

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.

Copyright 2020 held by Authors, DOI: 10.18420/ecscw2020_ep05

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, contact the Authors.

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 guidelines¹ 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.

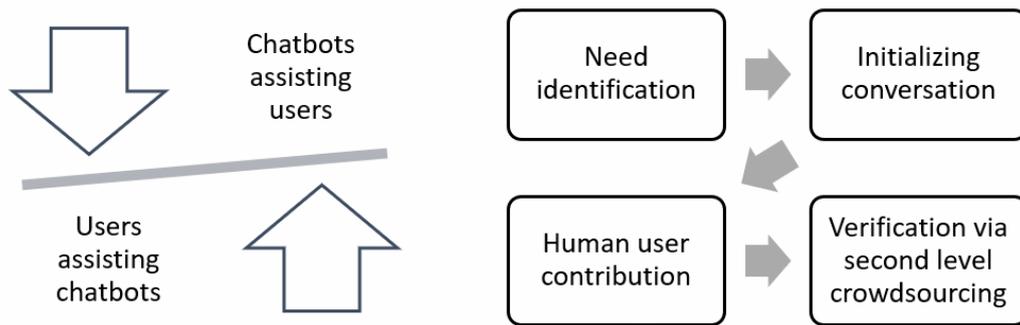


Figure 1. Inversion of the human-chatbot interaction paradigm and proposed chatbot-streamlined contribution process.

(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. Wikipedia³ 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 breeds⁴, 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	
			
Golden retriever		Origin Scotland, United Kingdom	
Inne nazwy	Yellow retriever, Russian Retriever	Traits [hide]	
Kraj patronacki	Wielka Brytania	Weight	Male 65–75 lb (29–34 kg) ^[1] Female 55–65 lb (25–29 kg) ^[1]
Kraj pochodzenia	USA^[1]	Height	Male 22–24 in (56–61 cm) ^[2] Female 20-22 in (51-56 cm) ^[2]
Wymiary		Coat	straight or moderately wavy
Wysokość	56 - 61 cm (psy), 51 - 56 cm (suki) ^[2]	Colour	any shades of gold or cream
Masa	29 - 45 kg (psy), 24 - 29 kg (suki)	Life span	average 10-12 years
Klasyfikacja		Classification / standards [hide]	
FCI	Grupa VIII, Sekcja 1, nr wzorca 111	FCI	Group 8, Section 1 standard ↗ Retrievers #111
AKC	Sporting	AKC	Sporting standard ↗
ANKC	Grupa 3 - Gundogs	ANKC	Group 3 (Gun dogs) standard ↗
CKC	Grupa 1 - Sporting Dogs	CKC	Group 1 – Sporting standard ↗ dogs
KC(UK)	Gundog	KC (UK)	Sporting dog standard ↗
NZKC	Gundog	UKC	Sporting and fishing
UKC	Grupa 4 - Gun Dog	Domestic dog (<i>Canis lupus familiaris</i>)	
FCI ↗ • AKC ↗ • ANKC ↗ • CKC ↗ • KC(UK) ↗ • NZKC ↗ • UKC ↗			

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

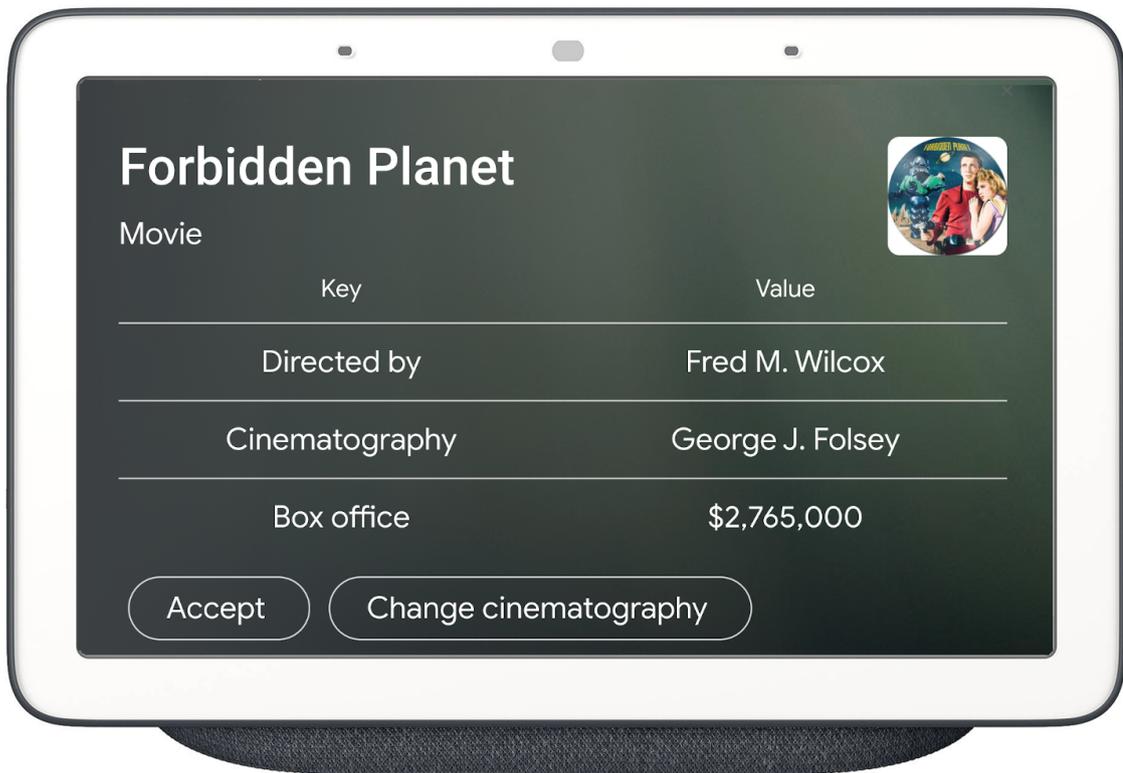


Figure 3. Gizmo on Google Home Hub. Sample screen showing the process of infobox editing. DialogFlow has predefined styles for tables so the table design is visually different from the standard infobox on Wikipedia.

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 getting 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 heterogeneous 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 find 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.
- Kopec, 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.
- Kopec, 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. Kopec, 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, L., 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.