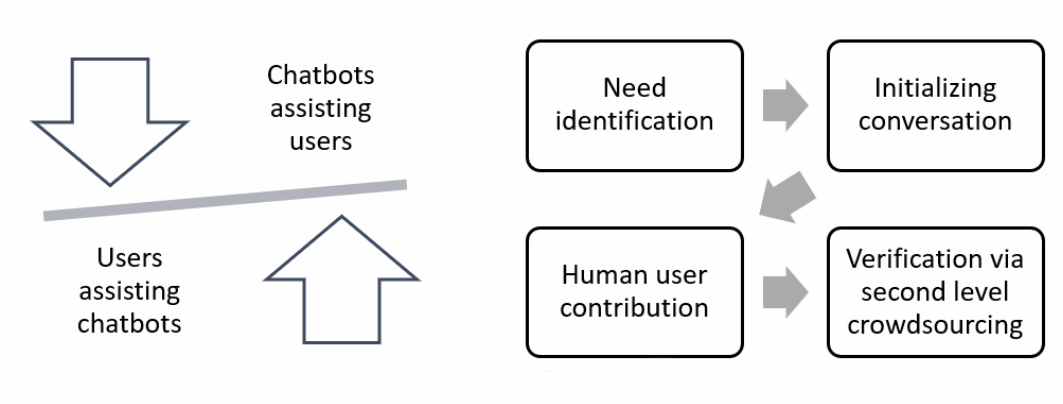


Figure 1. Inversion of the human-chatbot interaction paradigm and proposed chatbot-streamlined contribution process.

(2019). This implementation follows the logic on how chatbots are used to ease interaction with computers by shifting the weight of cognitive work of performing tasks from users to conversational agents. Morrissey and Kirakowski (2013) identify two main strategies that aid in this application, i.e "naturalness of interaction and sharing knowledge space". Hereby, it is easy to conceptualize that chatbots may also negotiate many circumstances in which it is necessary to learn artificial conventions, including crowdsourcing. The chatbot also embraces the idea of Chatbots for Social Good explored by Følstad et al. (2018). They point to the need to make various services more accessible via chatbots for the benefit of the users, such as increasing their opportunities for education, autonomy as well as their feelings of social relatedness. But in our case the user also gives back to the society by assisting in incrementally raising the quality of Wikipedia entries.

So, the aim of our chatbot is twofold: it explores the potential of employing a conversational interface for crowdsourcing with the use of micro-tasks, and in particular, it makes Wikipedia editing more accessible to everyone, but especially to older adults who, in general, do well with text and voice-driven interaction, and who may benefit from engaging in online volunteering, especially based on their inherent knowledge and skills.

Related work

Wikipedia and crowdsourcing

Apart from Wikipedia's status as a massive crowdsourcing effort in itself, there are also some indirect ways its content can be edited via separate crowdsourcing projects. For example, workers can help to validate the credibility of editors. Suzuki and Nakamura (2016) developed a system of rating contributors based on their historical edits in which vandalism can be predicted based on a database of labeled pairs of sentences. Another type of micro-tasks on Wikipedia is related to

verification of information. Redi et al. (2019) described the problem of missing citations and references in articles and had workers check or add requests for references assigned to sentences or articles. With the help of DBpedia, a queryable linked dataset based on Wikipedia, preparing such micro-tasks can be even easier.

There are few studies with older adults and Wikipedia and the existing ones, for example by Nielek et al. (2017), show problems with the accessibility of the interface including the colors, coherence and feedback. Such problems were also identified by Vora et al. (2010) who conducted a qualitative study with first-time editors. Additionally, Wikipedia is suffering from a decline in new editors, not only because senior editors discourage them, as noticed by Halfaker et al. (2013), but also a myriad of other interaction-driven problems, including two large categories related to confusing policies (overwhelming amount to read before one can start with no help) and technical difficulties (poor visual editor, having to use markup, lack of automation), as noticed by the Wiki community themselves. [2]

In this light, a chatbot seems like a perfect intermediary for editing Wikipedia, and solving some of its problems, which can be translated into micro-tasks and done without prior knowledge, especially for older adults, who often appreciate the ease and the social aspect of conversational interaction, as explored in Atay et al. (2016) and voice interaction, as in Kowalski et al. (2019)

Older adults

Volunteering is linked to improved physical and mental health, as shown by Lum and Lightfoot (2005), as well as well-being, which was explored by Morrow-Howell et al. (2003) and Greenfield and Marks (2004) with older adults. But for older adults to benefit from crowdsourcing tasks, some barriers to their participation need to be overcome, and these include lack of or lower ICT skills, lack of motivation due to unclear personal benefit or unsocial and repetitive nature of the tasks, as on Mechanical Turk study by Brewer et al. (2016). To encourage older adults to contribute Japanese researchers created a platform for proofreading scanned books as described by Itoko et al. (2014) and Kobayashi et al. (2013). They also researched different ways of motivating older adults to engage with it in a following study by Kobayashi et al. (2015). But the motivation of older adults is usually studied in context of physical exercises as in Schutzer and Graves (2004), or as in a study by Navarro et al. (2007), learning as in Kim and Merriam (2004) or work by Kanfer et al. (2013). Most of these works refer to the classic types of motivations, such as intrinsic and extrinsic, so ideas need to be adapted for chatbots. As for crowdsourcing tasks, researchers in Singapore, namely Yu et al. (2016), created a mobile application where older adults could tag historical photos. In general, older adults are willing to complete crowdsourcing tasks when the

[2] A full community analysis of this still relevant problem can be found here: https://en.wikipedia.org/wiki/Wikipedia:Why_is_Wikipedia_losing_contributors_-_Thinking_about_remedies (accessed on: 01.02.2020, revision ID=888559458). Similar issues are also touched upon in an MIT technology review article here: https://www.technologyreview.com/2013/10/22/175674/the-decline-of-wikipedia/.

system is more engaging and interaction more familiar to them, as shown in Skorupska et al. (2018). To address the lack of ICT-skills Hiyama et al. (2013) tested a question-based passive interaction model for gathering knowledge from seniors, which also inspired the design of our chatbot.

Chatbot design

Brandtzaeg and Følstad (2017) concluded that unlike other devices and software, chatbots ought to be created as "tools, toys and friends" at the same time. The friendship arc is explored in two scenarios by Lee et al. (2019) where people cared for, or were supported by, Vincent to increase their self-compassion in a study that may indicate human-chatbot relationships have the potential to affect users' well-being. Fichter and Wisniewski (2017) point out that two key aspects of chatbots are using friendly language, and the ability to handle failures in "an endearing manner" and increasing their acceptance by lowering expectations as in Broadbent et al. (2009). To aid in this Reshmi and Balakrishnan (2016) proposed an inquisitive chatbot "aware" of missing information and able to retrieve it by asking follow-up questions. Further repair strategies are explored by Ashktorab et al. (2019), who point out that misunderstandings ought to be acknowledged, repaired with user understanding and control, as well as adaptive to individual preferences. In their research Richards et al. (2016) explored ways to feed chatbots knowledge and noticed that the user satisfaction level is "proportional to the amount of fact knowledge" of the chatbot, which is likely connected to their efficiency within the domain in which they operate. To ease this process, Kobayashi et al. (2015) proposed an efficient way to extract question and response pairs from online forums to supplement chatbot knowledge. On top of this Morrissey and Kirakowski (2013) postulated four dimensions for evaluating the quality of conversations with chatbots: conscientiousness, originality, manners and thoroughness. Furthermore, Clark et al. (2019) identify key elements of a successful conversation to form design recommendations for chatbots which (or who) can move beyond task-oriented agents, even into the realm of being seen as community members, as postulated by Seering et al. (2019).

Still, the main challenge with human-computer interaction via text or voice conversational interface is natural communication. Moore et al. (2017) deems it necessary to overcome this challenge to increase user well-being and adoption of chatbots in everyday life. One thing that can help with that is the presence of response delays, which as Gnewuch et al. (2018) show ought to be dynamic (often created unnecessarily) to positively affect users' perception. Users then feel that other humans generate the responses. Also the visual aspects can change the perception of the conversation - a study by Candello et al. (2017) explored the use of a more robotic typeface which made the users more likely to evaluate a conversation with an advisor as bot-like, while a handwriting font did not show an opposite effect. The use of animated avatars potentially may also help humanize interactions by displaying emotions. Nevertheless, Ciechanowski et al. (2018)

show that in contrast to hardware robots, text, or voice bots without face are perceived as less dangerous and uncanny, and sometimes voice-only interaction or a humorous avatar is preferred to a human seeming one, as in Wang et al. (2019).

All of these considerations influenced the design of our "Gizmo" chatbot and can be summarized in the following set of recommendations:

1. Focus on their function as an interface to increase the users' productivity while satisfying interpersonal interaction needs.
2. Provide vast knowledge of the "inhabited" domain.
3. Ensure context awareness and retain key information about the user.
4. Allow more pre-defined dialogue choices and open-ended replies to increase naturalness.
5. Create a consistent and polite personality to grant positive social interaction.
6. Ensure careful and levelled emotional expression to match the shared narrative, as in Lee et al. (2019).
7. Avoid "the uncanny valley" by not using a human-like avatar.
8. Aim for amusing, endearing and entertaining demeanour.
9. Humanize the interaction with the use of dynamic response delays.

Conceptual design

While the best-case scenario would involve the possibility to generate all content for Wikipedia also via the chatbot interface, there are some practical considerations to take into account. First of all, Wikipedia is not a new project, but it has been around since 2001, as such, there is little need to elaborate on articles drawing from common, or easily accessible, knowledge in the most popular languages. Areas where there is largest need for contributions are translating the database of knowledge from English into local languages, which is still a massive task as examined by Wulczyn et al. (2016), and local interest information into other languages and verifying scientific and historical accuracy of existing articles, which may be lacking due to progress in the field, as shown by research reviewing articles in the Astronomy category by Thelwall (2016), bias or political agendas.

On the side of the users, there is preference for using their skills and knowledge that are immediately accessible to them without additional research and searching for references, which could break the immersion of interaction with the chatbot and make the task more work-like.

What these have in common are two large areas of concern: **translation and evaluation**, which address both the needs of the Wikipedia project, as well as make it possible for anyone to contribute without much research or specialized knowledge. These two aspects are what we will focus on.

Possible crowdsourcing tasks

In principle, the information that can be crowdsourced via a chatbot ought to come in chunks, so that it can be broken into micro-tasks, and be easily identifiable as the answer sought after. Examples of such tasks are listed below:

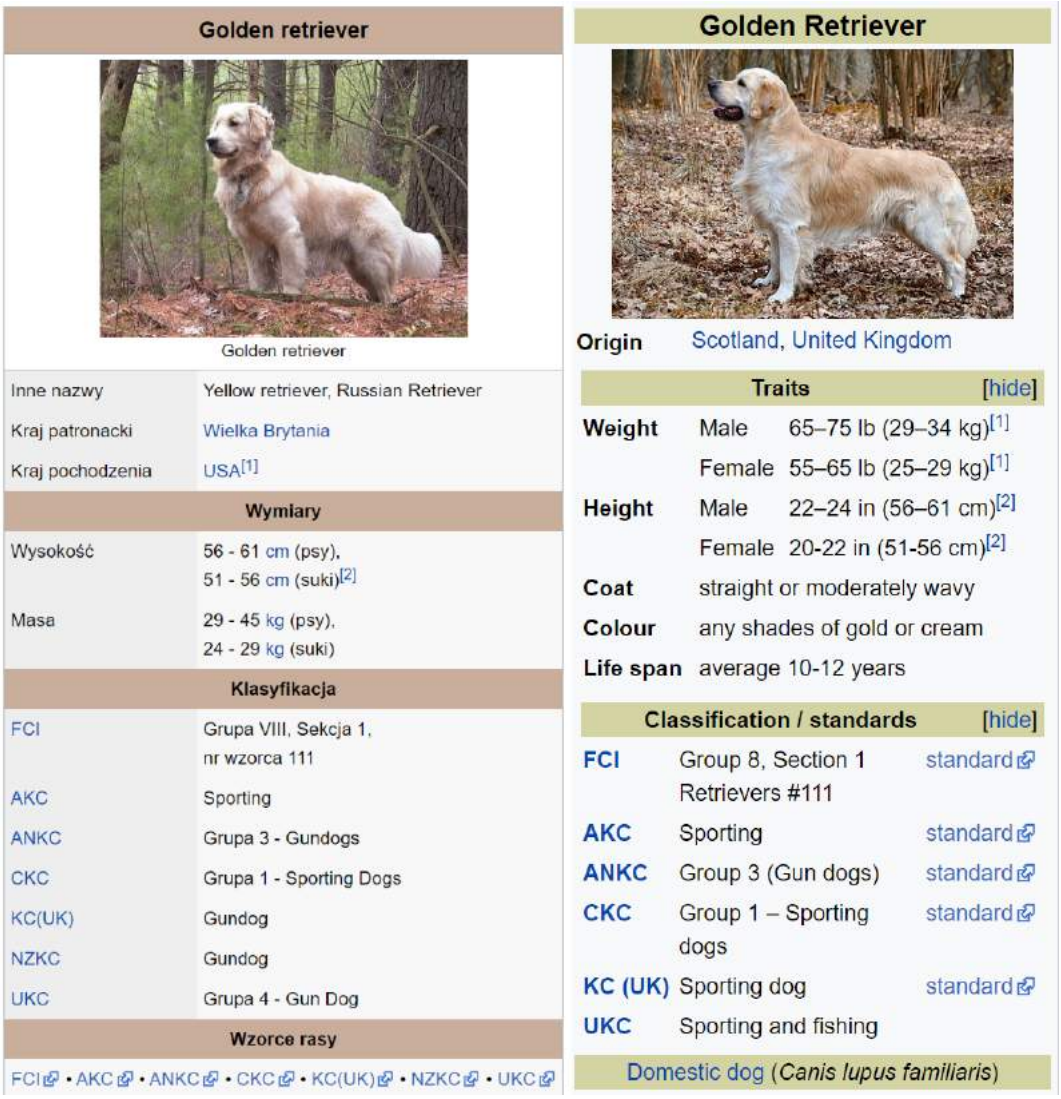
- **Historical information:** Knowledge about local history of the area, especially relevant in small towns and villages, including the history of institutions such as schools, churches or museums.
- **Local information:** Information about local attractions, buildings, museums projects, or people who are important locally, such as activists and officials.
- **Contested information:** Conducting surveys on the credibility of contested information, where the users can aid the chatbot to locate credible sources.
- **Sentiment verification:** Asking the users if information follows the guideline for the neutral point of view, or if an edit may have been an act of vandalism.
- **Meta-data:** Deciding on the classification of the articles, or aiding in the disambiguation as well as providing candidates for internal links within Wikipedia articles.
- **Tagging articles for clean-up:** Voting whether certain articles, perhaps from their pre-selected categories, need to be restructured or if they are understandable.
- **Translation verification:** Verifying translated chunks of information after machine translation, for example originating from infoboxes.

Pilot crowdsourcing task

Infoboxes are fixed-format tables in Wikipedia. They are most commonly used to display article summaries and navigation between pages. Each infobox consists of pairs of key/value. Infoboxes are very helpful for the automatic creation of knowledge graphs in the DBpedia project. Many categories of articles require the use of an infobox template as they ensure brevity and clarity of information. Wikipedia[3] has lists of categories and articles without infoboxes in multiple languages. These lists are a reference for users who want to expand Wikipedia, as many of these articles have equivalent pages in other languages which contain infoboxes. As this issue is common in multiple languages Zhang et al. (2017) introduced machine translation of values found in such tables and mechanisms for assigning keys in infoboxes. Doing this automatically would increase the internal consistency of information between the many languages of Wikipedia. Now, even in articles on neutral subjects, such as dog breeds[4], there are minor inconsistencies

[3] This Wikipedia article: https://pl.wikipedia.org/wiki/Kategoria:Artykuły_bez_infoboksu (accessed on 15.01.2020, revision ID=49749209) lists articles without an infobox for Polish.

[4] The exact search term used was "Golder Retriever, and the resulting url for Polish was: https://pl.wikipedia.org/wiki/Golden_retriever, accessed on: 1.02.2020, revision ID=58293665, and for English: https://en.wikipedia.org/wiki/Golden_Retriever, accessed on: 1.02.2020, revision ID=938596709).



Golden retriever

Golden retriever

Inne nazwy	Yellow retriever, Russian Retriever
Kraj patronacki	Wielka Brytania
Kraj pochodzenia	USA[1]
Wymiary	
Wysokość	56 - 61 cm (psy), 51 - 56 cm (suki)[2]
Masa	29 - 45 kg (psy), 24 - 29 kg (suki)
Klasyfikacja	
FCI	Grupa VIII, Sekcja 1, nr wzorca 111
AKC	Sporting
ANKC	Grupa 3 - Gundogs
CKC	Grupa 1 - Sporting Dogs
KC(UK)	Gundog
NZKC	Gundog
UKC	Grupa 4 - Gun Dog
Wzorce rasy	

FCI • AKC • ANKC • CKC • KC(UK) • NZKC • UKC

Golden Retriever

Origin Scotland, United Kingdom

Traits		[hide]
Weight	Male	65–75 lb (29–34 kg)[1]
	Female	55–65 lb (25–29 kg)[1]
Height	Male	22–24 in (56–61 cm)[2]
	Female	20-22 in (51-56 cm)[2]
Coat	straight or moderately wavy	
Colour	any shades of gold or cream	
Life span	average 10-12 years	

Classification / standards		[hide]
FCI	Group 8, Section 1 Retrievers #111	standard
AKC	Sporting	standard
ANKC	Group 3 (Gun dogs)	standard
CKC	Group 1 – Sporting dogs	standard
KC (UK)	Sporting dog	standard
UKC	Sporting and fishing	

Domestic dog (*Canis lupus familiaris*)

Figure 2. Side by side comparison of infoboxes about the Golden Retriever from the Polish and English Wikipedia. Information such as coat, colour and life span are missing in the Polish version.

between the content, as can be seen in Figure 2. With machine translation and consistent format of infoboxes in multiple languages the verification of information and updates could also be easier across many languages. **So, in our pilot implementation we propose the design of a micro-task to verify machine translations and information management in infobox migration from the English to the Polish version of Wikipedia.**

Stages of task execution - first interaction with the chatbot:

1. Algorithm downloads available categories and presents them to the user (from such topic classifications as for example: entertainment, food and drink, sports or people)

2. User selects one interesting category and a subcategory according to their interests (for example: "film" from "entertainment")

3. System pulls articles for editing (finds a similar page in English with an infobox) and categorizes them by difficulty level (based on number of keys and value types).

4. The system creates a shortlist of articles, starting from the easiest difficulty (to provide an early success and increase motivation).

5. The user chooses an article from the shortlist. (for example: the page for the "Forbidden Planet" movie)

6. The user can edit each of the rows. They can remove optional parameters, edit translations, skip the task and move to the next one or finish.

Initiating interaction

Establishing the best method of initiating interaction is a significant problem which is not only related to the question of when the users are most open to contributing to crowdsourcing, but also when they are willing to engage in either text or voice (or both!) conversations with a chatbot on regular basis.

So, there is the question of choosing the appropriate trigger and pattern of interaction. In case of integration with Wikipedia it could be entering the Wiki page which has some needs associated with it, for example it lacks an infobox, or the Wikipedia as a whole. In other cases it could be exiting an application on the device, location-based (coming back home in the evening) or using Messenger. One of the key strengths of chatbots populating the popular IMs is the fact that no additional software needs to be installed for the chatbot-user interaction to be possible. Thus, chatbots may avoid a massive barrier to quick adoption of new solutions.

