5-2-2023

FIGHTING FALSE INFORMATION - DESIGNING A CONVERSATIONAL AGENT FOR PUBLIC SECTOR ORGANIZATIONS

Alexander Kocur
University of Potsdam, alexander.kocur@uni-potsdam.de

Sünje Clausen
University of Potsdam, suenje.clausen@uni-potsdam.de

Lennart Hofeditz
University of Potsdam, lennart.hofeditz@uni-potsdam.de

Felix Brünker
University of Potsdam, felix.bruenker@uni-potsdam.de

Jennifer Fromm
University of Duisburg-Essen, jennifer.fromm@uni-due.de

See next page for additional authors

FIGHTING FALSE INFORMATION - DESIGNING A CONVERSATIONAL AGENT FOR PUBLIC SECTOR ORGANIZATIONS

Research in Progress

Alexander, Kocur, University of Potsdam, Germany, alexander.kocur@uni-potsdam.de

Sünje, Clausen, University of Potsdam, Germany, suenje.clausen@uni-potsdam.de

Lennart, Hofeditz, University of Potsdam, Germany, lennart.hofeditz@uni-potsdam.de

Felix, Brünker, University of Potsdam, Germany, felix.bruenker@uni-potsdam.de

Jennifer, Fromm, University of Duisburg-Essen, Germany, jennifer.fromm@uni-due.de

Stefan, Stieglitz, University of Potsdam, Germany, stefan.stieglitz@uni-potsdam.de

Abstract

The digital transformation poses challenges for public sector organizations (PSOs) such as the dissemination of false information in social media which can cause uncertainty among citizens and decrease trust in the public sector. Some PSOs already successfully deploy conversational agents (CAs) to communicate with citizens and support digital service delivery. In this paper, we used design science research (DSR) to examine how CAs could be designed to assist PSOs in fighting false information online. We conducted a workshop with the municipality of Kristiansand, Norway to define objectives that a CA would have to meet for addressing the identified false information challenges. A prototypical CA was developed and evaluated in two iterations with the municipality and students from Norway. This research-in-progress paper presents findings and next steps of the DSR process. This research contributes to advancing the digital transformation of the public sector in combating false information problems.

Keywords: False Information, Conversational Agents, Crisis Communication, Media Literacy.

1 Introduction

Individuals are confronted with an abundance of information online. The rise of social media has been accompanied by an increased circulation of false information, especially in times of crises (Ceron, de-Lima-Santos and Quiles, 2021). False information can harm society and the common good as the (unintended) spread of false information can, for example, mislead large populations (Meel and Vishwakarma, 2020) and threaten public health (Naeem, Bhatti and Khan, 2021).

As PSOs should serve the society and either create or contribute to public value or the common good (Jørgensen and Bozeman, 2007; Moore, 1995), they are also concerned with addressing the problem of false information. However, PSOs often lack the resources to handle 1:1 communication with the public (Hofeditz et al., 2022). Here, information technology can extend the range, availability, and quality of services for citizens and thus, enable PSOs to maximize their contribution to public value (Bannister and Connolly, 2014; Rose et al., 2015).

For instance, CAs are increasingly being used by governments to handle large volumes of citizen contact (Henman, 2020). CAs are interactive systems which can communicate with human users by using and processing natural language (McTear, Callejas and Griol, 2016). The public sector can benefit from

using CAs in the form of more flexible services, improved service levels, improved sense of safety and trust in government (Tendai et al., 2020). For example, many Norwegian municipalities already use CAs as an additional channel on their websites for providing digital public services to their citizens such as answering frequently asked questions[1].The lessons learned from a recent evaluation of the use of CAs in the Norwegian government have shown that CAs can contribute to improve interactions with citizens and to support service agents in their tasks, and thus have significant potential for further advancement in the direction of hybrid service teams (Vassilakopoulou et al., 2023).

In the context of false information, CAs have been shown to have the potential to amplify and distribute trustworthy information in times of uncertainty (Stieglitz et al., 2022) and have been evaluated as a useful tool for reducing the spread of false information by, for example, providing accurate information (Roque et al., 2021). Therefore, we suggest that CAs might also be a promising tool for supporting PSOs in combating false information online. Accordingly, we formulate the following research question (RQ):
RQ: How can a CA be designed that supports PSOs in combating false information online?

To answer this research question, we applied a DSR approach based on Peffers et al. (2007). We worked together with the municipality of Kristiansand in Norway for designing and evaluating the CA and for evaluating learnings and risks associated with using the CA for this specific purpose. Since the municipality already provides a CA for digital public services and faces the challenges of dealing with false information online, this public organization represents an adequate use case for answering the proposed research question.

We conducted a workshop with communications experts, crisis managers, and IT experts from the municipality of Kristiansand to specify the problems related to false information and to derive objectives for the CA. Based on the objectives, we developed a prototypical CA as an artifact and evaluated it in two iterations with the experts from the municipality and students from Norway. This research-in-progress paper presents the current progress and the next steps of the applied DSR process.

2 Background

2.1 False Information in Social Media Communication

In the current digital era, the access to information has gradually changed due to new digital technologies and the social connectivity of individuals on the internet. Accordingly, individuals encounter huge flows of information, created not only by traditional means but also by users which are able to spread any kind of information on social platforms (Ceron, de-Lima-Santos and Quiles, 2021). The increase in active users and user-generated content has amplified the problems related to false information. In information systems (IS) research, the terms misinformation, disinformation, fake news and rumors are often used interchangeably. While some researchers distinguish the terms based on the dimensions of falseness, harm, and the intention to deceive (DiFonzo and Bordia, 2007; Allcott and Gentzkow, 2017; Tandoc, Lim and Ling, 2018), this study uses the umbrella term *false information* to include all of the previously mentioned concepts.

False information online poses threats to the common good, as it can cause economic damages, emotional suffering, and loss of trust (Tran et al., 2019), mislead large populations (Meel and Vishwakarma, 2020), threaten public health and pose a danger to life (Naeem, Bhatti and Khan, 2021). Particularly during crisis events, such as the COVID-19 pandemic, false information has gained a lot of impact, as social media is increasingly used by individuals to seek information during crisis events in order to create an understanding of the situation, which is referred to as sense-making (Stieglitz et al., 2017). False information has been previously identified as a serious challenge that can create a barrier for the public's sense-making process on social media (Allcott and Gentzkow, 2017; Stieglitz et al., 2017). Consequently, there is growing literature that addresses the diffusion and impact of false information in order to derive implications on how to combat this issue from different perspectives such

[1] https://www.kristiansand.kommune.no/ (visited on November 17, 2022)

as governmental corrective actions like debunking and deletion of false information or proactive approaches to prevent the dissemination of false information (Dumitru, Ivan and Loos, 2022). While the advent of information and communication technologies facilitates the creation and dissemination of false information, new communication technologies, such as CAs, can also be investigated to combat this problem, as they offer the possibility of reaching a wide mass of people by being publicly and permanently available in a digital context.

2.2 Conversational Agents

In IS research, CAs are interactive systems such as chatbots (Duan, Edwards and Dwivedi, 2019) or virtual assistants (Mirbabaie et al., 2021) which are capable of interacting in natural language. By implementing social cues such as small talk, personal greetings, or asking to start or pursue the dialog, CAs can enhance the user experience (Feine, Morana and Maedche, 2019; Diederich, Brendel and Kolbe, 2020). Thus, CAs transmit their intent to humans through both text and social cues (Brachten, Kissmer and Stieglitz, 2021; Mirbabaie et al., 2021). Therefore, IS research suggest to follow distinct design principles for the development of CAs to tailor them to the specific use case (e.g. Feine, Morana and Maedche, 2019; Diederich, Brendel and Kolbe, 2020; Tavanapour, Poser and Bittner, 2020).

Past comprehensive IS research on CAs often focused on collaborative work (Brachten et al., 2020). However, possible applications for CAs in the context of the health communication of PSOs, such as governments during the COVID-19 pandemic, have also been examined (Hofeditz et al., 2022). CAs have the potential to amplify and distribute trustworthy information in times of uncertainty. Thus, to fight the spread of false information in social media communication, PSOs can use CAs to inform the public and improve their responses to uncertain situations (Stieglitz et al., 2022) as well as improve the flexibility, reach, service level and trust in government (Tendai et al., 2020). In particular during crisis situations, the degree of uncertainty increases by the complex and diverse information that are spreading on social media (Mirbabaie, Stieglitz and Brünker, 2022). Thus, CAs can be used to relentlessly inform the public about reliable information (Hofeditz et al., 2019). In situations with high uncertainty such as a crisis, CAs may reduce upcoming contradictory messages by providing a) real-time information about the current state of the crisis, and b) educate the public about the detection of false information on social media (Shu et al., 2020). Based on the literature, we aim to use this knowledge on CAs for PSOs communication to examine how these systems can help the public sector to combat false information to address challenges such as economic damages, emotional suffering, and loss of trust (Tran et al., 2019).