The initiation of the conversation also has to come at a right time for the users to contribute, based on their habits and preferences. It also ought to be initiated between their regular activities, so that they can commit to a new activity. Gizmo will be able to predict a good time to propose the interaction with the user. It could also happen based on the preferences of the users who could invite the chatbot to ping them at the time of their choice.

Currently, however, none of the voice assistants are allowed to initialize conversation without user input. The user only can initiate interaction via text or voice with a specific phrase (for example: "Ok, Google. Talk with Gizmo"). We resolve the technical problem of initializing the conversation by implementing push notifications with an encouragement to run the voice assistant. Each user will have a unique ID number. The central server would send notification to the user device at the predicted best time for interaction based on historical user data.

Additionally, the assistant can offer to finish the interaction by suggesting a break if it detects a decrease in the response pace or quality. Then it can offer a suggestion to continue at a chosen time.

Recovery techniques

The more complex the task, the more likely it is that the conversation may fall flat. For this reason, since apart from crowdsourcing, the chatbot aims to increase well-being of the users it must offer appropriate recovery techniques. In this case the recovery techniques ought to aid users in finding a task they can complete to help them feel good and to satisfy their need for cognitive closure. So, when the communication breaks depending on the type of failure there are a couple of techniques which can be used to get back on track:

- an apology and polite request to rephrase
- a switch to a different category
- a switch to an easier task
- a switch to a verification task
- a switch to a different rich response type
- planning a break and next interaction date

Chatbot evaluation

Designing successful HCI interactions is both about the numbers and the reasons behind them. For example, DialogFlow analytics consist of usage data, NLU Data: showing the frequency of use of "intents", the exit percentages of users, and response time. These can aid in finding problem areas, however, key measures for us are related to the user engagement and satisfaction with the conversation, and only then the number of successful contributions to Wikipedia and its relation to the previous two measures.

To evaluate these we intend to look at chat logs and hold individual IDIs with our pilot users once Google Home Hub is available in Polish. It is of key importance for us to deliver a pleasant experience, which could enrich the lives of older adults and aid them in making meaningful and visible contributions online, which they could share with others. We also expect that as the chatbot becomes more successful at encouraging contributions, the users may come to expect of it to also provide a pleasant social interaction outside of its main domain. This can be achieved by analysing the logs and, for example, evaluating the social dimensions of the attempted conversations, as in Aiello (2020). Finally, participatory design workshops with older adults could provide insights into the interaction patterns desired by them as well as key areas of needs that the chatbot could attempt to answer.

Design components

Our pilot implementation focuses on verifying machine translation of infobox contents. The choice of this implementation is partly inspired by our research, described in Skorupska et al. (2018), with older adults, who detected errors in subtitles, of which some were machine translated. It is evident that to evaluate the naturalness and logic of a translation it is often enough to know the target language, and older adults are confident in their native language abilities. On the other hand infoboxes are good candidates for machine translation verification tasks, because of their straightforward build of short key and value pairs. This implementation consists of the following modules:

Chatbot module

This module handles the full conversation flow, user contributions, including successful task switching and recovery and later feedback handling, as the motivational component. The module ought to allow for autocomplete to aid with typing, but also make use of different communication formats and allow multimodal interaction.

Design framework and communication formats

Leading tech companies have worked on launching chatbots since the dawn of smartphones. Siri was launched by Apple in 2010, and Google followed two years later with Google Now. Cortana and Alexa both appeared in 2015, followed by Google Home. Meanwhile chatbot frameworks were being developed alongside, with Wit.ai leading the way, followed by Api.ai bought by Google and renamed to DialogFlow, and Luis.ai. These frameworks provide HCI researchers with an amazing opportunity to use cutting-edge AI and ML algorithms in their studies. Without the need to develop these solutions from scratch on their own, they can focus on creating a specific chatbot for their research purposes. For this research Google's DialogFlow was chosen as a free tool that can be trained, without the need to create own NLP algorithms, as observed also by Mitrevski (2018) and which can be used with IM platforms whose popularity surpassed social networks as noticed by Pereira (2016).

The communication formats available on DialogFlow partly motivated our choice of the design framework, as they match the needs of the project and allow for multimodal communication with a touchscreen, which is important as some older adults encounter difficulties with typing or using a mouse. Here also voice interaction comes handy, but not in isolation. Voice-only communication relies on good memory and not much context, which is why it ought to be supplemented by a screen as shown by Kowalski et al. (2019). Chatbots, other than standard hardware voice assistants (smart speakers), offer many response types. Hybrid interfaces (for example: Google Smart Hub or Google Home) built into smartphones allow to generate visual responses, which are integral to voice

Figure 3. Gizmo on Google Home Hub. Sample screen showing the process of infobox editing. DialogFlow has predefined styles for tables so the table design is visually different from the standard infobox on Wikipedia.

commands. Dialogflow provides a few rich response types very useful for our pilot implementation of infobox translation verification:

- **Basic cards**: a bordered box with heading, text, optional image, and button. Excellent UI to summarize information about an article edited or display short Wikipedia information.
- **Browsing carousel**: a list of boxes with heading, text and small pictures. This component is useful for selecting categories.
- **Suggestion chips**: all of the responses can be decorated with many small buttons - suggestions of replies. This type of response is needed in the onboarding process. New users often do not have information about the list of acceptable responses for the current context. Secondly, chips are the best interface for closed tasks. At the same time chips ought not to block any other voice responses, as they are only suggestions.
- **Media responses**: a chatbot can play music and video as a response to the user request which is useful because Wikipedia has media content.
- **Table cards**: displaying tabular data is helpful for editing infoboxes or other structured data on Wikipedia, as seen in Figure 3

A chatbot using rich response types may offer multiple ways of continuing the conversation, either via voice, or by allowing the users to select one (or more) out of many responses with the use of voice, touch, mouse or keyboard, which can be imbued with multimedia elements. Devices such as Google Home Hub or Amazon Echo show added screen support to voice assistants. It is an impressive combination. Each voice command can be accompanied by a visual message (tables, pictures, etc.) This flexibility of possible communication formats makes modern chatbots uniquely accessible.

Wiki edition module

This element gathers the needed contributions from DBpedia or MediaWiki API such as missing information from charts, citations, meta-data or candidates for internal links and collects the needs related to vandalism detection and human verification of a neutral point of view. In our pilot implementation it handles the verification needs related to machine translations of infoboxes. It also handles the input information verification and the resulting Wiki edits as well as controls the generation of the feedback for the user, which includes reminders and statistics on the views of the page/pages the user helped edit to keep their motivation high.

Motivation and well-being

Building a solution that will have a real impact on Wikipedia's quality requires addressing the problem of encouraging users to interact more often.

Based on our previous studies we have identified a handful of heuristics that may work for older adults in this particular use case:

- **competition** – Gizmo: *With 63 corrected errors you are among the top 10 most active editors in Warsaw*"
- **self-imposed (or social) obligation** – Gizmo: *How many articles would you like to edit this week? / You wanted to correct 7 articles and so there are 2 more left for the week.*
- **sense of being useful** – Gizmo: *Last week, 245 people read an article you corrected. Thank you! / Can you improve the article about ...? On average, over 500 people read it every week.*
- **not wanting to be a burden** – Gizmo: *Cognitive training will allow you to enjoy good health for longer. Would you like to correct one more article?*

The above examples of chatbot motivating older people to edit Wikipedia articles show that one can reasonably refer to the main types of motivations known in psychology. However, the effectiveness of this solution will depend on the selection of the right arguments for each person and some of these may be shortlisted during participatory design workshops. Discussion on the detailed methods of automatic selection of the right approach to motivation, however, goes beyond the scope of this article.

To prevent users from betting frustrated and bored while they interact to help increase their well-being it is necessary to monitor indicators that can help formulate relevant, correct, and fast responses and aid the users in completing their goals. We list such disaffection indicators below:

- **Average response time** - an increase relative to the average may suggest boredom, which can lower accuracy - then Gizmo can suggest a category change, or an inappropriate choice of task difficulty, which Gizmo can then correct.
- **Abandoning tasks** - the user can interrupt the conversation at any time. Analysis of the timing (in a conversation, or relative to the time of the day) and task difficulty and category can reduce the abandonment rate.
- **Correct answers** - if the user provides irrelevant/wrong answers (in tasks where verification is possible) then Gizmo can decrease the task difficulty or suggest a break in conversation.
- **Response types** - The user can respond via voice, insert text, or choose predefined responses (chips). In the following tasks and conversations, the bot should not require non-preferred response types for this user.
- **Sentiment Analysis** - Each conversation may be logged and analysed to detect emotional overtones, both in terms of text content as well as speech as in Tashev et al. (2017). Based on this analysis Gizmo can adjust its own tone, suggest a change of category or encourage the user to take a break.

After identifying disaffection, the system can display tasks requiring verification. The user then can verify the correctness of the solution. These tasks can come from historical data of current and other users. This is recommended as Gadiraju et al. (2017) indicates that users of crowdsourcing systems in which the self-assessment system has been implemented solve the tasks 15% more accurately. Secondly, users can be quite interested in verifying the validity of other users work, as in Skorupska et al. (2018).

Additionally, assessing users' emotional reactions is important in HCI since well-being and lack of frustration are necessary to ensure the frequent use of non-essential applications. In chatbot communication, especially outside of the lab, measuring emotional states is challenging as sentiment analysis is limited to text and the collection of hard data (eg. from biometric sensors) is difficult. A unique method of measuring this is the mobile emotion wheel proposed by Warpechowski et al. (2019), which could appear after the conversation with Gizmo concludes. A screen-supported voice interface is an excellent environment to implement this solution.

Example conversation

An example conversation on how the chatbot interacts with the user is given below. The user first chooses the category and subcategory that interest them, and then dives into a specific task. In this conversation about the infobox for the movie

"Forbidden Planet", or "Zakazana Planeta" in Polish, the user corrects the common mistranslation of the word "cinematography" as "kinematografia" to the value of "zdjęcia" that is "pictures", which is the natural way to say it in Polish in this context:

- Gizmo: Greetings! Can you help me with fixing Wikipedia?
- User: Sure
- Gizmo: Great! Tell me a bit about your interests? Are you interested in entertainment, food and drink or people?
- User: entertainment
- Gizmo: Okay, and more specifically, perhaps dance, film, or music?
- User: film
- Gizmo: I have something for you. Look at this infobox about the movie "Zakazana Planeta". Do we need to change anything? [The chatbot displays the infobox on Google Home Hub, see Figure 3]
- User: Change the "kinematografia" to "zdjęcia".
- Gizmo: Okay I did - take a look. Should I change anything else?
- User: No, it is great.
- Gizmo: I saved your changes. Do you want to check another article in this category?
- User: Not today, I have no time.
- Gizmo: Thank you for helping. When would you like to talk again?
- User: Tomorrow morning is fine.
- Gizmo: Okay, I will catch you then. Thank you and bye!
- User: bye

Discussion

In this section we discuss the proposed approach in the context of opportunities related to co-designing such HCI solutions for older adults as well as the growing importance of this area of crowdsourced-based interaction and contribution.

Older adults and conversational crowdsourcing

Many studies have been conducted at the intersection of HCI and ageing, but according to Vines et al. (2015) they largely focus on stereotypes related to health, socialization and technology instead of looking for opportunities for ICT to, for example, tap into the older adults' potential increasing their well-being. One such opportunity is related to improving the accessibility of ICT-based solutions via chatbots; especially ones which allow older adults to remain active, both physically and cognitively which is a determinant of successful ageing, as in the

study by Lum and Lightfoot (2005) on how volunteering helps combat depression and slows negative effects of aging.

Despite the fact that older adults are a very heterogenous group, there are some things many of them have in common. For example, crystallized intelligence, which is largely measured through general world knowledge and language tasks, according to Baghaei and Tabatabaee (2015), may stay the same, or even benefit from aging as indicated by McArdle et al. (2000) as well as Wang and Kaufman (1993). So, older adults may be well-suited to provide contributions which hinge on the understanding of texts, as also the ability to conceptualize and grasp the high level meaning is shown not to degrade significantly with age, as shown by Radvansky (1999) and in some cases to increase. So, they are a good target to verify the logic of information presented in infoboxes in their native language.

At the same time, the fact that interaction with chatbots, especially multimodal interaction supported by touchscreens, allows to see the context of the conversation and does not overload short-term memory is very beneficial for older adults, as much of the cognitive capacity change in ageing happens in this realm. Another benefit of chatbots as interfaces for older adults is that there is no need for the users to learn artificial conventions of communication, as noticed by Morrissey and Kirakowski (2013), such as hard to navigate menus, to achieve their goals of interacting with computers and engaging in online projects, such as contributing to Wikipedia. This solution attempts to map the interaction with the system into natural conversations between people, in which they ask and answer questions. Because all people engage in conversational social interaction the barrier to entry to engaging with a system mirroring this interaction pattern is much lower, and they are encouraged to engage right away, as they learn-by-doing, without the fear of breaking anything.

Participatory design of conversational interfaces for crowdsourcing

One unique opportunity of designing text-based interaction is its potential to co-design it with users, without the need to explain UI elements to facilitate participation, as it was done in a study of participatory design during a hackathon with older adults by Kopeć et al. (2018). Because the text conventions are familiar to everybody the barrier to entry to actively take part in co-design workshops is lower, and there are multiple approaches to co-design that have been applied with older adults, as reviewed by Duque et al. (2019). Such workshops can be conducted in two parts as a mix of multiple activities inspired by the SPIRAL method of enabling user participation, after Kopeć et al. (2018), including:

Part 1 - Ideating and prototyping

- round-table discussion about Wikipedia to evaluate the appearing attitudes,
- quick presentation of Wikipedia, crowdsourcing and some areas of needs as user empowerment

- brainstorming key areas of interest for editing Wikipedia and collecting them as post-its (to fidn out what topics the chatbot could suggest first),
- affinity diagramming the received responses (to decide on "themes" the teams could work on),
- work in teams of older adults and designers to design the interaction flow (for the applications and topics identified above)

Part 2 - (Re-)designing and evaluating

- work in teams to write out possible responses for the elements of the interaction flow,
- using "Wizard of Oz" techniques to verify the ideas between different teams and find problematic areas,
- creating Mad Libs tasks where the users fill out some incomplete conversations with problematic dialogue elements left blank according to their preferences and doing these between teams
- discussing what feedback about their contributions the users would like to receive and how often and voting on the most attractive mode of receiving feedback (eg. every participant gets five star stickers to give out to different options)
- brainstorming the best time and method of initiating interaction with users based on their habits

During these workshops the potential users can co-design the interaction flow and the personality of the chatbot, to best fit their preferences.

One current limitation is that the screen interface implemented in Google Home Hub can not be co-designed, as it is not fully customizable for developers. Still, for this application an HTML box to render formatted Wikipedia articles with infoboxes would be very helpful, as it could print the outcome of the changes made thanks to Gizmo. Additionally, Home Hub does not have all the languages of Google Assistant on Android devices available, which is extremely important, as, at least in Poland, few older adults know foreign languages at a level that would allow them to converse with Gizmo if it is not available in Polish.

Yet, even in their native language selecting the users who would be interested in this form of spending free time can be a challenge, but participatory design workshops can help with making the solution more appealing to older adults. Still, the lower the barrier to entry, the larger the potential that more users will stick and continue to contribute, despite some of the possible problems. This solution is not a universal one for every user, or every older adult, but it aims to be a viable alternative to the Wikipedia's graphical user interface, which has been shown to be confusing, in particular, in a study with older adults by Nielek et al. (2017).

Overall, it is possible to take some issues Wikipedia is facing and attempt to translate them into a conversational interface mediating contributions, which has the potential to simplify and streamline crowdsourcing for some groups of older

adults who would like to contribute, but find it too difficult now, especially if they have the chance to co-design this solution.

The raising importance of making human verification accessible

Wikipedia is the largest repository of knowledge, and due to its simplicity, comprehensiveness and the massive size it is used not only a reference and a source of knowledge for individuals, but also companies, the media and governments. Over the years it has gained trust and expanded significantly. Because of its place in our lives it sets the tone of more than everyday debates, and with increased internet access more people flock to it for information, but often as passive users, without the know-how (eg. new internet users) or technical ease (eg. using mobile devices) to easily contribute their story. Because of the increased volume and readership without the increase in editing volume in the past 10 years, as described by Mandiberg (2020), Wikipedia's inequality in terms of who produces the content in contrast to who consumes it is even more profound.

This is why making Wikipedia editing more accessible to all user groups is a genuine concern, as the most active editors, and the most motivated individuals with greatest resources not only set its tone, but also choose what is worthy of being described, expanded upon or even mentioned.

While it is relatively easy to do remain objective when stating scientific theories or bare facts about someone's life, sections such as Controversies or Critical Reception are more problematic. So, on one hand it is curated by a select group of individuals, while on the other, because of inherent unconscious bias, even the best editors may produce subjective content.

So, making Wikipedia more accessible, by allowing alternate ways of interacting with the editing system and more reliable by aiding in data verification and evaluating if the content is accurate and/or free of bias is important. Especially at a time that organized online campaigns may be used to vandalize or skew information in support of a certain worldview or an interested party.

Although, the proposed solution is not a global Wikipedia editor for hybrid interfaces we recommend this method for verifying infobox translations. This application is in line with one of key areas of needs of Wikipedia, as infoboxes are difficult to maintain, not only due to the many languages where they may appear in different formats, but also because of the need to manually synchronize them with the contents of the article, whenever the article is updated with key information which ought to appear in the infobox.

Conclusions

Our solution reaches promising ground in conversational interface applications for collaborative work, and we believe that the CSCW domain is moving towards the emerging Web 5.0 with its focus on tailored emotive interaction accessible to more users. After all, one of the key benefits of chatbots as interfaces is that there is no

need for the users to learn artificial conventions of communication Morrissey and Kirakowski (2013). Thus, **chatbots hold great potential for simplifying interactions, especially when complex systems and designs arise incrementally, as was the case with Wikipedia or other collaborative environments**. However, there exist multiple challenges to enabling crowdsourcing contributions via conversational interfaces such as the identification of appropriate tasks, in which the need and expertise areas of both the system and the users have to meet, the monitoring of disaffection indicators and finally maintaining the interest and motivation of the users.

To verify these, as a next step we are planning to conduct tests with older adults once Google Home Hub supports the Polish language or with the use of "Wizard of Oz" techniques, if this support does not come in the nearest future. The users will interact with the chatbot and share their experience with the researchers conducting IDIs and analysing raw dialogue data to fine-tune the interaction to be more friendly, intuitive and motivating to also increase older adults' well-being. Additionally, we are looking into organizing participatory design workshops to further adjust the interaction with our chatbot towards what our potential users may expect.

Another aspect for further study are the contribution types which the chatbot could facilitate in connection with what the older adults would be interested to work on, and what other needs of the Wikipedia project, or other crowdsourcing projects, can be addressed with conversational crowdsourcing. Especially in the context of the emerging Web 5.0, we believe conversational crowdsourcing interfaces are an exciting area of inquiry for the CSCW community in the years to come.

References

Aiello, L. M. (2020): 'Ten Social Dimensions of Conversations'. *Medium*.

Ashktorab, Z., M. Jain, Q. V. Liao, and J. D. Weisz (2019): 'Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 254:1–254:12, ACM.

Atay, C., D. Ireland, J. Liddle, J. Wiles, A. Vogel, D. Angus, D. Bradford, A. Campbell, O. Rushin, and H. Chenery (2016): 'Can a Smartphone-based chatbot Engage Older Community Group Members? The Impact of Specialized Content'. *Alzheimer's & Dementia*, vol. 12, pp. P1005–P1006.