3 Research Design

The aim of this research is to develop a useful, applicable artifact that addresses the practical problems of public service organizations in the context of false information online. Accordingly, it is situated in the "design science as a research methodology" genre (Peffers et al., 2007; Peffers, Tuunanen and Niehaves, 2018) and follows the design science research methodology proposed by Peffers et al. (2007). Figure 1 delineates the applied research process.

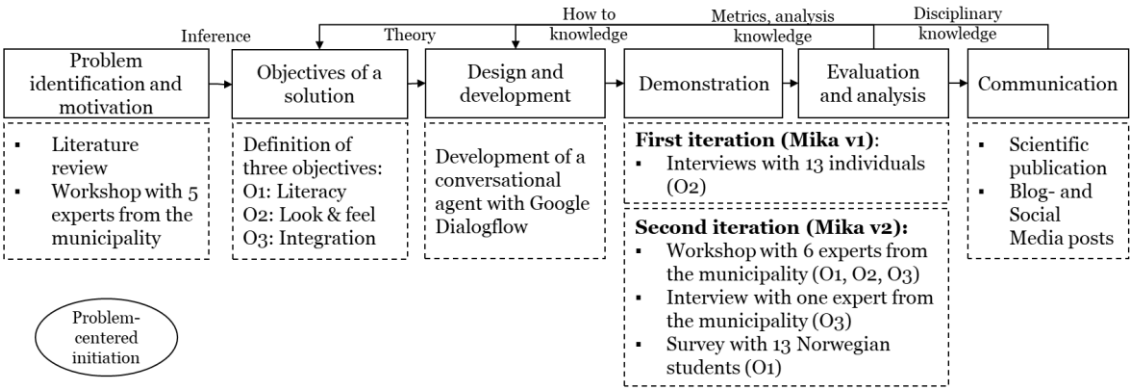

Figure 1. Overview of the applied iterative DSR process (based on Peffers et al., 2007)

3.1 Problem Identification and Motivation

The problem identification and motivation were partially described the previous sections of the paper. However, to refine our understanding of the specific challenges faced by PSOs, we conducted a workshop with practitioners from the municipality of Kristiansand, Norway (Table 2). The one-hour hybrid workshop was conducted in May 2022.

ID	Role	Expertise	W	FG
01	Health service manager	Crisis management, health	x	x
02	Chief physician	Crisis management, health	x	x
03	Communications officer	Communications, information dissemination	x	x
04	IT-advisor	IT-systems of the municipality, security	x	x
05	Emergency manager	Crisis management, detection of false information	x	x
06	Intercultural advisor	Communications, culture		x

Table 2. Participants of the requirement workshop (W) and focus group discussion (FG).

To ensure a common understanding, we first provided information on different types of false information online and on CAs. Then, each participant was interviewed on specific false information challenges and requirements that a CA would have to meet for being useful to the municipality. The interviews were documented with detailed notes and summarized.

The participants mentioned the following specific challenges in the context of false information online: i) false information (e.g., on vaccinations) is shared on official channels of the municipality, ii) politicians spread false information about the municipality, iii) ensuring a timely reaction to false information can be challenging as online communication cannot be monitored 24/7, iv) balancing freedom of speech and the necessity of correcting or deleting false information is difficult, and v) the municipality struggles reaching certain groups in Norwegian society (esp. young people and immigrants).

We will not be able to meet all of these specific challenges in the scope of this project. However, we noticed that challenges three and five would remain challenges in reactive strategies for combating false information online (e.g., fact checking CA). Thus, based on these challenges we suggest that a proactive, preventive approach would be the most promising strategy for PSOs. The objectives of a solution were subsequently determined based on the insights from the workshop and related literature. In the evaluation of the completed artifact, we will also discuss opportunities and limitations for addressing the remaining challenges with a CA.

3.2 Objectives of a Solution

Previous research found that preventive approaches such as developing media or information literacy effectively increase individuals' knowledge about media, awareness for the potential influence of media, perceived realism of media content, and criticism towards the media (Jeong, Cho and Hwang, 2012). We define media literacy as "*knowledge and skills that can help critical understanding and usage of the media*" (Jeong, Cho and Hwang, 2012, p. 455). In the context of false information, a study by Khan and Idris (2019) found that the ability to evaluate the validity of information online predicts whether unverified information is shared online. Thus, increasing media literacy is a promising strategy for PSOs to address problems related to false information online. Accordingly, we formulate the following objective:

First objective (O1): Interacting with the CA increases media literacy.