Baghaei, P. and M. Tabatabaee (2015): 'The C-Test: An integrative measure of crystallized intelligence'. *Journal of Intelligence*, vol. 3, no. 2, pp. 46–58.

Brandtzaeg, P. B. and A. Følstad (2017): 'Why people use chatbots'. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.

Brewer, R., M. R. Morris, and A. M. Piper (2016): 'Why would anybody do this?: Understanding Older Adults' Motivations and Challenges in Crowd Work'. In: *Proc. of the 2016 CHI Conf. on Human Factors in Computing Systems*. pp. 2246–2257.

Broadbent, E., R. Stafford, and B. MacDonald (2009): 'Acceptance of Healthcare Robots for the Older Population: Review and Future Directions'. *International Journal of Social Robotics*, vol. 1, no. 4, pp. 319.

Candello, H., C. Pinhanez, and F. Figueiredo (2017): 'Typefaces and the Perception of Humanness in Natural Language Chatbots'. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*.

Ciechanowski, L., A. Przegalinska, M. Magnuski, and P. Gloor (2018): 'In the Shades of the Uncanny Valley: An Experimental Study of Human-Chatbot Interaction'. *Future Generation Computer Systems*.

Clark, L., N. Pantidi, O. Cooney, P. Doyle, D. Garaialde, J. Edwards, B. Spillane, E. Gilmartin, C. Murad, C. Munteanu, V. Wade, and B. R. Cowan (2019): 'What Makes a Good Conversation?: Challenges in Designing Truly Conversational Agents'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 475:1–475:12, ACM.

Duque, E., G. Fonseca, H. Vieira, G. Gontijo, and L. Ishitani (2019): 'A systematic literature review on user centered design and participatory design with older people'. In: *Proceedings of the 18th Brazilian Symposium on Human Factors in Computing Systems*. Vitória Espírito Santo Brazil, pp. 1–11, ACM.

Fichter, D. and J. Wisniewski (2017): 'Chatbots Introduce Conversational User Interfaces.'. *Online Searcher*.

Følstad, A., P. B. Brandtzaeg, T. Feltwell, E. L.-C. Law, M. Tscheligi, and E. A. Luger (2018): 'SIG: Chatbots for Social Good'. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, Association for Computing Machinery.

Gadiraju, U., B. Fetahu, R. Kawase, P. Siehndel, and S. Dietze (2017): 'Using Worker Self-Assessments for Competence-Based Pre-Selection in Crowdsourcing Microtasks'. *ACM Transactions on Computer-Human Interaction*, vol. 24, pp. 1–26.

Gnewuch, U., S. Morana, M. Adam, and A. Maedche (2018): 'Faster Is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction'.

Greenfield, E. A. and N. F. Marks (2004): 'Formal volunteering as a protective factor for older adults' psychological well-being'. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 59, no. 5, pp. S258–S264.

Halfaker, A., R. S. Geiger, J. T. Morgan, and J. Riedl (2013): 'The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline'. *American Behavioral Scientist*, vol. 57, no. 5, pp. 664–688.

Hiyama, A., Y. Nagai, M. Hirose, M. Kobayashi, and H. Takagi (2013): 'Question first: Passive interaction model for gathering experience and knowledge from the elderly'. In: *Pervasive Computing and Communications Workshops (PERCOM Workshops), 2013 IEEE International Conference on*. pp. 151–156.

Itoko, T., S. Arita, M. Kobayashi, and H. Takagi (2014): 'Involving senior workers in crowdsourced proofreading'. In: *International Conference on Universal Access in Human-Computer Interaction*. pp. 106–117.

Kanfer, R., M. E. Beier, and P. L. Ackerman (2013): 'Goals and motivation related to work in later adulthood: An organizing framework'. *European Journal of Work and Organizational Psychology*, vol. 22, no. 3, pp. 253–264.

Khan, R. and A. Das (2018): 'Introduction to chatbots'. In: *Build Better Chatbots*. Springer, pp. 1–11.

Kim, A. and S. B. Merriam (2004): 'Motivations for learning among older adults in a learning in retirement institute'. *Educational gerontology*, vol. 30, no. 6, pp. 441–455.

Kobayashi, M., S. Arita, T. Itoko, S. Saito, and H. Takagi (2015): 'Motivating multi-generational crowd workers in social-purpose work'. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. pp. 1813–1824.

Kobayashi, M., T. Ishihara, T. Itoko, H. Takagi, and C. Asakawa (2013): 'Age-based task specialization for crowdsourced proofreading'. In: *International Conference on Universal Access in Human-Computer Interaction*. pp. 104–112.

Kopeć, W., B. Balcerzak, R. Nielek, G. Kowalik, A. Wierzbicki, and F. Casati (2018): 'Older Adults and Hackathons: A Qualitative Study'. *Empirical Software Engineering*, vol. 23, no. 4, pp. 1895–1930.

Kopeć, W., R. Nielek, and A. Wierzbicki (2018): 'Guidelines towards better participation of older adults in software development processes using a new SPIRAL method and participatory approach'. *Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering - CHASE '18*.

Kowalski, J., A. Jaskulska, K. Skorupska, K. Abramczuk, C. Biele, W. Kopeć, and K. Marasek (2019): 'Older Adults and Voice Interaction: A Pilot Study with Google Home'. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. LBW0187:1–LBW0187:6, ACM.

Lee, J. and D. Seo (2016): 'Crowdsourcing not all sourced by the crowd: An observation on the behavior of Wikipedia participants'. *Technovation*.

Lee, M., S. Ackermans, N. van As, H. Chang, E. Lucas, and W. IJsselsteijn (2019): 'Caring for Vincent: A Chatbot for Self-Compassion'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 702:1–702:13, ACM.

Lum, T. Y. and E. Lightfoot (2005): 'The effects of volunteering on the physical and mental health of older people'. *Research on aging*, vol. 27, no. 1, pp. 31–55.

Mandiberg, M. (2020): 'Mapping Wikipedia - An unprecedented data set shows where the encyclopedia's editors are, where they aren't, and why.'. *The Atlantic*.

McArdle, J. J., F. Hamagami, W. Meredith, and K. P. Bradway (2000): 'Modeling the dynamic hypotheses of Gf–Gc theory using longitudinal life-span data'. *Learning and Individual Differences*, vol. 12, no. 1, pp. 53–79.

Mitrevski, M. (2018): *Developing Conversational Interfaces for IOS: Add Responsive Voice Control to Your Apps*. Springer.

Moore, R. J., R. Arar, G.-J. Ren, and M. H. Szymanski (2017): 'Conversational UX design'. In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. pp. 492–497.

Morrissey, K. and J. Kirakowski (2013): ''Realness' in chatbots: Establishing quantifiable criteria'. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.

Morrow-Howell, N., J. Hinterlong, P. A. Rozario, and F. Tang (2003): 'Effects of volunteering on the well-being of older adults'. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, vol. 58, no. 3, pp. S137–S145.

Navarro, J. E. J.-B., J. L. G. Sanz, J. M. Del Castillo, A. C. Izquierdo, and M. M. Rodríguez (2007): 'Motivational factors and physician advice for physical activity in older urban adults'. *Journal of aging and physical activity*, vol. 15, no. 3, pp. 241–256.

Nielek, R., M. Lutostańska, W. Kopeć, and A. Wierzbicki (2017): 'Turned 70?: It is Time to Start Editing Wikipedia'. In: *Proceedings of the International Conference on Web Intelligence*. New York, NY, USA, pp. 899–906, ACM.

Pereira, J. (2016): 'Leveraging chatbots to improve self-guided learning through conversational quizzes'. In: *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '16*.

Radvansky, G. (1999): 'Aging, Memory, and Comprehension'. *Current Directions in Psychological Science*, vol. 8, pp. 49–53.

Redi, M., B. Fetahu, J. Morgan, and D. Taraborelli (2019): 'Citation needed: A taxonomy and algorithmic assessment of Wikipedia's verifiability'. In: *The World Wide Web Conference*. pp. 1567–1578.

Reshmi, S. and K. Balakrishnan (2016): 'Implementation of an inquisitive chatbot for database supported knowledge bases'. *Sādhanā*, vol. 41, no. 10, pp. 1173–1178.

Richards, B., R. M. Bernstein, and C. R. D'Amato (2016): 'Curious Cat Mobile, Context-Aware Conversational Crowdsourcing'. *Journal of Experimental and Theoretical Artificial Intelligence*.

Schutzer, K. A. and B. S. Graves (2004): 'Barriers and motivations to exercise in older adults'. *Preventive medicine*, vol. 39, no. 5, pp. 1056–1061.

Seering, J., M. Luria, G. Kaufman, and J. Hammer (2019): 'Beyond Dyadic Interactions: Considering Chatbots As Community Members'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, pp. 450:1–450:13, ACM.

Skorupska, K., M. Núñez, W. Kopeć, and R. Nielek (2018): 'Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality'. *Proc. ACM Hum.-Comput. Interact.*, vol. 2, no. CSCW, pp. 159:1–159:23.

Suzuki, Y. and S. Nakamura (2016): 'Assessing the quality of Wikipedia editors through crowdsourcing'. In: *Proceedings of the 25th International Conference Companion on World Wide Web*. pp. 1001–1006.

Tashev, I. J., Zhong-Qiu Wang, and K. Godin (2017): 'Speech emotion recognition based on Gaussian Mixture Models and Deep Neural Networks'. In: *2017 Information Theory and Applications Workshop (ITA)*. pp. 1–4.

Thelwall, M. (2016): 'Does Astronomy research become too dated for the public? Wikipedia citations to Astronomy and Astrophysics journal articles 1996-2014'. *El Profesional de la Información*, vol. 25, no. 6, pp. 893.

Vines, J., G. Pritchard, P. Wright, P. Olivier, and K. Brittain (2015): 'An Age-Old Problem: Examining the Discourses of Ageing in HCI and Strategies for Future Research'. *ACM Trans. Comput.-Hum. Interact.*, vol. 22, no. 1, pp. 2:1–2:27.

Vora, P., N. Komura, and S. U. Team (2010): 'The n00b Wikipedia editing experience'. In: *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*. p. 36.

Wang, I., J. Smith, and J. Ruiz (2019): 'Exploring Virtual Agents for Augmented Reality'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA, Association for Computing Machinery.

Wang, J.-J. and A. S. Kaufman (1993): 'Changes in fluid and crystallized intelligence across the 20- to 90-year age range on the K-BIT'. *Journal of Psychoeducational Assessment*, vol. 11, no. 1, pp. 29–37.

Warpechowski, K., D. Orzeszek, and R. Nielek (2019): 'Tagging emotions using a wheel user interface'. In: *Proceedings of the 13th Biannual Conference of the Italian SIGCHI Chapter: Designing the next interaction*. pp. 1–5.

Wikipedia.org (2020): 'Wikipedia: Statistics'. Wikimedia Foundation, Retrieved 27.05.2020 from: https://en.wikipedia.org/wiki/Wikipedia:Statistics.

Wulczyn, E., R. West, L. Zia, and J. Leskovec (2016): 'Growing Wikipedia Across Languages via Recommendation'. *Proceedings of the ... International World-Wide Web Conference. International WWW Conference*, vol. 2016, pp. 975–985.

Yu, H., C. Miao, S. Liu, Z. Pan, N. S. B. Khalid, Z. Shen, and C. Leung (2016): 'Productive aging through intelligent personalized crowdsourcing'. In: *30th AAAI Conference on Artificial Intelligence (AAAI-16)*.

Zhang, Y., T. Paradis, L. Hou, J. Li, J. Zhang, and H. Zheng (2017): 'Cross-lingual infobox alignment in Wikipedia using entity-attribute factor graph'. In: *International Semantic Web Conference*. pp. 745–760.

5.3 A Comparative Study of Younger and Older Adults' Interaction with a Crowdsourcing Android TV App for Detecting Errors in TEDx Video Subtitles (INTERACT 2019, 140 pts)

Title	A Comparative Study of Younger and Older Adults' Interaction with a Crowdsourcing Android TV App for Detecting Errors in TEDx Video Subtitles
Authors	Kinga Skorupska, Manuel Nunez, Wieslaw Kopec, and Radoslaw Nielek
Conference	IFIP TC13 Conference on Human-Computer Interaction [INTERACT 2019] (CORE A, 140 points on the list of the Polish Ministry of Science and Higher Education, as accessed on 5.04.2022)
Published in	INTERACT 2019, Human-Computer Interaction – INTERACT 2019. Lecture Notes in Computer Science, vol 11748. Springer, Cham.
Author Credit	I was the lead author of this article and my key contributions constituted the research design, study protocol design, UI and UX testing, data gathering, data analysis, literature review, conclusions, manuscript drafting and revision. I have also presented this research at the conference (in person).
Description	This paper reports the results of a pilot study comparing the older and younger adults' interaction with an Android TV application which enables users to detect errors in video subtitles. Overall, the interaction with the TV-mediated crowdsourcing system relying on language proficiency was seen as intuitive, fun and accessible, but also cognitively demanding; more so for younger adults who focused on the task of detecting errors, than for older adults who concentrated more on the meaning and edutainment aspect of the videos. We also discuss participants' motivations and preliminary recommendations for the design of TV-enabled crowdsourcing tasks and subtitle QA systems.
Contribution and Impact	This contribution is comprises an age-comparative study of the Android TV Crowdsourcing interface proposed in the earlier study. Overall it indicates that older adults are a good target for engaging with volunteer crowdsourcing, especially if the activities resemble leisure and address their needs and motivations.

A Comparative Study of Younger and Older Adults' Interaction with a Crowdsourcing Android TV App for Detecting Errors in TEDx Video Subtitles

Kinga Skorupska[1,2][0000−0002−9005−0348], Manuel Núñez[1][0000−0002−1197−7488], Wiesław Kopeć[1][0000−0001−9132−4171], and Radosław Nielek[1][0000−0002−5794−7532]

Polish-Japanese Academy of Information Technology
kinga.skorupska@pja.edu.pl
http://www.pja.edu.pl

Abstract. In this paper we report the results of a pilot study comparing the older and younger adults' interaction with an Android TV application which enables users to detect errors in video subtitles. Overall, the interaction with the TV-mediated crowdsourcing system relying on language proficiency was seen as intuitive, fun and accessible, but also cognitively demanding; more so for younger adults who focused on the task of detecting errors, than for older adults who concentrated more on the meaning and edutainment aspect of the videos. We also discuss participants' motivations and preliminary recommendations for the design of TV-enabled crowdsourcing tasks and subtitle QA systems.

Keywords: Crowdsourcing · Smart TV · Android TV· Design evaluation· Subtitles · Older adults · Younger adults.

1 Introduction and related works

With the increasing amount of video content it is necessary to ensure its accessibility to the deaf, the hard of hearing and international audiences through quality same language and multilingual subtitles. Therefore, crowdsourcing subtitle quality assurance (QA) models are an important research frontier, especially as subtitles are often created by volunteers, as in the case of TED and TEDx [7] or generated automatically. At the same time, there are groups who may benefit from more fun and accessible crowdsourcing projects.

For example, older adults, who comprised 19.2% of the EU-28 population in 2016 [1], benefit from all forms of volunteering, as it slows the negative effects of aging and helps combat depression [10]. Yet, there exist multiple barriers to their inclusion in typical crowdsourcing tasks, such as lower ICT skills, uncomfortable and costly setup of such solutions [16], unfamiliar interfaces and lack of motivation due to unclear personal benefit [4], unsocial nature of the task [18] or their perception of not being qualified [8].

Younger adults, on the other hand, who are more open to online crowdsourcing and microtasking, comprise a significant number of online video viewers, as, according to We Are Flint about 96% of people in UK and US aged 18-34 watch YouTube videos [2]. Both groups are relevant to the development of TV-enabled subtitle QA crowdsourcing tasks as potential contributors and audience.

Therefore, the key research goal was to validate a novel interface for creating no-grind crowdsourcing solutions, ones that do not rely on tedious repetition, with two relevant user groups. To do this, we deployed a Smart TV-based system based on best practices of designing for older users [6] [12] with a comfortable at-home setup, large screen size, and remote relying on familiar interaction patterns [13] with engaging edutainment crowdsourcing tasks. This lowered ICT and other participation barriers and allowed us to signal some possible differences in the participants' approach, motivation, mode of use, experience and expectations. We lay ground to the discussion of the extent to which one may build a universal crowdsourcing system suited to the needs of these different groups, to tap into their potential, facilitate social inclusion and build social capital.

2 Methods

2.1 Comparative study design

To explore these considerations we conducted a comparative qualitative study in the course of which we compared results from a study involving older adults [17] to the results of a study with younger adults conducted in February-March 2019.

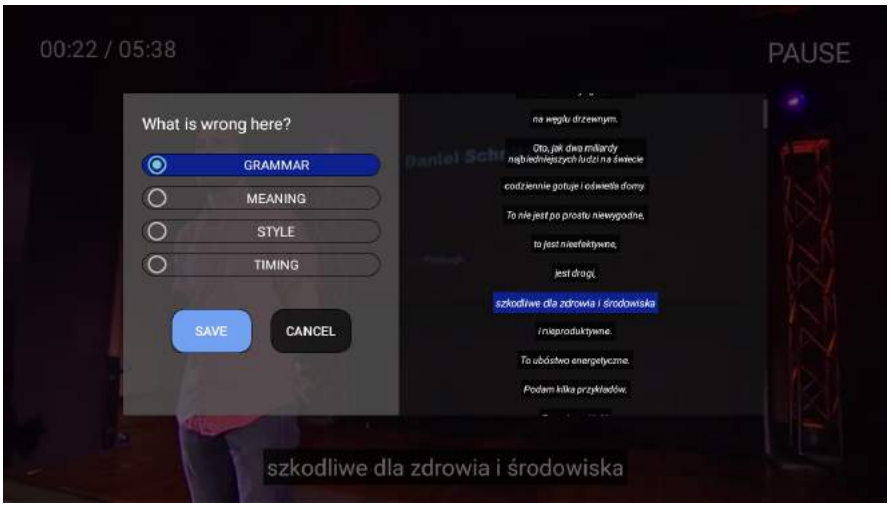


Fig. 1. The error category selection overlay in our Dream TV application

The study examined the interaction with the DreamTV application we created [17] which allows users to watch TEDx videos with volunteer-created subtitles retrieved from Amara API. Once they spot an error they can pause the video to display an overlay (Fig. 1) where they choose the error category among grammar, meaning, style and timing. These error categories were chosen based on preliminary tests and research to be more intuitive than existing models of quality assessment of subtitles by professionals [15] and to aid in improving the subtitles later within the pipeline or during post-editing.

The research protocol, which took about two hours to complete, involved individual testing at participants' homes, where an Android TV set-top box was connected to participants' TV sets, to provide the most natural use conditions, as proposed in multiple studies on Living Labs [9] [3]. It consisted of the DigComp survey[1], a semi-structured interview to evaluate experience with subtitles, the explanation of the project, that is the study and its benefits, an introduction to subtitles and a subtitle error detection written exercise, an app demonstration and a hands-on test, free interaction with the application and our pre-selected test videos (two in Polish, three in English) with redacted Polish subtitles.

For our study we selected five videos to represent different challenges. They were controlled for topic, length, source language (spoken), ease of comprehension and errors: saturation, category and source, either machine (using SubtitleEdit and Google Translate) or organic human or introduced by researchers based on common errors lists on TED Translators' wiki[2]. The videos selected and errors introduced allowed us to observe a variety of factors at play, in order to gather diverse insights to determine interesting areas of further inquiry.

2.2 Participants

We invited seven older adults (O1-O7) and seven younger adults (Y1-Y7) to participate in our study, in each case three female participants and four male participants. We controlled for age, occupation and ICT skills ("above basic proficiency", which is the highest level in DigComp). All participants live in Warsaw, the capital city of Poland. For older adults all owned TVs, including two Smart TVs, and had a dedicated entertainment space in their living room. There was a 20 years age span: the youngest participant was 60 years old and the oldest one was 79, mean 70.85 (SD=6.87). For younger adults we recruited a group that would share the most relevant characteristics with our older adults, especially in terms of their housing situation and entertainment setup, which meant that in Poland they had to be between 25-35 years of age. All but one participants owned Smart TVs and had their own dedicated entertainment space in the living room. All were professionally active and none of them had children. The age span was 5 years, as the youngest participant was 28, and the oldest 33, mean 30.71 (SD=2.28).

[1] A survey measuring indicators of Digital Competence based on the Digital Competence Framework [5].

[2] The TED Translators' wiki containing lists of common errors can be found at: https://translations.ted.com

3 Results and discussion

Overall, using the application was enjoyable, intuitive and easy for both younger and older participants, however there were differences in their approach to the task. While our group of younger adults saw it as an enjoyable activity one could do to improve subtitles, brag or supplement their income in a fun way, our group of older adults viewed it less as work and more an opportunity to learn something and did not expect payment for contributing. For older adults it was more interesting, as they were given access to resources they were unlikely to reach to on their own (TEDx videos) whereas younger adults agreed that they know less demanding or better entertainment. Younger adults detected more mistakes than older adults as they viewed the task to be more work-like and in consequence, demanding. Older adults seemed more lenient, especially when it came to style and punctuation, and focused more on the content of the videos, rather than correcting mistakes. There were also differences in feedback. Where older adults focused on ways to find videos that would be a better fit for them thematically, younger adults focused more on critiquing the error categories chosen and comparing the application to Netflix. This is due to the differences in experience with such services. Both groups found the interaction via the remote to be very convenient and well-suited for this activity and they learned to comfortably use the application in just one session, with older adults in general taking more time to learn and later to navigate, but with no significant other differences.

3.1 Error detection

Reflexes Overall, all of the older participants paused the videos one subtitle too late, and had to use the dialog list to navigate back to the subtitle where they wanted to mark the error. The same was true of all but one younger adults, as Y2 paused even before the speaker finished the sentence, indicating that they read rather than listened. This suggests that access to the full dialog list is necessary in this type of crowdsourcing for all age groups.

Number of errors found In general, younger adults found more errors than older adults which may be related to their attitude towards this activity. While younger adults focused on the task of finding errors, older adults engaged with the content of the videos more and felt that they are learning new interesting things (O1-O4). This is in contrast with younger adults, except for Y4, who admitted to focus more on the content and commented that they "should watch such videos more often as they are interesting". Consequently, younger adults found many more punctuation errors, which older adults often ignored. This may be as punctuation errors do not interfere with understanding. Older adults, who focused more on understanding the content, often chose the "meaning" category, when something was not clear to them (e.g. "it is not explained what is this photon" or "Spiderman, this is not Polish" by O2 and "kryptonite, must be a mistake" by O5, O6), suggesting the application could benefit from a built-in

dictionary. Older adults' focus on meaning is in line with Radvansky's research on the effect of aging on memory and comprehension, suggesting that while lower levels of memory, which may be responsible for remembering specifics such as punctuation, deteriorate with age, the ability to form situation models on a higher level, aiding in meaning and general comprehension is less affected [14]. Moreover, different people found very different errors, depending on their interests and background (science for Y6: "the Sun vs the sun", detailed punctuation rules for Y2 with linguistic background) which shows that the effect of scale by relying more on quantity and not quality of contributions may work well here.

Error categories All but one of the younger participants (Y1-Y6) encountered errors in subtitles to which they wished to assign more than one error category, to remove the analysis paralysis of choosing the best fitting category ("People like me would deliberate 3 years over a single word" Y1) and likely to satisfy their need for cognitive closure [11], as many younger participants found the categories to be "fuzzy". The other participant, Y7, said that "these are short lines so if someone marks a mistake it is easy to know what it is" and proposed to remove categories, the same could be seen in O3's eagerness to just mark mistakes quickly and continue watching the videos.

Younger adults remarked that "synchronization is the most intuitive" (Y1). Other error categories requested were "punctuation" (Y6) and "subtitle division" (line breaking) (Y3) and "technical errors" such as subtitle convention errors as a separate category (Y1, Y2) and both Y7 and Y3 said that knowing subtitle conventions requires a lot of practice, and pre-teaching, for which Y3 suggested a mini-game, while older adults wished for an in-application tutorial to ensure they do not make mistakes when marking mistakes (O1, O4). One participant, Y6, also said there ought to be a way to mark recurring errors ("Here I would have to mark a lot of things, because the Sun should be written with capital letter, and it repeats a lot"), on the other hand O3 remarked "He made the same mistake, but I'll overlook it now", eager to continue watching.

Older adults (O1-O7) did not question the error categories even though they often could not decide which category to choose (O4, O5) and sometimes deliberated aloud (O3). This may be because older adults are less likely to criticize design choices in the context of technology, as they feel they lack experience in it so they are not confident enough to know they can contribute. This was also observed in the context of participatory design by Kopec et al [8]. Also, even though some older adults had to sit closer to the screen to read (O1, O3) it was a younger adult (Y6) who voiced that they would like the interface to be bigger.

In conclusion, to ease the choice of error categories we propose to present them in the order of importance, with the top category being "meaning" - answering the question "Is this subtitle understandable?", followed by "grammar", as it includes common punctuation mistakes, and then "style", which would have to be explained as relating to technical errors, and including also other problems. We postulate that because of conflicts of simultaneous work it is very difficult to find synchronization errors, while also looking for other types of errors ("It

is difficult to catch problems with synchronization - you focus on all the other mistakes" Y3, and "I had to read" O5). This was seen in the tests with older adults, who found no synchronization errors (O1-O7), and younger adults who rarely marked them as they found it tiring to both read, and listen (Y7: "I did not listen to the guy", Y6: "difficult to focus on what the person was saying") Signalling the relationship between enjoyment, interest and errors found Y6 said: "this topic was interesting, sometimes I did not focus on finding mistakes". Both older (O1-O4, O6) and younger adults (Y3, Y4, Y6) seemed to find fewer errors the more they enjoyed the video, with Y4 saying that they were "forgetting to read". The enjoyment was also negatively correlated with the number of errors marked, with Y2 saying that "The errors were so thickly distributed, it is a very tiring video" and that "If there were fewer errors it would be more fun than work" and Y5 mentioning that "If you have to focus only on subtitles it is more like work, but if you get to mark glaring errors only it is more entertainment".

3.2 Fun or work?

Y1 and Y5 found the application to be very fun, commenting that "you can point out someone's mistakes without arguing with that person, everyone loves that!" (Y1), adding that it is true especially when there are people around, and "How fun! I like it! I could do it all my life" (Y5). Y6 also said "it's cool, I like nitpicking". The other participants commented that it would be work if you "had to do it, like an editor in a paper" and "the movies are not long, and you can take breaks" (Y7). Similarly, Y3 mentioned that "you should be able to choose how long video you want". This aspect of controlling time was also present in older adults' feedback, as they enjoyed the ability to pause the video at will, take breaks, and O3 even said "The movies should be shorter, then I could watch anything! Just give me ten 5 minute films and I can do that for an hour". Older adults overall focused on the educational aspect of the task, saying that it is good practice and one can "learn a lot" (O1-O4) from these videos. This aspect was less prominent with younger adults, who often treated the experience almost job-like as it was "mentally demanding" and felt more like "work", or that it is a bit like an "exam" (Y1) and felt judged when they did not understand a subtitle (Y3) ("I don't know what they mean by "last mile" and since it was in quotation marks it must be something that everyone knows, so now I feel stupid"). In contrast, only O4 mentioned that "It is tiring, I am not that young anymore." drawing attention to the task's cognitive load.

3.3 Motivation, Gamification and Rewards

While older adults' participants motivation was mostly based on the value for them, in terms usefulness, relevance to their interests and staying active, for younger adults there was almost no concern about the topic as they viewed the task to be more "work-like" and focused finding errors more than understanding and enjoying the content - likely because they have other entertainment readily available. Detailed comparison of approaches and attitudes is visible in Table 1.

Table 1. Comparison of older and younger adults' motivations, rewards and wishes

	Younger adults	**Older adults**
Pointing out mistakes	Y1, Y2, Y5, Y6	O3
Social activity	Y1: "to do with friends"	O2: "with grandchildren"
Helping somebody	Y1: "If some friend asked me to do this for them, I would help them", Y4	
Learning new things	Our group of younger adults could watch such videos, but just watch as Y4: "they are interesting" to Y6: "focus on the content".	O1-O4, O1: "I learned a lot", O3 "I would watch movies about health, global issues, climate change or politics" but: O5 "The topics would have to be useful"
Getting paid	Y1-Y7, except for Y5: "Nobody would pay much, it's better to have bonuses, like a subscription or a small gift because earning little money is meh"	
Improving the world	Y1: "I like it, if I was convinced myself that this is making the world a bit better, then this is a convenient way to help"	
Challenging oneself cognitively	Y2 and Y3, but about other people, Y3: "blue-collar workers" and "stay at home moms" who can do it for fun and Y2: "retired people to stay active".	O3: "This task is great for old people, but only those who are mentally fit, so that they dont deteriorate"
Passion for the topic	Y3 mentioning feminists: "people who are very passionate about a topic can contribute"	
Statistics of confirmed contributions	Y4: "ranking like on Memrise", Y6: "ranking of best reviewers", Y3: "a community to care about my achievements listed on my profile". Interestingly, both Y1 and Y2 mentioned they do not need statistics.	
Helping improve subtitles being used	Y3: "that there were 100 people who watched this film with improved subtitle in a month would mean something"	
Access to training	Y3, "in the community access to games that help you develop skills to contribute better"	O1: "It would be good to have a testing mode, to be able to train without consequences", O4
Addressing glaring errors	in videos they are watching with subtitles anyway (Y4, Y5)	
Reliance on linguistic experience	Y3: "I like that I don't have to learn anything to start doing it, I know the language"	O3: "There should be more subtitle testers like me, but not young people because they have little experience"

3.4 Sustainability

Overall, although most of the participants found this activity to be fun, there are doubts whether they would do it in the long run without other incentives. The tests with older adults suggest that some may continue using the application as an easy foray into the world of edutainment and to stay active, except for O5 who stated "I manage, but it is not my thing - the topics would have to be useful" and O4 who expected to be bored as one has to "be focused". On the other hand, some younger adults commented "I wouldn't do it because it is time consuming, when you watch something to gain knowledge it is easier to understand the content if you are just watching" (Y6) or "it's not my type of thing, I am not a linguist and correcting errors is not my passion" (Y7). They also mentioned shortage of time (Y2, Y3) and the demanding nature of this task (Y2, Y3, Y6, Y7) as a problem. For younger adults, who have formed habits regarding their access to other forms of entertainment, it may work best as a feature integrated into their familiar experience. Both Y4 and Y5 suggested that such activity could be "integrated into a player" they use anyway", on YouTube for Y4 ("it could be great if YouTube had something like that in their automatic subtitles, which now suck") or on VOD for Y5, who noted that "Sometimes I am tempted to mark something on VOD - there are few people who would bother to go to a film distributors' website and report errors in subtitles". Y5 concluded that "If it was easily accessible then a lot of people would do it, if they could just mark something on their remote".

4 Conclusions

As this is a pilot study with a small number of participants it is important to verify the following preliminary findings. While this task is fun for both younger and older adults, the former treat it more like work and expect payment. This group would benefit from having a similar solution integrated into their entertainment medium of choice. On the other hand, older adults are a promising target for this type of crowdsourcing, as it not only provides them with content they may otherwise miss, but also allows them to learn and stay active.

Future work ought to explore TV-mediated crowdsourcing in larger studies, and focus on the patterns of interaction with this solution, including the timing of engagement and quantitative relationship between the enjoyment of the video and the number of subtitle errors found. It is also important to verify if this TV-mediated crowdsourcing solution can hold older adults' interest over time, and if so, what are other ways such mode of interaction can be used to allow older adults to stay active for longer, contribute to society and learn new things.

This research in part was supported by the Polish National Science Center grant 2018/29/B/HS6/02604 and the European Unions Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 690962.

References

1. Population structure and ageing. http://ec.europa.eu/eurostat/statistics-explained/index.php/Population_structure_and_ageing (2017)
2. Social 2018 main findings. https://weareflint.co.uk/main-findings-social-media-demographics-uk-usa-2018 (2018)
3. Alaoui, M., Lewkowicz, M.: A livinglab approach to involve elderly in the design of smart tv applications offering communication services. In: International Conference on Online Communities and Social Computing. pp. 325–334. Springer (2013)
4. Brewer, R., Morris, M.R., Piper, A.M.: Why would anybody do this?: Understanding older adults' motivations and challenges in crowd work. In: Proc. of the 2016 CHI Conf. on Human Factors in Computing Systems. pp. 2246–2257. ACM (2016)
5. Ferrari, A.: Digcomp: A framework for developing and understanding digital competence in europe (2013)
6. Fisk, A., Czaja, S., Rogers, W., Charness, N., Sharit, J.: Designing for Older Adults: Principles and Creative Human Factors Approaches, Second Edition. Human Factors and Aging Series, CRC Press (2009)
7. Cámara de la Fuente, L.: Multilingual crowdsourcing motivation on global social media, case study: Ted otp. Sendebar **25**, 197–218 (2014)
8. Kopeć, W., Balcerzak, B., Nielek, R., Kowalik, G., Wierzbicki, A., Casati, F.: Older adults and hackathons: A qualitative study. Empirical Software Engineering **23**(4), 1895–1930 (2018). https://doi.org/10.1007/s10664-017-9565-6
9. Kopeć, W., Skorupska, K., Jaskulska, A., Abramczuk, K., Nielek, R., Wierzbicki, A.: Livinglab pjait: Towards better urban participation of seniors. In: Proceedings of the International Conference on Web Intelligence. pp. 1085–1092. WI '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3106426.3109040
10. Lum, T.Y., Lightfoot, E.: The effects of volunteering on the physical and mental health of older people. Research on aging **27**(1), 31–55 (2005)
11. M. Webster, D., Kruglanski, A.: Individual differences in need for cognitive closure. Journal of personality and social psychology **67**, 1049–62 (01 1995). https://doi.org/10.1037/0022-3514.67.6.1049
12. Pak, R., McLaughlin, A.: Designing Displays for Older Adults. Human Factors and Aging Series, CRC Press (2010)
13. Pan, Z., Miao, C., Yu, H., Leung, C., Chin, J.J.: The effects of familiarity design on the adoption of wellness games by the elderly. In: 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). vol. 2, pp. 387–390. IEEE (2015)
14. Radvansky, G.: Aging, memory, and comprehension. Current Directions in Psychological Science **8**, 49–53 (04 1999). https://doi.org/10.1111/1467-8721.00012
15. Romero-Fresco, P., Pchhacker, F.: Quality assessment in interlingual live subtitling: The ntr model. Linguistica Antverpiensia, New Series Themes in Translation Studies **16**(0) (2018)
16. Sandhu, J., Damodaran, L., Ramondt, L.: Ict skills acquisition by older people: Motivations for learning and barriers to progression. International Journal of Education and Ageing **3**(1), 25–42 (2013)
17. Skorupska, K., Nunez, M., Kopec, W., Nielek, R.: Older adults and crowdsourcing: Android tv app for evaluating tedx subtitle quality. Proc. ACM Hum.-Comput. Interact. **2**(CSCW), 159:1–159:23 (Nov 2018). https://doi.org/10.1145/3274428
18. Vines, J., Blythe, M., Dunphy, P., Monk, A.: Eighty something: banking for the older old. In: Proceedings of the 25th BCS Conference on Human-Computer Interaction. pp. 64–73. British Computer Society (2011)

5.4 Older Adults' Motivation and Engagement with Diverse Crowdsourcing Citizen Science Tasks (INTERACT 2021, 140 pts)

Title	Older Adults' Motivation and Engagement with Diverse Crowdsourcing Citizen Science Tasks
Authors	Kinga Skorupska, Anna Jaskulska, Rafał Masłyk, Julia Paluch, Radosław Nielek, Wiesław Kopeć
Conference	IFIP TC13 Conference on Human-Computer Interaction [INTERACT 2021] (CORE A, reranked to B in 2021, 140 points on the list of the Polish Ministry of Science and Higher Education, as accessed on 5.04.2022)
Published in	Ardito C. et al. (eds) Human-Computer Interaction – INTERACT 2021. INTERACT 2021. Lecture Notes in Computer Science, vol 12933. Springer, Cham.
Author Credit	I was the lead author of this article and my key contributions constituted the research concept and design, study tool and protocol design, data gathering, qualitative and partial quantitative data analysis, figure design, literature review, conclusions, manuscript drafting and revision. I have also presented the research outcomes at the conference (remotely).
Description	In this exploratory study we evaluated the engagement, performance and preferences of older adults who interacted with different citizen science tasks. Out of 40 projects recently active on the Zooniverse platform we selected top ones to be represented by 8 diverse, yet standardized, microtasks, 2 in each category of image, audio, text and pattern recognition. Next, 33 older adults performed these microtasks at home and evaluated each task right after its completion to, finally, share what could encourage them to engage with such tasks in their free time. Based on the results we draw preliminary conclusions regarding older adults' motivations for engaging with such crowdsourcing tasks and suggest recommendations for task design while discussing interesting avenues for further inquiry in the area of crowdsourcing for older adults.
Contribution and Impact	This study is unique in that the evaluation of older adults' crowdsourcing experience and impressions about solving different crowdsourcing tasks were done right after each task, and overall, after the completion of the remote study scenario. Moreover, the study provides details on the performance of older adults in different task types, showcasing that they are a good target for online crowdsourcing.

Older Adults' Motivation and Engagement with Diverse Crowdsourcing Citizen Science Tasks

Kinga Skorupska[1 [0000−0002−9005−0348]], Anna Jaskulska[2 [0000−0002−2539−3934]], Rafał Masłyk[1 [0000−0003−1180−2159]], Julia Paluch[1 [0000−0002−7657−7856]], Radosław Nielek[1 [0000−0002−5794−7532]], and Wiesław Kopeć[1 [0000−0001−9132−4171]]

[1] Polish-Japanese Acaemy of Information Technology, Warsaw, Poland
kinga.skorupska@pja.edu.pl
[2] Kobo Association, Warsaw, Poland

Abstract. In this exploratory study we evaluated the engagement, performance and preferences of older adults who interacted with different citizen science tasks. Out of 40 projects recently active on the Zooniverse platform we selected top ones to be represented by 8 diverse, yet standardized, microtasks, 2 in each category of image, audio, text and pattern recognition. Next, 33 older adults performed these microtasks at home and evaluated each task right after its completion to, finally, share what could encourage them to engage with such tasks in their free time. Based on the results we draw preliminary conclusions regarding older adults' motivations for engaging with such crowdsourcing tasks and suggest some guidelines for task design while discussing interesting avenues for further inquiry in the area of crowdsourcing for older adults.

Keywords: Crowdsourcing · Older adults · Citizen science · Motivation.