Additionally, the participants of the first workshop formulated several requirements for a CA. Specifically, i) users should be aware that they are interacting with a CA, ii) the CA should not collect personal information, iii) the CA should reference trustworthy and official sources, iv) the conversational style should be appropriate for the municipality (e.g., formality, humanness), and v) it

should be possible to integrate the CA on different communication channels of the municipality (i.e., website, Facebook, Instagram). We summarized these aspects in the following objective for the artifact:

Second objective (O2): The 'look and feel' of the CA is appropriate for both the audience and the PSO.

Last but not least, it would be crucial that the PSO sees value in the CA and would integrate it into their communication services. While fulfilling O1 and O2 already increases the likelihood of integration, we included this aspect as an additional objective:

Third objective (O3): The PSO would integrate the CA into their services.

3.3 Design and Development

Based on the three objectives, we developed a CA called "Mika" with Google Dialogflow[2]. Figure 2 provides overview of the functionality and appearance of Mika.

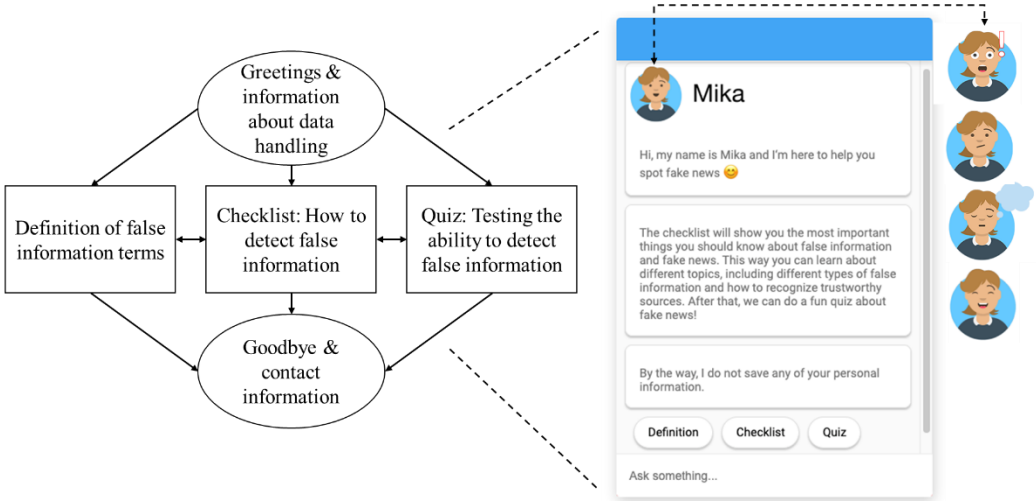

Figure 2. Overview of the functionality, the appearance of the chat window in which users could interact with Mika, and exemplary facial expressions of Mika.

The name, appearance, and conversational style of the CA was designed according to the municipality's requirements (O2) and considering design principles for CAs, such as social cues (e.g., greetings, farewell, different facial expressions of the CA in response to user input) to create a natural conversational flow and user experience (Feine, Morana and Maedche, 2019). The main interaction was designed with buttons to better guide unexperienced users, reduce errors, and avoid unintended fallbacks. In addition, the users could interact with the CA via textual input. Regarding media literacy (O1), Mika provides an overview of types of false information according to Zarocostas (2020) and a checklist on how to recognize false information online. As previous research showed that a CA could use a quiz to educate users on a certain topic (Schuetzler et al., 2021), we included a quiz for users to test their ability in identifying exemplary false information in the topics: Politics, climate, sports, health, society, and history. To avoid the risk of loss of trust in the CA due to non-transparency of the source of the information (Hofeditz et al., 2022), the CA gives additional resources to the information provided.

3.4 Demonstration

To test the ability of the artefact to support the municipality in combating false information online, we evaluated the CA in two iterations (Mika v1 and Mika v2) in different settings, focusing on different objectives.

[2] https://dialogflow.cloud.google.com/ (visited on November 17, 2022)

3.4.1 First Iteration: Evaluation with Potential Users

In the first iteration (Mika v1), we focused on the second objective, that is, the look and feel of the CA and that users understand the different functions of the CA and are able to interact with them in order to subsequently, evaluate the functionality. For this purpose, we interviewed 13 individuals (M_{age} = 32.15, SD_{age} = 13.27, eight females and five males). All participants provided their informed consent to record and analyze the interview. After some general questions on participants' previous knowledge about CAs and false information, participants were asked to interact with the CA while thinking aloud. Thus, participants were asked to state their thoughts out loud while performing a specific task and recall thoughts immediately following the completion of that task (Eccles and Arsal, 2017). This allowed the interviewers to understand and track thoughts, emotions, and concerns in relation to the CA. After the interaction, participants were asked open questions about specific aspects of the CA including the look and feel (i.e., usability, trustworthiness) and the perception of specific features of the CA (e.g., content of the quiz, navigation, educational value) and general feedback. An exemplary question was, "*How did you perceive the navigation through the CA's functions?".* The key findings of the interviews were noted, summarized, and analyzed in order to adjust the artifact accordingly for the next iteration.