1 Introduction and Related Works

The area of crowdsourcing for older adults is both underappreciated and underexplored and developing sustainable solutions for older adults is still challenging [14, 13]. This may be due to multiple barriers both specific to the required ICT-skills [1] and the nature of crowdsourcing microtasks. Older adults differ from the younger generation in their online behavior and decision-making [9] and they seem more selective when choosing their engagements [5], which, alongside their generally lower ICT skills, may explain how little interest they expressed in the Mechanical Turk platform populated by tedious and repetitive tasks [3] and lacking a suitable motivation to participate in crowdsourcing, as tasks are not challenging, fun or easily relatable. This is in line with research placing the average age of crowd workers at around 20-30 years [15] [24]. On the other hand, crowd-volunteering tasks, often called citizen science tasks, such as the ones found on the Zooniverse platform [26] can appeal to a more balanced representation of contributors, as about 15% of the platform contributors self-report

as retired.[3] There are also some crowdsourcing systems designed specifically for older adults which mitigate technology barriers, as in Hettiachchi et al. [10] and tap into their knowledge and skills, such as tagging historical photos as in Yu et al. [29], proofreading, as in Itoko et al. [12] and Kobayashi et al. [17], or both as in Skorupska et al. [28] They often rely on motivations that are pro-social, as in Kobayashi et al. [15] and also social, as in Seong et al. [25] which is a trademark of Zooniverse. The Zooniverse platform allows crowd workers to support science projects at a larger scale by solving difficult tasks thanks to the impressive potential of such contributions [2] on a diverse crowdsourcing landscape of Zooniverse (www.zooniverse.org), which is why we have chosen this platform a to serve as the basis for this research. So, there is an opportunity to tap into the potential of older adults as crowd workers with a lot to offer and time on their hands - especially that their share in the society is increasing, and in 2019, "more than one fifth of the EU-27 population was aged 65 and over". [6]

The question whether crowdsourcing tasks are effective in keeping older adults cognitively engaged is relevant, as volunteering activities [22] in general may increase older adults' well-being [23], improve their mental and physical health [21] and can be seen as a protective factor for their psychological well-being [7, 8], potentially delaying the onset of age-related issues [19]. Therefore, in this study we want to gain insights into older adults' motivation and engagement with online citizen science tasks and uncover some guidelines for designing and presenting crowdsourcing citizen science tasks to this group. In designing our research we took care to uniformly present the wide-range of real crowd-volunteering tasks often appearing in citizen science projects. Only after older adults have completed each task we asked them how to improve it, and finally what would motivate them to engage with such tasks in the future.

2 Methods

In this study 33 older adults were asked to complete and evaluate 8 diverse, but standardized citizen science tasks at home, in an unsupervised environment. The study consisted of a short socio-demographic survey including questions about the participants' age, sex, education, activity, ICT-use, and crowdsorucing preferences based on Seong et al. [25]. These questions were followed by a set of 8 different tasks chosen based on expert knowledge of the research team, localized into Polish and presented in an uniform way, broken into 4 pages - each page for a different type of a task. There were two tasks (one easier, and one more difficult/abstract) in each category of **image recognition (PIC)** for tasks T1 and T2, **audio recognition (AUD)** for tasks T3 and T4, **document transcription (DOC)** for tasks T5 and T6 and **pattern recognition (PAT)** for tasks T7 and T8, visible in Fig. 1 in order. The tasks were selected out of 40 community-chosen projects active on the Zooniverse platform in the 2019-

[3] Survey results were presented in a post: https://blog.zooniverse.org/2015/03/05/who-are-the-zooniverse-community-we-asked-them/

20 academic year and spotlit in the publication "Into the Zooniverse Vol. II",[4] published on the 17th of November 2020.

The final standardized tasks were as follows:

Fig. 1. Visual overview of the tasks; T1-T8 from the left to the right.

- **T1** Recognizing animal sillouettes - multiple choice of animal silhouettes, including human visible, no animal and other (representing MichiganZoomIn)
- **T2** Recognizing cat fur types on cat images - multiple choice with abstract images of a cat pelts with fur patterns (similar to image recognition tasks)
- **T3** Recognizing radio programs (97s-long recording) - checkboxes and a follow-up open answer about specifics (representing Vintage Cuban Radio)
- **T4** Recognizing local urban sounds (10s-long recording) - checkboxes with pre-defined answers and an other option (representing Sounds of NYC)
- **T5** Transcribing key information from a hand-written birth certificate of a person born in 1887 - 4 open short answer questions about the dates, name, and the location (representing tasks such as Every Name Counts)
- **T6** Transcribing a longer (346 characters) typewritten text on a specific subject - 1 open long answer question (representing tasks relying on longer transcription of typewritten documents)
- **T7** Recognizing Aurora Borealis patterns (6s-long recording) - multiple choice question with names of patterns and colours (representing Aurora Zoo)
- **T8** Recognizing eye elements in eye pictures on a coordinate grid: two drop-down questions about coordinates and one multiple choice on the visibility of veins. (based on Eye for Diabetes; image by Mikael Häggström [11])

There was a short standardized introduction to each task explaining its importance and purpose, as not to bias the participants with the quality of the project presentation, which can vary considerably between projects. Then, the participants performed each example microtask. After each task we asked the participants to judge, on a 3-point scale, its: **attractiveness**, **importance**, **ease**

[4] The book is available for download here: https://blog.zooniverse.org/2020/11/17/into-the-zooniverse-vol-ii-now-available/

of performing it, **engagement**, and **if they would like to perform similar tasks in the future** and to suggest ways in which each task could be improved. The study protocol was positively evaluated by our ethics committee. The study itself was built in Google Forms and it took between 20-35 minutes to complete, depending on the amount of feedback given after each task and the ICT proficiency of the participants. Finally, after completing all tasks the participants were asked what could encourage them to engage in such tasks in general and whether they have done similar tasks in the past. The suggestions of motivating factors were inspired by an article by Campo et al [4] as well as a wide body of research on crowdsourcing and volunteering.

3 Results and Discussion

3.1 Participants

There were 33 participants who completed and evaluated the chosen crowdsourcing tasks between Dec. 2020 and Feb. 2021. They were recruited from among the participants of our Living Lab [18] via e-mail as unpaid volunteers as we did not want to interfere with their motivation with a financial incentive. 22 participants were in the 60-69 age group, and 10 in the 70-79 group and 1 in 80+ group.[5] All of them were based in Poland, Polish and all but 6 of them came from larger cities (over 200k) and 21 of the participants had higher education. In Table 1 we can see the results concerning volunteering motivation before performing tasks for our 33 participants contrasted with results by Seong et al [25].

	What would encourage you to engage with online or offline volunteer projects? n=33		Values older adults wanted from game experience [25] n=12	
	No. of P.	% of P.	No. of P.	% of P.
Physical improvement	10	30.3%	3	25.0%
Cognitive improvement	15	45.5%	4	33.3%
Opportunity to learn something new	26	78.8%	5	41.7%
Opportunity to communicate and interact with people	14	42.4%	8	66.7%
Opportunity to participate and contribute to society	11	33.3%	4	33.3%
None of the above	4	12.1%	-	-

Table 1. Motivations of volunteer participants before the volunteer experience.

Our participants use the following devices: 28 use a smartphone, 25 a laptop, 18 a desktop PC, 13 a tablet, 8 a SmartTV while 4 a smartwatch or a smartband and 2 a VR headset. They are also avid Internet users as 28 of them use the Internet either a few times a day or every day, and only 5 a few times a week or less often. As such, our participant group would be a good target for online volunteering and crowdsourcing tasks. Yet, after having completed the study 28 participants reported that they have never done similar tasks before, 3 of them said they did similar tasks at work and 2 did such tasks while volunteering.

[5] We have chosen to use multiple choice for age groups as not to bias the participants with an assumption that the research was targeted at older adults.

3.2 Performance and Feedback

Image Recognition Tasks In **T1** 26 participants correctly identified the animal silhouette, 4 pointed to other silhouettes, 1 answered that there was no animal present while 2 more chose the "other" option where they have given in one case each: the name of the animal, the more detailed description of the animal. After completion 2 participants suggested to have a video instead - and 2 others wished the task was more challenging, while 1 complained the question was imprecise. Additionally, 1 person wished there were more animals to spot and "better hidden". In **T2** there was less agreement with 13 people choosing pattern 3, and 7 each voting for patterns 1 and 2, 2 for pattern 4, 3 saying it is "hard to tell" and 1 deciding it was some "other" pattern. The suggestions were to have "a different view of the cat in the picture" (1), a "couple of different pictures" (2) and comments appeared that "if someone does not like cats nothing can improve this task", but also "I liked it, even more so, because I like cats".

Audio Recognition Tasks In both **T3** and **T4** our participants had no trouble listening to the recordings. In both tasks the majority of participants successfully identified the key audio elements (in T3: "many male voices" (31) in T4: "bells" (28) and "traffic" (25)). In T4 about half identified other elements ("birds singing", "people talking", "music"), while only one person noticed "barking", and one indicated "there was nothing specific" in the recording. Additionally, over half of the participants (18) chose to provide additional comment about the exact content of the radio recording from **T3**. For two people the **T3** recording was too short, for another too long and one wished it was accompanied by visuals. The feedback for **T4** was to have a longer recording (5), 1 person also wished for "more variety, to make it more difficult, but also more interesting" and to show a visual connected to the sound. One participant admitted that they "heard birds singing only upon the second hearing" but they were not sure.

Document Transcription Tasks Both transcription tasks were done very well. In **T5** only one participant decided that the text of the birth certificate was "intelligible" and only in 1 out of 4 places, while two people provided only the first name of the person. Among the others there is almost perfect agreement about what the text says with a varying level of detail for the name of the place and date notation. Additionally, one person provided a full transcription of the document, even though the task did not require it. Appearing suggestions were to have more information about the person in the birth certificate (2), to have more similar documents (1) and the participant suggesting it said that they have "transcribed 247 pieces of disappearing poetry before" and are experienced and now working on transcribing "very difficult historical letters". Another participant suggested to have documents related to the participants' own personal history (1). In **T6** most people (22) provided a complete transcription of the 346-character long text, on top of the transcription one person commented "(placing this dot here is incorrect - transcriber's comment)", while 3 wrote that the text was legible,

and 6 provided an incomplete transcription, of these 2 added that the text was legible. Additionally, 2 wrote that it was intelligible. Two people wished for a more challenging text with a harder to read font, and one said other types of content interest them, other appearing comment was "For those who have not been involved in reading old manuscripts and other documents, this is a remarkably interesting activity (...) engaging and motivating, others will put it off or give up. I like it, it draws you in" while another participant mentioned that "such tasks require patience, they are not for everybody".

Pattern Recognition Tasks In **T7** the count of choices was 12, 7, 6, 3, 3 for the dominant aurora pattern, while one person said that it is "hard to tell" and one person saw no aurora in the video. When asked about the colour 30 people agreed it was green, and over half added other colours (yellow, violet blue and pink). Here four people suggested to have a longer video, especially that "one would like to look longer, as we don't have that here and it is very interesting". In **T8** all participants (33) correctly identified the section coordinates with the described features; 26 said that the veins are clearly visible, 7 claiming to the contrary. One participant suggested that a longer analysis would improve this task, one more expressed that they are not sure where the macula of the retina was, and another wished for an analysis of some other organ.

Summary Overall, the older adults from our study in most cases provided high quality contributions with no training. Only T2 and T7 proved to be somewhat challenging and with these participants asked for more data. Many wished for other tasks to be more challenging (harder font (T6), more audio variety (T4), more and better hidden animals (T1), longer analysis (T8)) and the "easy" dimension had the weakest correlations with willingness to do similar tasks in the future. It seems that older adults would not mind, and even preferred it, if the tasks posed more of a challenge (eg. T5 vs T6), especially if it would allow them to learn something interesting. They also wished for the shorter tasks to be extended, either by additional data (T1, T2) steps (T8) or longer duration (T7, T3, T4), not only because they enjoyed them and wanted to learn more, but also to allow them to provide higher quality contributions by adding more data to verify their choices. It seems therefore, that microtasks, designed to be brief for efficiency, could be extended and elaborated upon to increase the contributors' satisfaction, especially if they rely on image, video or audio data.

3.3 Evaluation of Tasks

Participants rated T8 the highest, while T2 the lowest. They distributed most points in the category "I would do similar tasks" (380) followed by "easy" (370), "attractive" (363), "engaging" (328) and "important" (283). Our participants, once exposed to each task, reported a high willingness to engage with similar tasks (with an average of 1.44 out of 2) suggesting, that older adults would engage with such tasks more, if they were made more easily available to them.

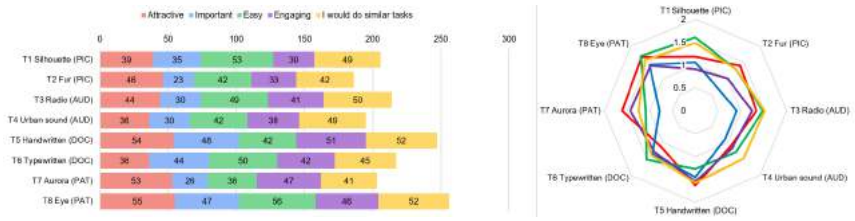


Fig. 2. Left: Total points awarded by our participants after the completion of each task. Right: Average scores for the same tasks.

As seen in Fig. 3 the correlations with the willingness to do similar tasks in the future are either positive, or close to zero, while the strongest correlation is with the visual or thematic "attractiveness" of the task. It was also slightly important whether the task was "engaging" or "important", especially if it was not found to be "attractive" and to a lesser extent if it was "easy". This suggests that older adults' main motivation is rather intrinsic, connected to their own interest in the task, which of course is moderated by other variables.

Correlation r <-1,1>; r = 0 no correlation
r < 0 negative correlation; r > 0 positive correlation

	I would do similar tasks and								
	Attractive		Important		Easy		Engaging		Interpretation of |r|
Correlation	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	
T1 Recognizing animal silhouettes (PIC)	0,5995	0,5653	0,5017	0,4315	0,2400	0,2712	0,3315	0,3424	> 0.9 – very strong colleration
T2 Recognizing cat fur types (PIC)	0,4282	0,4334	0,2965	0,2662	0,5141	0,4796	0,1354	0,1633	0.7 – 0.9 – strong correlation
T3 Recognizing messages on the radio (AUD)	0,6936	0,6091	0,3993	0,4078	0,2764	0,3475	0,2550	0,1958	0.4 – 0.7 – moderate correlation
T4 Recognizing local urban sounds (AUD)	0,2602	0,2351	0,4840	0,4699	0,0845	0,0969	0,6205	0,5937	0.2 – 0.4 – weak correlation
T5 Transcribing handwritten birth certificate (DOC)	**0,7965**	**0,8077**	0,0904	0,0680	0,2342	0,2651	0,2876	0,3369	< 0.2 – no linear relationship
T6 Transcribing a longer typewritten text (DOC)	0,3871	0,3253	0,2745	0,2485	0,2929	0,3476	0,1861	0,1832	
T7 Recognizing Aurora Borealis patterns (PAT)	0,4812	0,5438	0,2336	0,2386	0,2895	0,2967	0,3967	0,3686	
T8 Recognizing eye elements (PAT)	0,3433	0,3146	0,5323	0,4415	0,0675	0,1200	0,6632	0,6731	
Arithmetic mean	0,4987	0,4793	0,3515	0,3215	0,2499	0,2783	0,3595	0,3571	
Standard Deviation	0,1835	0,1886	0,1540	0,1391	0,1388	0,1242	0,1923	0,1891	
Variance	0,0337	0,0355	0,0237	0,0194	0,0193	0,0154	0,0370	0,0358	

Fig. 3. Correlation matrix for the dimension "I would do similar tasks" and other dimensions, from older adults' evaluation of the tasks right after performing them.

3.4 Motivation

After having completed all of the tasks, the participants chose learning something new and information about the purpose of performing these tasks as the prevailing motivators. They would also like to receive feedback on their performance and to have detailed tutorials. Detailed results are reported in Table 2. Moreover, the ability to perform these tasks using interfaces (smartphone, Smart TV or audio) other than a computer screen was judged as not particularly important, however, this may be due to lack of familiarity with them for audio and TV devices, as there are studies which successfully implement them for crowdsourcing [10, 27] and the challenge of small-screen interaction for smartphones [16] which is still relevant. [20]

	No. of P.	% of P.
The opportunity to learn something interesting while performing these tasks	24	72.7%
More knowledge about the purpose of performing these tasks	24	72.7%
Receiving feedback on the use and usefulness of the tasks performed	21	63.6%
A short training to make sure I do the tasks well	17	51.5%
More interesting topic of tasks	9	27.3%
Online support and contact with other people performing these tasks	9	27.3%
Tasks suited to my skills	9	27.3%
Training and personal meetings for those performing the tasks	7	21.2%
Statistics showing the number of already completed tasks	7	21.2%
The ability to perform these tasks on the TV screen with the remote control	6	18.2%
Thanks from the researchers	6	18.2%
Ability to perform these tasks on a smartphone	5	15.1%
Ability to perform these tasks using the voice interface	0	0.0%
"None of the above" and "Other (own answer)"	0	0.0%

Table 2. Answers to "Which of these elements would encourage you to perform tasks similar to the sample tasks in this survey?" asked after completing all of the tasks.

4 Conclusions

In this exploratory research we have verified that crowdsourcing microtasks, especially those appearing in citizen science projects, can be well-suited for some groups of older adults - both in terms of the quality of older adults' contributions and their motivation. Yet, even among older adults with average and higher ICT skills - sufficient to contribute to such projects, such as the participants in our sample, the awareness of the existence of such crowdsourcing projects is quite low, as such citizen science tasks are not easily found and sampled. Older adults as a group often overlooked as potential contributors to larger scale crowdsourcing projects due to their often lower willigness to engage online and the perception of their ICT skills. However, the older adults in our study who received no compensation, provided high quality contributions with little training and were open to continue volunteering online.

To increase participation, and thus the representation of this age group's voice in citizen science, we suggest that crowdsourcing tasks ought to be advertised in line with older adults' preferences. These are related to the way in which completing these tasks may benefit, first, them individually, and then, the society as a whole. Based on our research, crowdsourcing microtasks' presentation should focus on the aspect of learning something interesting (which was confirmed by an arithmetic mean correlation of 0.47 for "I would do similar tasks" and "Attractive, thematically or visually"), rather than the aspect of being able to utilize ones' existing skills and knowledge. The contributors should also be provided with a high awareness of the tasks' purpose and ought to be made aware of the usefulness of their individual contributions to reassure the participants that it was time well spent. The tasks could also be more elaborate, to provide an appropriate challenge and increase immersion. Hence, in future research we would also like to examine a wider range of tasks of increasing complexity and duration, as well as the effects of engaging in crowdsourcing on participant's physical, mental or cognitive well-being in further comparative longitudinal research with larger groups of participants of all ages.

References

1. Aula, A.: Learning to use computers at a later age. In: HCI and the Older Population, pp. 3–5. University of Glasgow, Leeds, UK (Sep 2004)
2. Barber, S.T.: The zooniverse is expanding: Crowdsourced solutions to the hidden collections problem and the rise of the revolutionary cataloging interface. Journal of Library Metadata **18**(2), 85–111 (2018). https://doi.org/10.1080/19386389.2018.1489449
3. Brewer, R., Morris, M.R., Piper, A.M.: Why would anybody do this?: Understanding older adults' motivations and challenges in crowd work. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 2246–2257. ACM (2016)
4. Campo, S.a., Khan, V.J., Papangelis, K., Markopoulos, P.: Community heuristics for user interface evaluation of crowdsourcing platforms. Future Generation Computer Systems **95**, 775 – 789 (2019)
5. Djoub, Z.: ICT education and motivating elderly people. In: Ariadna; cultura, educación y tecnología. vol. 1, pp. 88–92 (Jul 2013)
6. Population structure and ageing, https://ec.europa.eu/eurostat/statistics-explained/index.php
7. Greenfield, E.A., Marks, N.F.: Formal volunteering as a protective factor for older adults' psychological well-being. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **59**(5), S258–S264 (2004)
8. Hao, Y.: Productive activities and psychological well-being among older adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **63**(2), S64–S72 (2008)
9. von Helversen, B., Abramczuk, K., Kopeć, W., Nielek, R.: Influence of consumer reviews on online purchasing decisions in older and younger adults. Decision Support Systems (2018)
10. Hettiachchi, D., Sarsenbayeva, Z., Allison, F., van Berkel, N., Dingler, T., Marini, G., Kostakos, V., Goncalves, J.: "hi! i am the crowd tasker" crowdsourcing through digital voice assistants. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–14. CHI '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3313831.3376320
11. Häggström, M.: Medical gallery of mikael häggström 2014 (Jul 2014), https://en.wikiversity.org/
12. Itoko, T., Arita, S., Kobayashi, M., Takagi, H.: Involving senior workers in crowdsourced proofreading. In: International Conference on Universal Access in Human-Computer Interaction. pp. 106–117. Springer (2014)
13. Knowles, B., Hanson, V.L.: Older adults' deployment of 'distrust'. ACM Trans. Comput.-Hum. Interact. **25**(4), 21:1–21:25 (Aug 2018). https://doi.org/10.1145/3196490, http://doi.acm.org/10.1145/3196490
14. Knowles, B., Hanson, V.L.: The wisdom of older technology (non)users. Commun. ACM **61**(3), 72–77 (Feb 2018). https://doi.org/10.1145/3179995, http://doi.acm.org/10.1145/3179995
15. Kobayashi, M., Arita, S., Itoko, T., Saito, S., Takagi, H.: Motivating multi-generational crowd workers in social-purpose work. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. pp. 1813–1824. ACM (2015)
16. Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M., Ifukube, T.: Elderly user evaluation of mobile touchscreen interactions. In: Proceedings of the 13th

IFIP TC 13 International Conference on Human-Computer Interaction - Volume Part I. p. 83–99. INTERACT'11, Springer-Verlag, Berlin, Heidelberg (2011)
17. Kobayashi, M., Ishihara, T., Itoko, T., Takagi, H., Asakawa, C.: Age-based task specialization for crowdsourced proofreading. In: International Conference on Universal Access in Human-Computer Interaction. pp. 104–112. Springer (2013)
18. Kopeć, W., Skorupska, K., Jaskulska, A., Abramczuk, K., Nielek, R., Wierzbicki, A.: Livinglab pjait: Towards better urban participation of seniors. In: Proceedings of the International Conference on Web Intelligence. pp. 1085–1092. WI '17, ACM, New York, NY, USA (2017). https://doi.org/10.1145/3106426.3109040
19. Kötteritzsch, A., Koch, M., Lemân, F.: Adaptive training for older adults based on dynamic diagnosis of mild cognitive impairments and dementia. In: International Workshop on Ambient Assisted Living. pp. 364–368. Springer (2014)
20. Kowalski, J., Jaskulska, A., Skorupska, K., Abramczuk, K., Biele, C., Kopeć, W., Marasek, K.: Older adults and voice interaction: A pilot study with google home. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 187:1–187:6. CHI EA '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3290607.3312973
21. Lum, T.Y., Lightfoot, E.: The effects of volunteering on the physical and mental health of older people. Research on aging **27**(1), 31–55 (2005)
22. Morrow-Howell, N.: Volunteering in later life: Research frontiers. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **65**(4), 461–469 (2010)
23. Morrow-Howell, N., Hinterlong, J., Rozario, P.A., Tang, F.: Effects of volunteering on the well-being of older adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences **58**(3), S137–S145 (2003)
24. Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., Tomlinson, B.: Who are the crowdworkers? shifting demographics in mechanical turk. In: CHI '10 Extended Abstracts on Human Factors in Computing Systems. p. 2863–2872. CHI EA '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1753846.1753873
25. Seong, E., Kim, S.: Designing a crowdsourcing system for the elderly: A gamified approach to speech collection. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. p. 1–9. CHI EA '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3334480.3382999
26. Simpson, R., Page, K.R., De Roure, D.: Zooniverse: Observing the world's largest citizen science platform. In: Proceedings of the 23rd International Conference on World Wide Web. p. 1049–1054. WWW '14 Companion, Association for Computing Machinery, New York, NY, USA (2014). https://doi.org/10.1145/2567948.2579215
27. Skorupska, K., Núñez, M., Kopeć, W., Nielek, R.: Older adults and crowdsourcing: Android tv app for evaluating tedx subtitle quality. Proc. ACM Hum.-Comput. Interact. **2**(CSCW), 159:1–159:23 (Nov 2018). https://doi.org/10.1145/3274428
28. Skorupska, K., Núñez, M., Kopeć, W., Nielek, R.: A comparative study of younger and older adults' interaction with a crowdsourcing android tv app for detecting errors in tedx video subtitles. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (eds.) Human-Computer Interaction – INTERACT 2019. pp. 455–464. Springer International Publishing, Cham (2019)
29. Yu, H., Miao, C., Liu, S., Pan, Z., Khalid, N.S.B., Shen, Z., Leung, C.: Productive aging through intelligent personalized crowdsourcing. In: 30th AAAI Conference on Artificial Intelligence (AAAI-16) (2016)

5.5 AFFORCE: Actionable Framework for Designing Crowdsourcing Experiences for Older Adults (WI-IAT 2021, 70 pts)

Title	AFFORCE: Actionable Framework for Designing Crowdsourcing Experiences for Older Adults
Authors	Kinga Skorupska, Radosław Nielek, and Wiesław Kopeć
Conference	IEEE/WIC/ACM International Joint Conferences on Web Inteligence and Intelligent Agent Technology [WI-IAT 2021] (CORE B, 70 points on the list of the Polish Ministry of Science and Higher Education, as accessed on 5.04.2022)
Published in	IEEE/WIC/ACM International Conference on Web Intelligence (WIIAT '21), December 14–17, 2021, ESSENDON, VIC, Australia. ACM, New York, NY, USA.
Author Credit	I was the lead author of this article and my key contributions constituted the literature review and review of our own findings to date, analysis of barriers to crowdsourcing, analysis of dropout points, resulting framework design, accompanying conclusions, figure design, manuscript drafting and revision. I have also presented the paper at the conference (remotely).
Description	Based on our experience with the design of crowdsourcing systems for older adults in exploratory cases and studies, the context of related work, as well as research at the intersection of older adults' use of ICT, crowdsourcing and citizen science, we propose a framework for designing crowdsourcing systems that would be attractive and engaging for older adults called AFFORCE (Actionable Framework For Crowdsourcing Experiences). We first categorize and map mitigating factors and barriers to crowdsourcing for older adults to finally present, discuss and gather system elements addressing them into a comprehensive reference framework towards better inclusion of older adults in online crowdsourcing activities, including citizen science.
Contribution and Impact	This contribution aims to become a reference for crowdsourcing system designers and researchers aiming to develop engaging and motivating crowdsourcing systems, that would encourage older adults' contributions.

AFFORCE: Actionable Framework for Designing Crowdsourcing Experiences for Older Adults

Kinga Skorupska
kinga.skorupska@pja.edu.pl
Polish-Japanese Academy of Information Technology
Warsaw, Poland

Radosław Nielek
Polish-Japanese Academy of Information Technology
Warsaw, Poland

Wiesław Kopeć
Polish-Japanese Academy of Information Technology
Warsaw, Poland

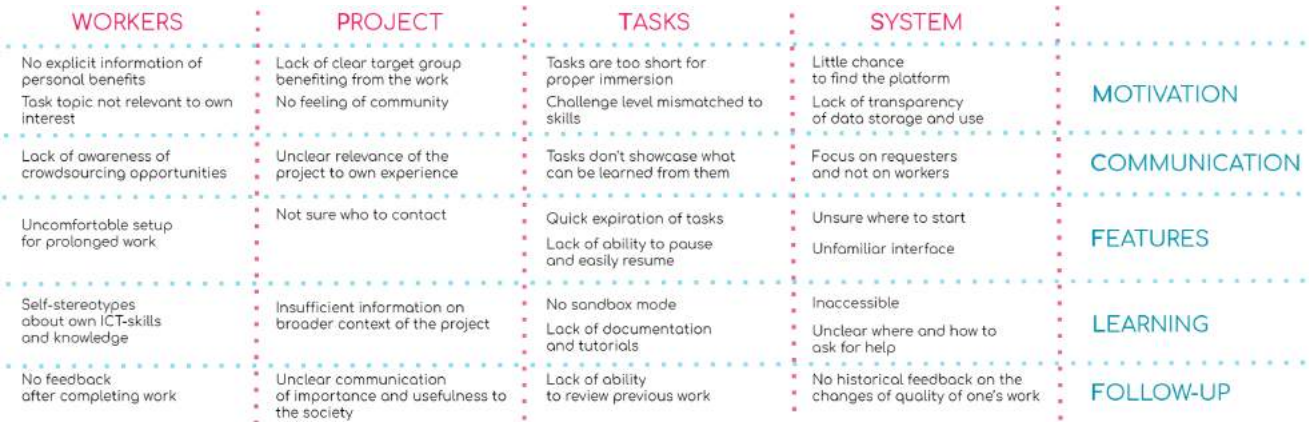


Figure 1: A summary of barriers to crowdsourcing which are especially discouraging to older adults.

ABSTRACT

In this article we propose a unique framework for designing attractive and engaging crowdsourcing systems for older adults, which is called AFFORCE (Actionable Framework For Crowdsourcing Experiences). We first categorize and map mitigating factors and barriers to crowdsourcing for older adults to finally discuss, present and combine system elements addressing them into an actionable reference framework. This innovative framework is based on our experience with the design of crowdsourcing systems for older adults in exploratory cases and studies, related work, as well as our and related research at the intersection of older adults' use of ICT, crowdsourcing and citizen science.

CCS CONCEPTS

• **Information systems** → **Crowdsourcing**; • **Social and professional topics** → *Seniors.*

KEYWORDS

crowdsourcing, older adults, user experience design, crowdsourcing framework, crowdsourcing system, crowdsourcing experience, active aging

ACM Reference Format:
Kinga Skorupska, Radosław Nielek, and Wiesław Kopeć. 2021. AFFORCE: Actionable Framework for Designing Crowdsourcing Experiences for Older Adults. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI-IAT '21), December 14–17, 2021, ESSENDON, VIC, Australia.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3486622.3494026

1 RATIONALE

Since in 2013 Kittur et al. [11] envisioned the future of crowdsourcing, technological advancements made crowdsourcing more efficient, cost-effective and streamlined. Yet, they lag in terms of humanizing the work outside of citizen science projects and efforts driven by volunteers and communities of practice or platforms catering to professional workers. Crowdsourcing platforms offering paid microtask crowd-work place little value in expendable workers, who in turn, may provide poor quality contributions, as the tasks are boring and repetitive [3], but the more they do, the better the pay. But even citizen science efforts could up their game to motivate and retain more groups of contributors driven by diverse intrinsic and extrinsic motivations. Thus, crowdsourcing systems are riddled with shortcomings which inhibit participation for all, but especially for older adults. These barriers, related e.g. to motivation, communication and accessible platform, project and task design are summarized in Fig 1 [3, 5, 8, 12, 14, 16, 19, 23, 29, 31, 32, 38].

WI-IAT '21, December 14–17, 2021, ESSENDON, VIC, Australia

ACM ISBN 978-1-4503-9115-3/21/12…$15.00
https://doi.org/10.1145/3486622.3494026

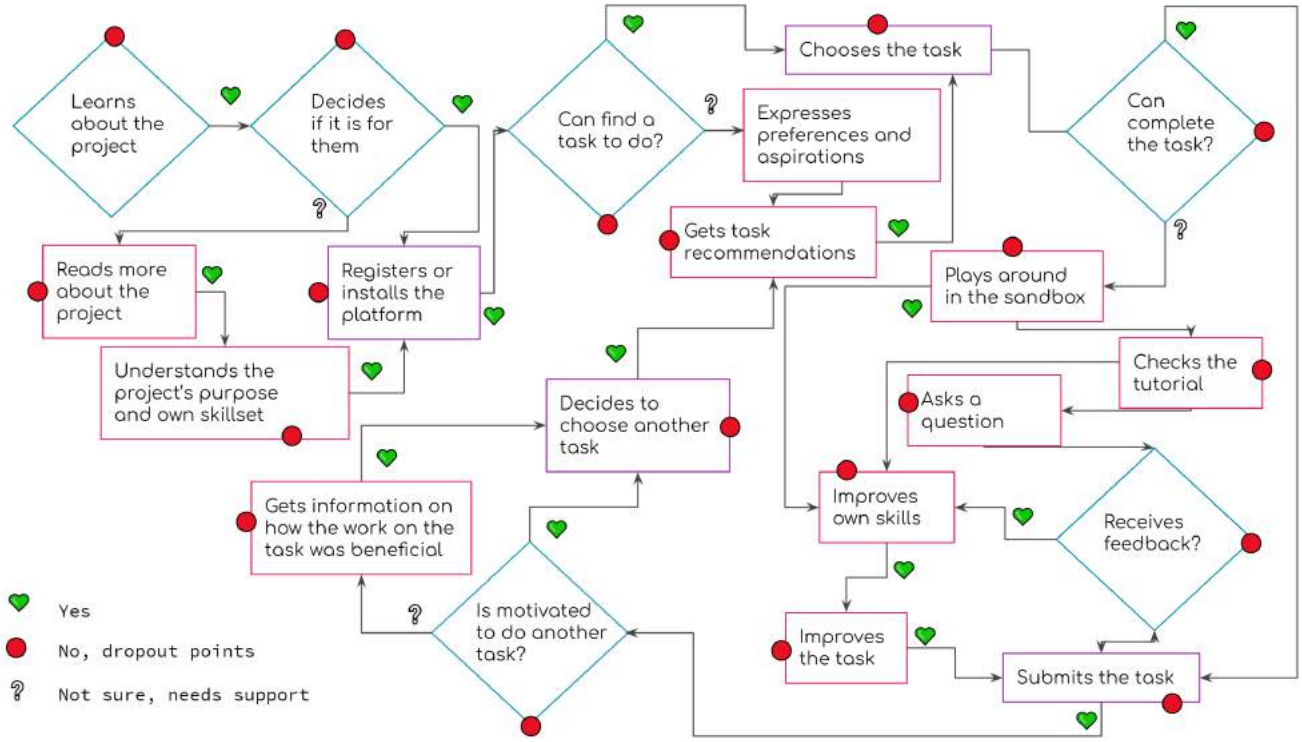


Figure 2: A diagram showing the user story on crowdsourcing platforms with indicated possible points of failure based on barriers shown in Fig 1. Potential support elements shown with pink borders.

As the share of older adults, here understood as people aged 65+, especially in western societies, is increasing it more and more important to design remote and scalable ICT-based activities, such as crowdsourcing, that provide them with cognitive challenges, enable positive interaction and active ageing [1, 4, 12, 20, 32]. Crowdsourcing projects, on the other hand can benefit from the massive, and largely untapped, potential of older adults' collective intelligence, experience and time [13, 20, 32]. Especially that they have proven to be dedicated, diligent and successful contributors [14, 20, 29, 30, 32] who are held back not only by inaccessible task-driven project design, which is incompatible with their prevailing motivations [11] and possible age-related cognitive decline in the processing speed and working memory [21] and the idea of worker expendability [3, 4, 19, 23] but also a lack of support mechanisms [4, 32], which all result in preventable dropouts, as shown in Fig 2.

Therefore, based on our experience with using, designing and developing crowdsourcing platforms, and a keyword-driven literature review done in a form of a mind-map, which allowed us to cluster relevant findings, in this article we envision the way forward to create more engaging, accessible and inclusive experience of crowdwork, encouraging contributions from underrepresented populations [31, 34]. Designing crowdsourcing experiences with older adults in mind, due to the curb-cut effect[1], will not only improve the experience for them, but also for the general population. To further this goal first we analysed the key barriers to crowdsourcing for older adults, as shown in Fig 1 and mapped them over the generalized crowd work user story depicted in Fig 2 to mark dropout points. Finally, we took Kittur et al.'s proposed framework [11] as the starting point for the construction of an actionable framework addressing these dropout hot-spots. Challenges are many, as visible in Fig 1, both related to the accessibility and the required ICT-proficiency [32] and to communication of projects, tasks and their importance. [3, 19] Crowdsourcing systems designed to encourage older adults' contributions ought to not only mitigate barriers but also highlight older adults strong suits: their lifelong experience, skills and patience[32]. They should meet their preferences regarding the experience of contributing, including their individual aspirations.[29]

The analysis and elements of the proposed framework are based on our experience with the design of crowdsourcing systems encouraging older adults' contributions in exploratory studies [17, 31–33], related work, as well as research at the intersection of older adults' use of ICT [2, 16, 25] and crowdsourcing [22, 24]. Our resulting framework aims to serve as an actionable reference for designing more sustainable and motivating crowdsourcing experiences that better engage older adults to boost the rates of their participation and help them reap the benefits and protective factors that come from online volunteering, lifelong learning, and staying active longer.

[1]The existence of the curb-cut effect has long been documented in the digital space [9] and it differs from universal design, as it first focuses on specific populations who could benefit most from a unique approach.

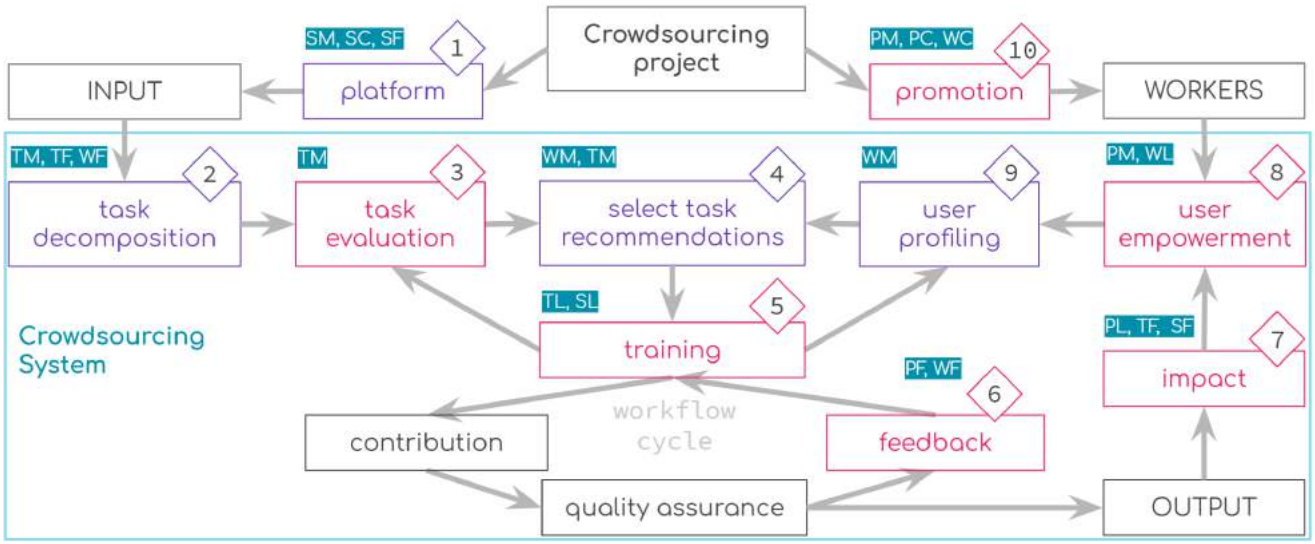


Figure 3: Crowdsourcing framework updated with elements important for older adults' engagement - elements in violet are changed, while in pink added in relation to the framework by Kittur et al. [11]. Moreover, the barriers addressed by each element are referred to by the first letters of titles of rows and columns in Fig. 1, so e.g. SM stands for System Motivation.