3.4.2 Second Iteration: Evaluation with the Municipality

In the second iteration, we conducted a focus group discussion with the municipality of Kristiansand in August 2022 to assess the integration of the CA (Mika v2) into existing services of the municipality and evaluate the CAs' ability to improve media literacy in dealing with false information. The discussion was conducted with six participants (Table 2). The workshop was introduced with a small questionnaire containing two questions about the perceived media literacy of citizen and the responsibility of the municipality in educating on false information. To create a common understanding of the issue and the conducted work in this research project, we provided some general information about false information online, the concept of media literacy and CAs as well as a recap about the previous workshop with members of the municipality. The participants of the focus group discussion were then asked to interact with the CA to evaluate the objectives and an open discussion was moderated. Finally, the focus group discussion was closed with two questions about the usefulness of the CA.

To gain further insight into the municipality's existing IT-infrastructure, the focus group participants suggested conducting another interview with an expert from Kristiansand municipality's communications department who was unable to join the focus group discussion. Therefore, a semi-structured interview was conducted in August 2022 to gain more expert knowledge about the challenges and opportunities of CA integration from an expert involved in the municipality's internal crisis communication processes and implementation and maintenance of its existing CA. Key findings of the focus group discussion and interview were noted and summarized.

3.4.3 Second Iteration: Evaluation with Potential Users

To evaluate the potential of the CA to increase media literacy in a relevant user group, we invited 13 students from a master-level course in the domain of IS of the University of Agder, Norway to interact with the CA and provide their feedback through a survey. The pre-post evaluation survey consisted of three parts: i) general questions about previous experiences with false information online and self-assessment of one's ability to detect false information, ii) self-guided interaction with the CA, and iii) self-assessment of one's ability and evaluation of the CA. The survey was conducted in September 2022. Participation in the survey was voluntary and we did not collect any personal information.

4 Evaluation and Analysis

4.1 First Iteration: Results

Overall, the comments and answers of the participants regarding the usability of the CA were inductively classified into different categories, such as findable, credible, useful, desirable, and accessible to create

a common understanding of the feedback. In general, the CA was perceived as appealing and useful. Anyway, the interviewees provided some feedback on improvements regarding the accessibility of the CA and the content. Six interviewees were facing difficulties regarding the accessibility because they did not have a clear idea of what the checklist was about. Additionally, three interviewees indicated that they were facing difficulties with the navigation through the buttons when fulfilling a task and trying to return to the menu. Regarding the content of the CA, two interviewees thought it would be useful to have additional context in the explanations of false information in the quiz. In addition, four respondents suggested changing technical terms to make the CA more accessible and two interviewees drew attention to grammatical errors to guarantee credibility. Besides, one interviewee felt the CA's language was judgmental when conducting the survey, perceiving "*you should do more research next time*" as undesirable. Based on these key findings, the CA was adapted by addressing all of the previously mentioned issues. Thus, an improvement of the appearance, the interaction with and the provided content by the CA could be ensured. However, other responses and observations were rated as useful, such as feedback on the scrolling, the image size or general feedback on colors and backgrounds.

4.2 Second Iteration: Results

In the following, we summarize the main results from the evaluation with the municipality (E1) and potential users (E2) in relation to the three objectives (O1, O2, O3).

The evaluation of O1 revealed that the approach of providing a CA to increase the citizens media literacy was perceived as relevant because the trust in media was recently perceived to decline due to clickbait, hate speech, and political polarization, especially on social media (E1). Consequently, the municipality has been concerned with providing accurate information to all citizens, but there are difficulties in reaching all groups in terms of age and nationality. The questionnaire (E1) showed that on a 4-point Likert Scale (1 = strongly disagree, 4 = strongly agree) the focus group participants had different assessments about the statement that Norwegian citizens know how to detect false information (three agreed, three disagreed) but there was an agreement that the municipality has the responsibility to educate citizens about false information (five agreed, one strongly agreed). It was also considered important for citizens themselves to be able to evaluate information properly (E1, E2). It was stated that there is no perceived strategic media literacy program provided by organizations of the public sector but there is a willingness to build internal responsibility for media literacy in the municipality (E1). In addition, it was mentioned that the educational CA can be used in schools (E1, E2) and educate in more than false information (E1). Moreover, the reference to current news was perceived as a desirable characteristic of the quiz (E1) and the CA was perceived as a useful and enjoyable tool for increasing media literacy (E1, E2). This was also reflected in the questionnaire of the focus group discussion, where all remaining participants (n = 5) agreed that the CA would help the municipality to educate about false information and would complement existing communication initiatives. In addition, the survey (E2) revealed that on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree) the participants on average agreed (= 5.61) that the CA is a useful tool for increasing media literacy.