2 PROPOSED FRAMEWORK

What follows is the description of the proposed actionable framework addressing the analysed (Fig. 1) and mapped (Fig. 2) barriers. The discussion of framework elements is supported by examples from related research and practice. Special consideration is given to the elements important for designing experiences for older adults.

2.1 Platform

When starting a crowdsourcing project the choice of the **platform (1)** is the project-defining moment, as some of the features described in the framework may not be present in the platform and will need to be supplemented outside of the crowdsourcing system. This is a common problem with the **feedback (6)** and **training (5)** module, as well as **impact (7)** measures and follow-up information on the performance and importance of completed tasks, and quite a challenge, when it comes to **user empowerment (8)** – a necessary step when working with older adults using technology [16, 25]. The **platform (1)** may also be custom-built, and there was some success with crowdsourcing platforms designed with older adults in mind, especially with co-design [13, 16], which accounted for such aspects as familiarity with the interfaces [26], to help bridge the digital divide, and cater to the characteristic motivation and strengths of older adults [29, 32]. The importance of this point is clearly visible when evaluating platforms and experiences which were created incrementally and their technical and procedure-driven complexity increased overtime, such as Wikipedia [24]. In the end this complexity constitutes another barrier contributing to the inequality of contributions in terms of less ICT-privileged or experienced user groups which then are underrepresented [31].

Even platform-specific communication makes a big difference. For example, Amazon Mechanical Turk's worker-facing communication focuses on communicating benefits to businesses, calling itself a "crowdsourcing marketplace". On the other hand, the Prolific platform caters to both requesters and participants (contributors) equally and acknowledges the diverse motivations of the crowd with a call to action: "Take part in engaging research, earn cash, and help improve human knowledge" [2]. On the other hand, accessibility still plays a role with crowdsourcing platforms active on the market now, small fonts, cryptic pictograms, unfamiliar interfaces and complex procedures make many crowdsourcing platforms inaccessible by default to many groups of users [4, 23]. This problem, however, is not limited to crowdsourcing platforms or tasks and persists even in applications marketed at older adults [5]. To alleviate this problem crowdsourcing in a non-computer scenario may bring good results, such as with a remote on a TV screen [32], making use of familiar modes of interaction, as visible in Fig 4 or via a voice interface [10, 31].

2.2 Tasks

The input of the crowdsourcing project should be prepared so that it may be divided into microtasks. However, based on previous research such tasks are often too far removed to understand their purpose [3] and too short to be meaningful [11]. On the other hand, easily relatable tasks, such as tagging historical photos [36] or proofreading texts [15], are quite encouraging for older adults. Therefore, **task decomposition (2)** step should strive to produce recognizable work units that the contributors can engage with and become immersed, but can be paused if needed, even for a longer period of time [38] – perhaps self-specified by the contributor within a pre-set limit, so that users can take a break if they are tired [32] or feel sick [38] and have an overview of the broader context

[2]Quote and data retrieved on 12.04.2021 from two platforms, commonly used for conducting research: https://www.mturk.com/ and https://www.prolific.co/

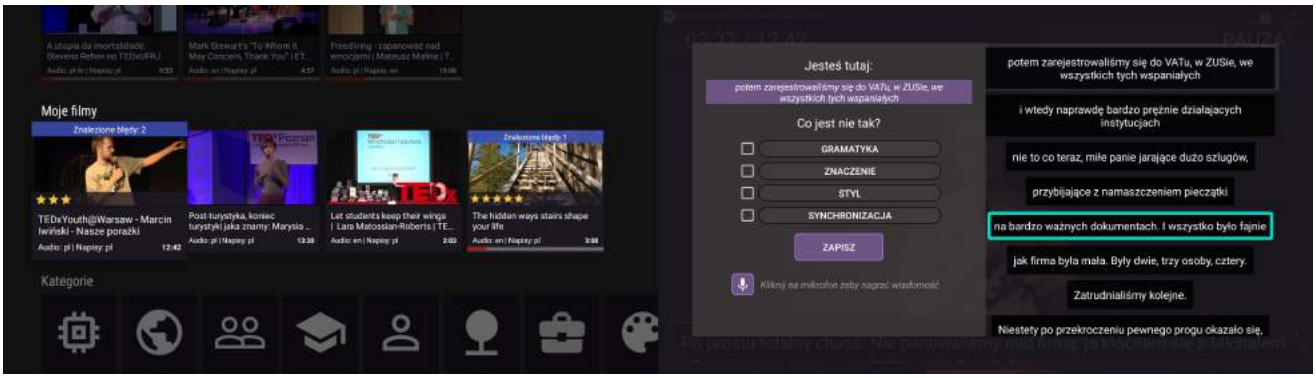


Figure 4: Screenshots from a crowdsourcing system designed for older adults and based on the idea of familiar interaction. It also allowed for pausing, viewing the context and free navigating through the task. Source: https://github.com/manununhez/dreamtv-app.

of their work upon return, so that they may again find their footing easily as shown in Fig 3 [32].

These **mezzo-tasks** should be between microtasks and macrotasks, short enough to complete in a self-defined session length of about 15-30 minutes, but long enough to offer a skill or a piece of knowledge that is valuable to the contributors. Especially older adults are often driven by intrinsic motivation of wanting to learn something interesting [11, 30]. For this reason **task evaluation (3)** is important, so that tasks can be categorized by the skills they utilize, the interest they belong to as well as the level of challenge they pose. Matching the interest, preferred skills and challenge level to the profile of the contributor is necessary for them to stay engaged and to support their development [35]. **Select task recommendations (4)** ought to be suggested to the contributors to choose from, so that they may exercise their autonomy [29, 32] and pick work based on their interests, knowledge, skills and energy levels that may change over time [12]. Very commonly tasks get abandoned [7] which results in effort and time being wasted, but also in discouragement of the contributors. Crowdsourcing platforms and projects get abandoned along with them. There are many reasons it can happen [4] including not receiving help, not finding an adequate project at all, or fast enough – but it can happen at almost any stage of the user journey, as shown in Fig 2. This is why **user profiling (9)** and a shortlist with **select task recommendations (4)** are so important. Once a user rejects a recommendation an adaptive system could suggest another one, with an increasingly better fit to ensure that they are able to find a task, and successfully complete it – increasing their well-being [31] and lowering dropout [4].

2.3 Empowerment

Older adults are often apprehensive towards technology and online services, especially if they had little prior contact with them [16]. In this context **user profiling (9)** ought to be preceded by **user empowerment (8)**, so that any self-stereotypes are caught before they affect the profiling. Held stereotypes may affect performance, beliefs about oneself as well as one's skills [13]. To address this danger of self-stereotyping and pave way for increased engagement the users ought to learn about crowdsourcing, its purpose and see representative examples. The empowerment step is recommended in many ICT-contexts where older adults are expected to contribute [16]. They should also be encouraged to trust their own life-course skills, experience and abilities [13] which benefit crowdsourcing projects, such as their high crystallized intelligence – equal to or higher than this of younger adults. Because crystallized intelligence encompasses aspects such as knowledge, understanding vocabulary and concepts and reasoning older adults may be perfect partners for human-in-the-loop systems, which could do the bulk of the crowdsourcing work, with **quality assurance** left to older adults in areas of their greatest strengths (e.g. language skills, life-course experience, sense-making, individual professional mastery). So, empowerment ought to:

(1) include ICT-empowerment by explaining the importance of crowdsourcing within the ICT practice,
(2) incorporate psycho-education to make older adults aware of their strengths which may benefit such projects,
(3) and at later stages involve **training (5)** to do the tasks better,
(4) as well as incorporate **feedback (6)** into the contribution cycle,
(5) allow to choose the type of contribution based on individual preferences, [12]
(6) and finally, inform the users of the **impact (7)** of their contribution, to showcase its societal importance [3].

2.4 Training

The aspect of **training (5)** and receiving **feedback (6)** is important to increase contributors' confidence and skills. Older adults may feel apprehensive of unfamiliar interfaces and new technologies so a sandbox mode, where they can practice without the fear of breaking anything is a requested feature [32]. Another need is related to staying mentally active through challenging oneself [32], while doing something of societal importance [3] and learning new things can be such challenge. So, **task evaluation (3)** should incorporate such tags as "knowledge gained" and "skills practiced" on ladder of different challenge levels, so that users may gain skills to become even better contributors and gain new roles [12]. The platform

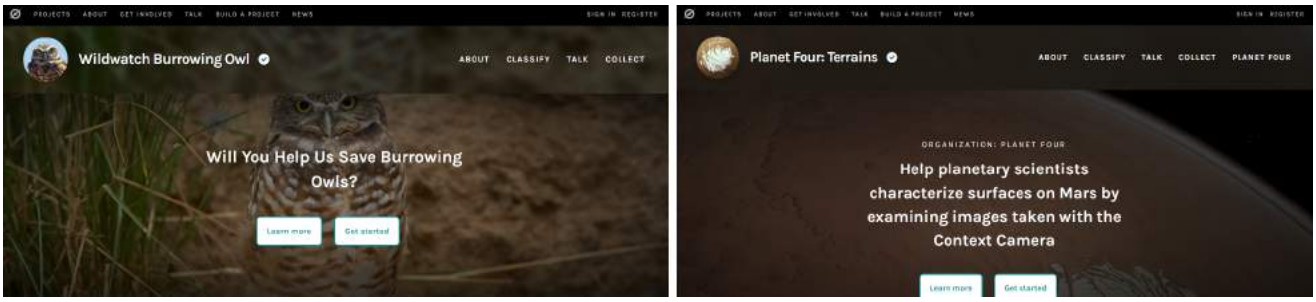


Figure 5: Two example calls to action from Zooniverse.org, a citizen science platform – both appealing to extrinsic motivation. Rephrasing them based on user-profiling, while still keeping true to the nature of the project, could bring better results.

ought to also support efforts such as community tutorial generation, either by providing a place to gather collective task insights [28] or fostering in-platform communication between users regarding specific tasks, or specific task parts [4], but in a positive learning environment, as not to discourage contributors [6]. Feedback and task performance should stay visible on the contributors' profile for their later reference [19]. The ability to grow one's skill through participation and self-correction is also important, as advancement path and the ability to choose one's role within the project may facilitate sustainable engagement [12]. Moreover, to increase engagement going beyond online space may be of value, especially to older adults who are isolated and may prefer face-to-face communication [8, 29]. Such offline supplementary activities can involve training, tutorial writing or content-creation marathons.

2.5 Profiling and Motivation

Older adults are a very heterogeneous group which spans decades of life [13] during which even successful aging may take very diverse forms in different people. As individual differences are prominent **user profiling (9)** has to catch diverse users' motivations, aspirations as well as abilities, skills and interests, to enable the system to adapt to its users. Profiling enables to **select task recommendations (4)** that would be appropriate for each individual, to still grant them freedom to choose their tasks, without causing overload. As small monetary rewards do not sufficiently motivate older adults [3, 29] focusing communication on other extrinsic and intrinsic motivators is important. Especially intrinsic motivation plays a greater role for older adults [11] so it is necessary to clearly establish aspirations, which especially in western societies most affected by the changing demographic situation, may include the need for entertainment, engaging with their interests or learning new things, and sometimes even staying mentally active to prevent cognitive decline [4, 29, 32]. Extrinsic motivation, such as the value and usefulness of the task also plays a role, especially if a specific beneficiary is clearly communicated [29]. So, the **impact (7)** that the project has in the real world, including the users' contribution should be made evident [31]. Also, as the energy levels, knowledge, skills, abilities and interests may change over time due to i.e. learning, or natural aging processes [12] the system should reassess the profile based on task performance, choices and engagement patterns [31], or even adapt to users' moods and emotions [37].

2.6 Discovery and Communication

Finally, older adults in general are not familiar with crowdsoucing [3], even if they have good ICT skills [30]. Here, **promotion (10)** of the crowdsourcing project is necessary in places that older adults already visit online, on platforms they are familiar with. A technology skill and familiarity barrier also exist, but other research suggests that older adults often teach each other tech skills [18] if they find them useful and engaging, so it may be enough to promote the project among a group of early adapters, who then take it further. Possibly, they could also recruit each other, as they know the places they frequent best [19]. In communication, it is good to underline such aspects as fun, entertainment or edutainment [32] and the fact that engaging in such activities may support "the increase of self-esteem and social engagement" [1] along with other benefits of active ageing and project-specific skill and interest highlights.

In general, communication efforts throughout the crowdsourcing platform ought to focus on the evident benefits for older adults themselves and then the society, or, better, a well-specified group within this society, such as children [29] or planetary researchers as visible in the message in Figure 5. Project communication should also be easy, open and two-way, as contributors quite often ask questions about the goals of a project, even in places not designed for this purpose, i.e. comment boxes [19]. Transparent communication is very important [3] and it should clearly uncover all study and work goals [19] and cover as much context of the work as possible, as limiting information on the project can negatively impact motivation [14]. In addition to communication related to recruitment and ongoing issues, such as current projects and open tasks, post-contribution communication is crucial, as older adults especially need to have a proof that their time was purposeful, meaningful and well-spent [27].

3 CONCLUSIONS

A crowdsourcing project aiming to be successful among older adults ought to be **well-promoted (10)**, discoverable, available on an accessible **platform (1)**, consisting of pausable, longer **self-contained mezzo-tasks (2)**, that are **evaluated and tagged (3)** based on the interest, skills involved, length and challenge levels. The users should be able to **select task recommendations (4)** provided based on **user profiling (9)** done after an **empowerment (8)** step that levels the playing field for contributors. The users

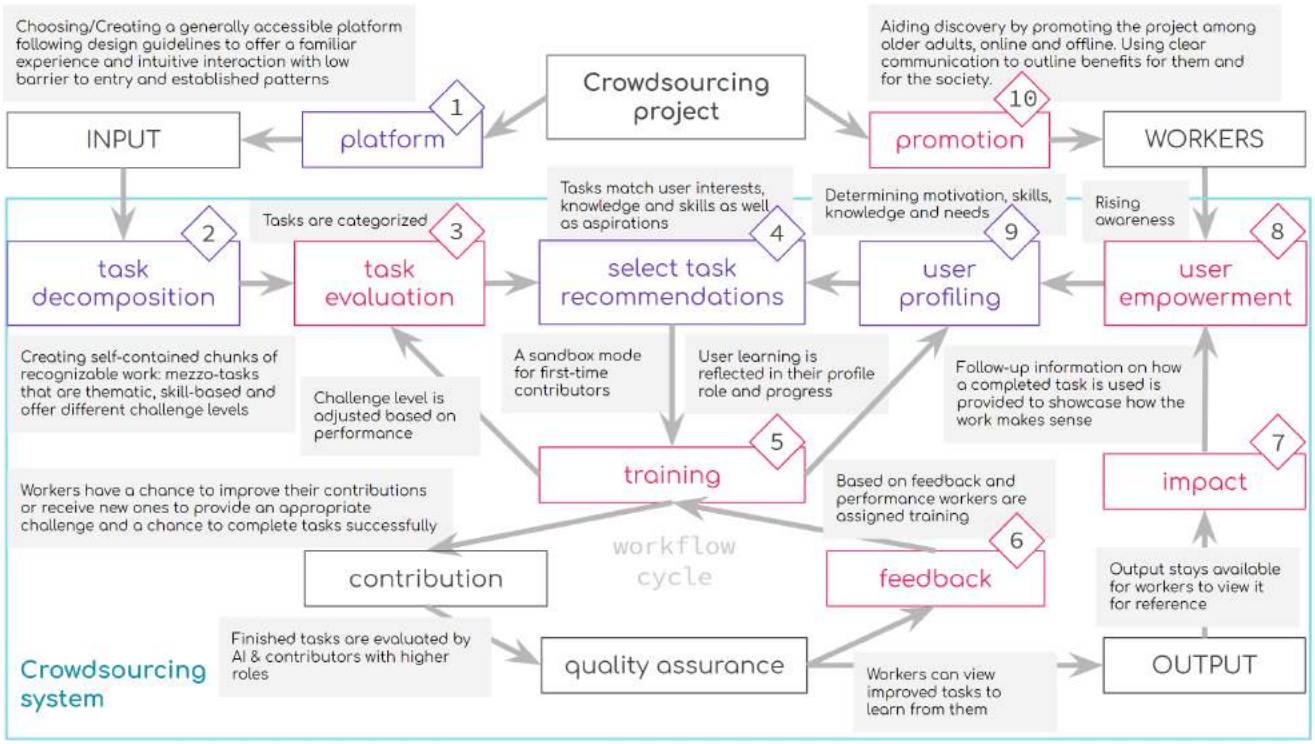


Figure 6: A summary of key assumptions of the framework discussed in Section 2.

should receive help in the form of tutorials, including **training (5)** in a sandbox mode. After completing a task, the users should get **feedback (6)** either from experienced users based on **quality assurance** or from automated systems, and a chance to **train (5)** the skills involved as well as information about the **impact (7)** of their contribution to fulfill both their need to learn, get better and progress in the ranks (intrinsic motivation) and contribute to Social Good (extrinsic motivation) as well as strengthen their motivation to continue contributing. The elements of the framework are summarized in Fig 6 together with key recommended functionalities.

Our hope is that this framework will be used as guidance for other researchers and practitioners delving into the area of creating crowdsourcing systems and experiences for older adults. Moreover, based on the future emerging studies we hope to develop the framework into a more comprehensive guide, towards better inclusion of older adults in ICT-mediated activities.

One limitation of the current crowdsourcing-focused framework is that there are no systems which incorporate all of the elements proposed – but it is also the motivation behind writing this paper. These proposed elements mitigate the analysed (Fig. 1) and mapped (Fig. 2) barriers making crowdsourcing more attractive to some groups of older adults, and due to the curb-cut effect, also to other groups of potential users. Therefore, they should be incorporated into crowdsourcing systems to cater not only to current devoted contributors or task requesters, but also the fringe of the user base – older adults who soon will be joining the ranks of contributors in greater numbers.

This research was partially supported by grants 2018/29/B/HS6/02604 and 2019/35/J/HS6/03166 from the National Science Centre of Poland.

REFERENCES

[1] Ana Maria Amorim and Vaninha Vieira. 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* (Vitória, Espírito Santo, Brazil) *(IHC '19)*. Association for Computing Machinery, New York, NY, USA, Article 65, 4 pages. https://doi.org/10.1145/3357155.3360478

[2] Bartłomiej Balcerzak, Wiesław Kopeć, Radosław Nielek, Sebastian Kruk, Kamil Warpechowski, Mateusz Wasik, and Marek Węgrzyn. 2017. Press F1 for help: participatory design for dealing with on-line and real life security of older adults. In *Computer Sciences and Information Technologies (CSIT), 2017 12th International Scientific and Technical Conference on*, Vol. 1. IEEE, IEEE, Lviv, Ukraine, 240–243.