In general, the look and feel of the CA (O2) and the general interaction with the CA was perceived as appealing and pleasant (E1, E2). Especially, the avatar's facial expression and humorous communication style were highlighted as positive. The interaction via buttons was described as inclusive as it allows easy communication (E1). Nevertheless, the option to ask questions directly via text input was found useful, as was the inclusion of multiple and diverse avatars to provide a range of different agents (E1). The use of a gender-neutral avatar was positively highlighted (E1). Moreover, participants suggested to include more than one language to make the CA more accessible (E1, E2).

The possibility to integrate the CA into many different platforms and messengers was perceived as very useful (O3). The false information CA could either be combined with the existing CA of the municipality or it could be referenced by the municipality CA. The minimization of maintenance effort and definition of responsibilities were considered important (E1). However, the integration interview revealed that the effort for maintaining the CA by providing additional information is considered to be low and therefore does not represent a challenge. Additionally, the municipality CA is provided by a company and used

by several municipalities which have access to a shared database (E1). Since the same provider and software is used, the integration of the false information CA in multiple municipalities can be simplified through cooperation with the corresponding company. This might allow for economies of scale and minimize maintenance resources. Lastly, it was emphasized that strict data protection is important, but also that the use of innovative technologies in the municipality's public sector is desirable and that the municipality is willing to test new initiatives (E1).

5 Conclusion and Next Steps

The aim of this research was to investigate how a CA could be designed to support PSOs in addressing the problems of false information. The municipality of Kristiansand (Norway) was selected as a case for this research. Based on a literature review and the specific challenges faced by the municipality of Kristiansand, increasing media literacy with CAs emerged as a promising strategy. Accordingly, we developed the CA "Mika" which provides information about identifying false information online and allows users to test their ability of identifying false information.

In this research-in-progress paper, we described two iterations of the CA development and evaluation as part of a DSR process. The preliminary findings indicate that the CA is perceived by both PSOs and potential users as a useful tool for improving media literacy, thus providing insights for the development of CA for PSOs to combat false information online. However, these insights are only a starting point for the assessment of the CA for PSOs. For example, we evaluated the CA with Norwegian students as potential users, however, they are not representative for society and can only provide initial evidence. Furthermore, some challenges for integrating the CA into the services of PSOs might be difficult to identify in a hypothetical scenario. Lastly, the CA currently has a relatively narrow focus which was useful for a proof-of-concept evaluation but can be adapted to include more methods that are useful for increasing media literacy that are relevant in the context of false information online, such as providing advice and tips on evaluating the correctness of a specific piece of information.

To conduct an in-depth evaluation of the first objective, the effectiveness of the CA in improving media literacy, we aim to complete the demonstration and evaluation of the second iteration by measuring the CA's potential to improve media literacy with a larger and more diverse group of Norwegian citizens and by evaluating the citizen-oriented perspective on the requirements for the CA. Because low user engagement with government digital services, such as websites, is a challenge (Huang and Benyoucef, 2014), the citizen-oriented evaluation will explore requirements for the CA to increase user engagement on the government website and social media channels. As we noticed that several Norwegian municipalities use CAs that are developed and maintained by the same provider, we aim to explore challenges and opportunities for integrating Mika into these services directly with the provider in more detail. Following the DSR process, we will finally evaluate if another adaptation of the artifact is necessary to satisfy the objectives and to derive general learnings on how PSOs could design CAs for combating false information and whether further iterations will be conducted to take advantage of the CA's capabilities. By completing the iterations of the DSR, this work contributes with general findings from a holistic assessment of a CA as a public service to combat false information by improving citizens' media literacy.

Acknowledgements

This research has been co-funded by the German Federal Ministry for Education and Research (BMBF) in the project PREVENT (16KIS1491K) and by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 823866.

References

Allcott, H. and Gentzkow, M. (2017) 'Social media and fake news in the 2016 election', *Journal of Economic Perspectives*, 31(2), pp. 211–236. doi: 10.1257/jep.31.2.211.

Bannister, F. and Connolly, R. (2014) 'ICT, public values and transformative government: A framework and programme for research', *Government Information Quarterly*, 31(1), pp. 119–128. doi: https://doi.org/10.1016/j.giq.2013.06.002.

Brachten, F., Brünker, F., Frick, N.R.J., et al. (2020) 'On the ability of virtual agents to decrease cognitive load: An experimental study', *Information Systems and e-Business Management*, 18(2), pp. 187–207. doi: 10.1007/s10257-020-00471-7.

Brachten, F., Kissmer, T. and Stieglitz, S. (2021) 'The acceptance of chatbots in an enterprise context – A survey study', *International Journal of Information Management*, 60(October 2021), p. 102375. doi: 10.1016/j.ijinfomgt.2021.102375.

Ceron, W., de-Lima-Santos, M.-F. and Quiles, M. G. (2021) 'Fake news agenda in the era of COVID-19: Identifying trends through fact-checking content', *Online Social Networks and Media*, 21, p. 100116. doi: https://doi.org/10.1016/j.osnem.2020.100116.

Diederich, S., Brendel, A. B. and Kolbe, L. M. (2020) 'Designing anthropomorphic enterprise conversational agents', *Business and Information Systems Engineering*, 62(3), pp. 193–209. doi: 10.1007/s12599-020-00639-y.

DiFonzo, N. and Bordia, P. (2007) *Rumor psychology: Social and organizational approaches.* Washington: American Psychological Association. doi: 10.1037/11503-000.

Duan, Y., Edwards, J. S. and Dwivedi, Y. K. (2019). 'Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda', *International Journal of Information Management*, 48, pp. 63–71. doi: 10.1016/j.ijinfomgt.2019.01.021.

Dumitru, E.-A., Ivan, L. and Loos, E. (2022) 'A generational approach to fight fake news: In search of effective media literacy training and interventions', in Gao, Q. and Zhou, J. (eds.) *Human Aspects of IT for the Aged Population Design, Interaction and Technology Acceptance*. Cham: Springer International Publishing, pp. 291–310. doi: 10.1007/978-3-031-05581-2_22.

Eccles, D. W. and Arsal, G. (2017) 'The think aloud method: What is it and how do I use it?', *Qualitative research in sport, exercise and health.*, 9(4), pp. 514–531. Available at: http://dro.dur.ac.uk/22254/.

Feine, J., Morana, S. and Maedche, A. (2019) 'Designing a chatbot social cue configuration system', in *40th International Conference on Information Systems, ICIS 2019*.

Henman, P. (2020) 'Improving public services using artificial intelligence: possibilities, pitfalls, governance', *Asia Pacific Journal of Public Administration*, 42(4), pp. 209–221. doi: 10.1080/23276665.2020.1816188.

Hofeditz, L., Ehnis, C., Bunker, D., et al. (2019) 'Meaningful use of social bots? Possible applications in crisis communication during disasters', in *27th European Conference on Information Systems, ECIS 2019*.

Hofeditz, L., Mirbabaie, M., Erle, L., et al. (2022) 'Automating crisis communication in public institutions – Towards ethical conversational agents that support trust management', in *Wirtschaftsinformatik 2022 Proceedings*.

Huang, Z. and Benyoucef, M. (2014) 'Usability and credibility of e-government websites', *Government Information Quarterly*, 31(4), pp. 584–595. doi: https://doi.org/10.1016/j.giq.2014.07.002.

Jeong, S. H., Cho, H. and Hwang, Y. (2012) 'Media literacy interventions: A meta-analytic review', *Journal of Communication*, 62(3), pp. 454–472. doi: 10.1111/j.1460-2466.2012.01643.x.

Jørgensen, T.B. and Bozeman, B. (2007) 'Public values: An inventory', *Administration \& Society*, 39(3), pp. 354–381. doi: 10.1177/0095399707300703.

Khan, M. L. and Idris, I. K. (2019) 'Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective', *Behaviour and Information Technology*, 38(12), pp. 1194–1212. doi: 10.1080/0144929X.2019.1578828.

McTear, M., Callejas, Z. and Griol, D. (2016) *The Conversational Interface: Talking to Smart Devices*. 1st edn. Springer Cham. doi: https://doi.org/10.1007/978-3-319-32967-3.

Meel, P. and Vishwakarma, D. K. (2020) 'Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities', *Expert Systems with Applications*, 153, p. 112986. doi: https://doi.org/10.1016/j.eswa.2019.112986.