[3] Robin Brewer, Meredith Ringel Morris, and Anne Marie Piper. 2016. "Why Would Anybody Do This?": Understanding Older Adults' Motivations and Challenges in Crowd Work. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 2246–2257. https://doi.org/10.1145/2858036.2858198

[4] Jennifer L. Davidson, Umme Ayda Mannan, Rithika Naik, Ishneet Dua, and Carlos Jensen. 2014. Older Adults and Free/Open Source Software: A Diary Study of First-Time Contributors. In *Proceedings of The International Symposium on Open Collaboration* (Berlin, Germany) *(OpenSym '14)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/2641580.2641589

[5] Amy Franklin and Sahiti Myneni. 2018. Engagement and Design Barriers of MHealth Applications for Older Adults. In *Proceedings of the Technology, Mind, and Society* (Washington, DC, USA) *(TechMindSociety '18)*. Association for Computing Machinery, New York, NY, USA, Article 9, 5 pages. https://doi.org/10.1145/3183654.3183695

[6] Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2013. The Rise and Decline of an Open Collaboration System: How Wikipedia's Reaction to Popularity Is Causing Its Decline. *American Behavioral Scientist* 57, 5 (2013), 664–688. https://doi.org/10.1177/0002764212469365 arXiv:https://doi.org/10.1177/0002764212469365

[7] Lei Han, Kevin Roitero, Ujwal Gadiraju, Cristina Sarasua, Alessandro Checco, Eddy Maddalena, and Gianluca Demartini. 2021. The Impact of Task Abandonment in Crowdsourcing. *IEEE Transactions on Knowledge and Data Engineering* 33, 5 (2021), 2266–2279. https://doi.org/10.1109/TKDE.2019.2948168

[8] Dave Harley, Kate Howland, Eric Harris, and Cara Redlich. 2014. Online Communities for Older Users: What Can We Learn from Local Community Interactions to Create Social Sites That Work for Older People. In *Proceedings of the 28th International BCS Human Computer Interaction Conference on HCI 2014 - Sand, Sea and Sky - Holiday HCI* (Southport, UK) *(BCS-HCI '14)*. BCS, Swindon, GBR, 42–51.

[9] B.W. Hesse. 1995. Curb cuts in the virtual community: telework and persons with disabilities. In *Proceedings of the Twenty-Eighth Annual Hawaii International Conference on System Sciences*, Vol. 4. IEEE, Wailea, HI, USA, 418–425 vol.4. https://doi.org/10.1109/HICSS.1995.375707

[10] Danula Hettiachchi, Zhanna Sarsenbayeva, Fraser Allison, Niels van Berkel, Tilman Dingler, Gabriele Marini, Vassilis Kostakos, and Jorge Goncalves. 2020. "Hi! I Am the Crowd Tasker" Crowdsourcing through Digital Voice Assistants. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376320

[11] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The Future of Crowd Work. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) *(CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 1301–1318. https://doi.org/10.1145/2441776.2441923

[12] Tiffany Knearem, Xiying Wang, and John M. Carroll. 2020. Sustaining Engagement in Volunteer Activities for Older Adults. In *Conference Companion Publication of the 2020 on Computer Supported Cooperative Work and Social Computing* (Virtual Event, USA) *(CSCW '20 Companion)*. Association for Computing Machinery, New York, NY, USA, 299–303. https://doi.org/10.1145/3406865.3418337

[13] Bran Knowles, Vicki L. Hanson, Yvonne Rogers, Anne Marie Piper, Jenny Waycott, Nigel Davies, Aloha Hufana Ambe, Robin N. Brewer, Debaleena Chattopadhyay, Marianne Dee, David Frohlich, Marisela Gutierrez-Lopez, Ben Jelen, Amanda Lazar, Radoslaw Nielek, Belén Barros Pena, Abi Roper, Mark Schlager, Britta Schulte, and Irene Ye Yuan. 2021. The Harm in Conflating Aging with Accessibility. *Commun. ACM* 64, 7 (June 2021), 66–71. https://doi.org/10.1145/3431280

[14] Masatomo Kobayashi, Shoma Arita, Toshinari Itoko, Shin Saito, and Hironobu Takagi. 2015. Motivating Multi-Generational Crowd Workers in Social-Purpose Work. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1813–1824. https://doi.org/10.1145/2675133.2675255

[15] Masatomo Kobayashi, Tatsuya Ishihara, Toshinari Itoko, Hironobu Takagi, and Chieko Asakawa. 2013. Age-Based Task Specialization for Crowdsourced Proofreading. In *Universal Access in Human-Computer Interaction. User and Context Diversity*, Constantine Stephanidis and Margherita Antona (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 104–112.

[16] Wiesław Kopeć, Radoslaw Nielek, and Adam Wierzbicki. 2018. Guidelines towards Better Participation of Older Adults in Software Development Processes Using a New SPIRAL Method and Participatory Approach. In *Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering* (Gothenburg, Sweden) *(CHASE '18)*. Association for Computing Machinery, New York, NY, USA, 49–56. https://doi.org/10.1145/3195836.3195840

[17] Wiesław Kopeć, Kinga Skorupska, Anna Jaskulska, Katarzyna Abramczuk, Radoslaw Nielek, and Adam Wierzbicki. 2017. LivingLab PJAIT: Towards Better Urban Participation of Seniors. In *Proceedings of the International Conference on Web Intelligence* (Leipzig, Germany) *(WI '17)*. ACM, New York, NY, USA, 1085–1092. https://doi.org/10.1145/3106426.3109040

[18] Jarosław Kowalski, Anna Jaskulska, Kinga Skorupska, Katarzyna Abramczuk, Cezary Biele, Wiesław Kopeć, and Krzysztof Marasek. 2019. Older Adults and Voice Interaction: A Pilot Study with Google Home. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3290607.3312973

[19] Qisheng Li, Krzysztof Z. Gajos, and Katharina Reinecke. 2018. Volunteer-Based Online Studies With Older Adults and People with Disabilities. In *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (Galway, Ireland) *(ASSETS '18)*. Association for Computing Machinery, New York, NY, USA, 229–241. https://doi.org/10.1145/3234695.3236360

[20] Ryan Moore and Jeffrey Hancock. 2020. Older Adults, Social Technologies, and the Coronavirus Pandemic: Challenges, Strengths, and Strategies for Support. *Social Media + Society* 6 (07 2020), 205630512094816. https://doi.org/10.1177/2056305120948162

[21] Daniel L Murman. 2015. The Impact of Age on Cognition. *Seminars in hearing* 36, 3 (Aug. 2015), 111–121. https://doi.org/10.1055/s-0035-1555115 Publisher: Thieme Medical Publishers.

[22] Radoslaw Nielek, Filip Georgiew, and Adam Wierzbicki. 2016. Crowd Teaches the Machine: Reducing Cost of Crowd-Based Training of Machine Classifiers. In *Artificial Intelligence and Soft Computing*, Leszek Rutkowski, Marcin Korytkowski, Rafał Scherer, Ryszard Tadeusiewicz, Lotfi A. Zadeh, and Jacek M. Zurada (Eds.). Springer International Publishing, Cham, 502–511.

[23] Radoslaw Nielek, Marta Lutostańska, Wiesław Kopeć, and Adam Wierzbicki. 2017. Turned 70? It is Time to Start Editing Wikipedia. In *Proceedings of the International Conference on Web Intelligence* (Leipzig, Germany) *(WI '17)*. Association for Computing Machinery, New York, NY, USA, 899–906. https://doi.org/10.1145/3106426.3106539

[24] Radoslaw Nielek, Marta Lutostańska, Wiesław Kopeć, and Adam Wierzbicki. 2017. Turned 70?: It is Time to Start Editing Wikipedia. In *Proceedings of the International Conference on Web Intelligence* (Leipzig, Germany) *(WI '17)*. ACM, New York, NY, USA, 899–906. https://doi.org/10.1145/3106426.3106539

[25] D. Orzeszek, W. Kopec, M. Wichrowski, R. Nielek, B. Balcerzak, G. Kowalik, and M. Puchalska-Kaminska. 2017. Beyond Participatory Design: Towards a Model for Teaching Seniors Application Design.

[26] Zhengxiang Pan. 2016. The Roles of Familiarity Design in Active Ageing. arXiv:1601.06869 [cs.HC]

[27] Pninit Russo-Netzer and Hadassah Littman-Ovadia. 2019. "Something to Live for": Experiences, Resources, and Personal Strengths in Late Adulthood. *Frontiers in Psychology* 10 (2019), 2452. https://doi.org/10.3389/fpsyg.2019.02452

[28] Shin Saito, Toshihiro Watanabe, Masatomo Kobayashi, and Hironobu Takagi. 2014. Skill Development Framework for Micro-Tasking. In *Universal Access in Human-Computer Interaction. Universal Access to Information and Knowledge*, Constantine Stephanidis and Margherita Antona (Eds.). Springer International Publishing, Cham, 400–409.

[29] Eunjin Seong and Seungjun Kim. 2020. Designing a Crowdsourcing System for the Elderly: A Gamified Approach to Speech Collection. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3382999

[30] Kinga Skorupska, Anna Jaskulska, Rafał Masłyk, Julia Paluch, Radosław Nielek, and Wiesław Kopeć. 2021. Older Adults' Motivation and Engagement with Diverse Crowdsourcing Citizen Science Tasks. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer International Publishing, Cham, 93–103.

[31] Kinga Skorupska, Kamil, Warpechowski, Radosław Nielek, and Wiesław Kopeć. 2020. Conversational Crowdsourcing for Older Adults: a Wikipedia Chatbot Concept.

[32] Kinga Skorupska, Manuel Núñez, Wiesław Kopeć, and Radoslaw Nielek. 2018. Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 159 (Nov. 2018), 23 pages. https://doi.org/10.1145/3274428

[33] Kinga Skorupska, Manuel Núñez, Wiesław Kopeć, and Radosław Nielek. 2019. A Comparative Study of Younger and Older Adults' Interaction with a Crowdsourcing Android TV App for Detecting Errors in TEDx Video Subtitles. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernando Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis Zaphiris (Eds.). Springer International Publishing, Cham, 455–464.

[34] Nitin Verma, Kristina Shiroma, Kate Rich, Kenneth R. Fleischmann, Bo Xie, and Min Kyung Lee. 2021. Conducting Quantitative Research with Hard-To-Reach-Online Populations: Using Prime Panels to Rapidly Survey Older Adults During a Pandemic. , 384–393 pages.

[35] Xiying Wang, Tiffany Knearem, and John M. Carroll. 2018. Ordinary Innovations of Older Adults: An Interview Study to Understand Interest Development in Aging. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing* (Jersey City, NJ, USA) *(CSCW '18)*. Association for Computing Machinery, New York, NY, USA, 253–256. https://doi.org/10.1145/3272973.3274068

[36] Han Yu, Chunyan Miao, Siyuan Liu, Zhengxiang Pan, N. Syahidah B. Khalid, Zhiqi Shen, and Cyril Leung. 2016. Productive Aging through Intelligent Personalized Crowdsourcing. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence* (Phoenix, Arizona) *(AAAI'16)*. AAAI Press, Phoenix, Arizona, 4405–4406.

[37] Han Yu, Zhiqi Shen, Simon Fauvel, and Lizhen Cui. 2017. Efficient scheduling in crowdsourcing based on workers' mood. In *2017 IEEE International Conference on Agents (ICA)*. IEEE, Beijing, China, 121–126. https://doi.org/10.1109/AGENTS.2017.8015317

[38] Kathryn Zyskowski, Meredith Ringel Morris, Jeffrey P. Bigham, Mary L. Gray, and Shaun K. Kane. 2015. Accessible Crowdwork? Understanding the Value in and Challenge of Microtask Employment for People with Disabilities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing* (Vancouver, BC, Canada) *(CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 1682–1693. https://doi.org/10.1145/2675133.2675158

6 Scientific Profile of the Candidate

Kinga H. Skorupska

		h-index	citations
Google Scholar	https://scholar.google.fi/citations?user=zEugAPUAAAAJ&hl	7	184
Research Gate	https://www.researchgate.net/profile/Kinga-Skorupska	6	161
Scopus	https://www.scopus.com/authid/detail.uri?authorId=57196017554	4	80 (69)

accessed on 18.05.2022

Selected Research & Publications

B	**Kinga Skorupska**, Radosław Nielek, Wiesław Kopeć: ***AFFORCE: Actionable Framework for Designing Crowdsourcing Experiences for Older Adults***. IEEE/WIC/ACM International Conference on Web Intelligence (WIIAT '21) (Dec 2021). https://doi.org/10.1145/3486622.3494026	2021
A	**Kinga Skorupska,** Anna Jaskulska, Rafał Masłyk, Julia Paluch, Radosław Nielek, Wiesław Kopeć (2021) ***Older Adults' Motivation and Engagement with Diverse Crowdsourcing Citizen Science Tasks***. In: Ardito C. et al. (eds) Human-Computer Interaction – INTERACT 2021. INTERACT 2021. Lecture Notes in Computer Science, vol 12933. Springer, Cham. https://doi.org/10.1007/978-3-030-85616-8	
A*	Wiesław Kopeć, Krzysztof Kalinowski, Monika Kornacka, **Kinga H. Skorupska**, Julia Paluch, Anna Jaskulska, Grzegorz Pochwatko, Jakub Filip Mozaryn, Paweł Kobyliński, and Piotr Gago. 2021. ***VR Hackathon with Goethe Institute: Lessons Learned from Organizing a Transdisciplinary VR Hackathon***. In CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '21 Extended Abstracts), May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3411763.3443432	
A*	Wiesław Kopeć, Jarosław Kowalski, Julia Paluch, Anna Jaskulska, **Kinga H. Skorupska**, Marcin Niewiński, Maciej Krzywicki, and Cezary Biele. 2021. ***Older Adults and Brain-Computer Interface: An Exploratory Study***. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21). Association for Computing Machinery, New York, NY, USA, Article 465, 1–5. DOI: https://doi.org/10.1145/3411763.3451663	
B	**Kinga Skorupska**, Kamil Warpechowski, Radosław Nielek, Wiesław Kopeć (2020): ***Conversational Crowdsourcing for Older Adults: a Wikipedia Chatbot Concept***. In: Proceedings of the 18th European Conference on Computer-Supported Cooperative Work: The International Venue on Practice-centred Computing on the Design of Cooperation Technologies - Exploratory Papers, Reports of the European Society for Socially Embedded Technologies (ISSN 2510-2591), DOI: 10.18420/ecscw2020_ep05	2020
	Kinga Skorupska, Daniel Cnotkowski, Julia Paluch, Rafał Masłyk, Anna Jaskulska, Monika Kornacka, and Wiesław Kopeć: ***All Factors Should Matter! Reference Checklist for Describing Research Conditions in Pursuit of Comparable IVR Experiments*** - Proceedings of the Mulitimedia, Interaction, Design and Innnovation, MIDI 2020, Warsaw, Poland, December 9-10, 2020	
	Anna Jaskulska, **Kinga Skorupska**, Barbara Karpowicz, Jarosław Kowalski, Cezary Biele, and Wiesław Kopeć: ***Exploration of Voice User Interfaces for Older Adults - A Pilot Study to Address Progressive Vision Loss*** - Proceedings of the Mulitimedia, Interaction, Design and Innnovation, MIDI 2020, Warsaw, Poland, December 9-10, 2020	
	Rafal Maslyk, **Kinga Skorupska**, Piotr Gago, Marcin Niewinski, Barbara Karpowicz, Anna Jaskulska, Katarzyna Abramczuk, and Wieslaw Kopeć: ***Deploying Crowdsourcing for Workflow Driven Business Process - a Brief Proposal*** - Proceedings of the Mulitimedia, Interaction, Design and Innnovation, MIDI 2020, Warsaw, Poland, December 9-10, 2020	
A	**Kinga Skorupska**, Manuel Nunez, Wieslaw Kopec, and Radoslaw Nielek. 2019. ***A Comparative Study of Younger and Older Adults' Interaction with a Crowdsourcing Android TV App for Detecting Errors in TEDx Video Subtitles***. The 17th IFIP TC.13 International Conference on Human-Computer Interaction (Interact 2019)	2019
A*	Jarosław Kowalski, Anna Jaskulska**, Kinga Skorupska,** Katarzyna Abramczuk, Cezary Biele, Wiesław Kopeć, and Krzysztof Marasek. 2019. ***Older Adults and Voice Interaction: A Pilot Study with Google Home***. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19). ACM, New York, NY, USA, Paper LBW0187, 6 pages.	
	Cezary Biele., Anna Jaskulska, Wiesław Kopeć, Jarosław Kowalski, **Kinga Skorupska**, Aldona Zdrodowska. ***How Might Voice Assistants Raise Our Children?***. In: Karwowski W., Ahram T. (eds) Intelligent Human Systems Integration 2019. IHSI 2019. Advances in Intelligent Systems and Computing, vol 903. Springer	
	Wieslaw Kopec, Marcin Wichrowski, Krzysztof Kalinowski, Anna Jaskulska, **Kinga Skorupska**, Krzysztof Marasek et al.*:* ***VR with Older Adults: Participatory Design of a Virtual ATM Training Simulation***. **IFAC HMS '19**: Proceedings of the 14th IFAC Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, Tallinn, Estonia, IFAC-PapersOnLine Volume 52, Issue 19, 2019, Pages 277-281, Elsevier, UK	
A	**Kinga Skorupska**, Manuel Nunez, Wieslaw Kopec, and Radoslaw Nielek. 2018. ***Older Adults and Crowdsourcing: Android TV App for Evaluating TEDx Subtitle Quality***. Proc. ACM Hum.-Comput. Interact. 2, CSCW, Article 159 (November 2018), 23 pages.	2018 & earlier
B	Wiesław Kopeć, **Kinga Skorupska**, Anna Jaskulska, Katarzyna Abramczuk, Radoslaw Nielek, and Adam Wierzbicki. 2017. ***LivingLab PJAIT: Towards Better Urban Participation of Seniors***. WI '17, Leipzig, Germany, August 23-26, 2017, 8 pages.	

Research Goals

In the area of ICT:	Evaluate the potential of novel interfaces such as Voice Assistants or Virtual and Augmented Reality for collaborative and social activities such as work, education, or crowdsourcing in the context of access to ICT-based solutions for users often excluded from the main technological discourse.
In social sciences:	Postulate best practices of designing ICT solutions used for social good (i.e. social inclusion; access to knowledge; lifelong learning, building Social Capital, pro-social attitudes) and strategies for lowering participation barriers and increasing user engagement.

Academic & Research Experience

Researcher & Lecturer at PJAIT – XR Lab, Multimedia Department	» does participatory research in the area of software and application design » plans, conducts and communicates exploratory studies in HCI » follows state of the art on novel interfaces' (VA, XR, BCI) design and verification, participatory design practices, ICT for Ageing Population & Education » facilitates research, educational and collaborative activities in the area of HCI » conducts classes with students of Computer Science, Interior Design and Information Management	**2019** **Present**
Researcher & Methodologist at PJAIT - The ALIEN Project	» researched experiential, participatory and inquiry-based methods » evaluated and advised on training outcomes » co-designed & described a social design method of collaboration between design and social science students facilitating migration studies in an international team » facilitated the creation of e-learning courses for the NOMAD platform » conducted classes with students of New Media Arts and Information Management	**2016** **2020**

Selected Formal Education

2018 **Present**	**Polish-Japanese Academy of Information Technology & SWPS University of Social Sciences and Humanities,** Computer Science, ICT & Psychology, **PhD Candidate,** Thesis on designing crowdsourcing systems for older adults
2004 **2010**	**University of Warsaw,** Neophilology**, Master's Degree,** Thesis on otherness and humanity in the context of social and technological changes

Other Interests

Gamification **Science & Arts** **Other interests**	» MUDs, board games, game mechanics, scenario design for pen and paper RPGs » podcasts, popular science, creative writing, science-fiction, lifelong learning » viticulture and winemaking, in-line skating, table tennis, curling