Mheidly, N. and Fares, J. (2020) 'Leveraging media and health communication strategies to overcome the COVID-19 infodemic', *Journal of Public Health Policy*, 41(4), pp. 410–420. doi:

10.1057/s41271-020-00247-w.
Mirbabaie, M., Stieglitz, S., Brünker, F., et al. (2021) 'Understanding Collaboration with Virtual Assistants – The Role of Social Identity and the Extended Self', *Business and Information Systems Engineering*, 63(1), pp. 21–37. doi: 10.1007/s12599-020-00672-x.
Mirbabaie, M., Stieglitz, S. and Brünker, F. (2022) 'Dynamics of convergence behaviour in social media crisis communication – a complexity perspective', *Information Technology and People*, 35(1), pp. 232–258. doi: 10.1108/ITP-10-2019-0537.
Moore, M. H. (1995). *Creating public value: Strategic management in government*. London: Harvard university press.
Naeem, S. Bin, Bhatti, R. and Khan, A. (2021) 'An exploration of how fake news is taking over social media and putting public health at risk', *Health Information \& Libraries Journal*, 38(2), pp. 143–149. doi: 10.1111/hir.12320.
Peffers, K., Tuunanen, T., Rothenberger, M., et al. (2007) 'A Design Science Research Methodology for Information Systems Research', *Journal of Management Information Systems*, 24(3), pp. 45–77. doi: 10.2753/mis0742-1222240302.
Peffers, K., Tuunanen, T. and Niehaves, B. (2018) 'Design science research genres: Introduction to the special issue on exemplars and criteria for applicable design science research', *European Journal of Information Systems*, 27(2), pp. 129–139. doi: 10.1080/0960085X.2018.1458066.
Rose, J., Persson, J.S., Heeager, L.T., et al. (2015) 'Managing e-Government: Value positions and relationships', *Information Systems Journal*, 25(5), pp. 531–571. doi: https://doi.org/10.1111/isj.12052.
Roque, G., Cavalcanti, A., Nascimento, J., et al. (2021) 'BotCovid: Development and evaluation of a chatbot to combat misinformation about COVID-19 in Brazil', in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2021, pp. 2506–2511. doi: 10.1109/SMC52423.2021.9658693.
Schuetzler, R. M., Giboney, J. Grimes, G., et al. (2021) 'Deciding whether and how to deploy chatbots', *MIS Quarterly Executive*, 20(1), pp. 1–15. doi: 10.17705/2msqe.00039.
Shu, K., Mahudeswaran, S., Wang, S., et al. (2020) 'FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media', *Big Data*, 8(3), pp. 171–188. doi: 10.1089/big.2020.0062.
Stieglitz, S., Mirbabaie, M., Schwenner, L., et al. (2017) 'Sensemaking and Communication Roles in Social Media Crisis Communication', in *Wirtschaftsinformatik 2017 Proceedings.*
Stieglitz, S., Hofeditz, L., Brünker, F., et al. (2022) 'Design principles for conversational agents to support Emergency Management Agencies', *International Journal of Information Management*, 63(December 2021). doi: 10.1016/j.ijinfomgt.2021.102469.
Tandoc, E. C., Lim, Z. W. and Ling, R. (2018) 'Defining "fake news": A typology of scholarly definitions', *Digital Journalism*, 6(2), pp. 137–153. doi: 10.1080/21670811.2017.1360143.
Tavanapour, N., Poser, M. and Bittner, E. A. C. (2020) 'Supporting the idea generation process in citizen participation - Toward an interactive system with a conversational agent as facilitator', in *27th European Conference on Information Systems, ECIS 2019*.
Tendai, M., Alireza, N., Desouza, K., et al. (2020) 'Chatbot-mediated public service delivery: A public service value-based framework.', *First Monday*, 25(12), pp. 4644–4646. doi: 10.1109/BigData47090.2019.9005561.
Tran, T., Rad, P., Valecha, R., et al. (2019) 'Misinformation Harms During Crises: When The Human And Machine Loops Interact', in *2019 IEEE International Conference on BigData (Big Data)*. Los Angeles, CA, USA: IEEE, pp. 4644–4646. doi: 10.1109/BigData47090.2019.9005561.
Vassilakopoulou, P., Haug, A., Salvesen, L.M., et al. (2023) 'Developing human/AI interactions for chat-based customer services: Lessons learned from the norwegian government', *European Journal of Information Systems*, 32(1), pp. 10–22. doi: 10.1080/0960085X.2022.2096490.
Zarocostas, J. (2020) 'How to fight an infodemic', *The Lancet*, 395(10225), p. 676. doi: 10.1016/S0140-6736(20)30461-X.