

Research Article

Jasmin Haunschild, Markus Henkel* and Christian Reuter

Breaking down barriers to warning technology adoption: usability and usefulness of a messenger app warning bot

<https://doi.org/10.1515/icom-2024-0067>

Received November 29, 2024; accepted February 3, 2025;

published online February 26, 2025

Abstract: In crisis situations, citizens' situational awareness is paramount for effective response. While warning apps offer location-based alerts, their usage is relatively low. We propose a personalised messaging app channel as an alternative, presenting a warning bot that may lower adoption barriers. We employ the design science research process to define user requirements and iteratively evaluate and improve the bot's usability and usefulness. The results showcase high usability, with over 40 % expressing an interest in utilising such a warning channel, stressing as reasons the added value of proactive warnings for personalised locations while not requiring a separate app. The derived requirements and design solutions, such as graphically enhanced user interface elements as guardrails for effective and error-free communication, demonstrate that a suitable warning chatbot does not necessarily require complex language processing capabilities. Additionally, our findings facilitate further research on accessibility via conversational design in the realm of crisis warnings.

Keywords: warning bot; conversational interface design; crisis communication; mobile warning; botplication

1 Introduction: problem identification and motivation

Informing the population of impending hazards is crucial for crisis response. Events, such as the European floods in

*Corresponding author: **Markus Henkel**, Science and Technology for Peace and Security (PEASEC), Technical University of Darmstadt, Darmstadt, Germany, E-mail: henkel@peasec.tu-darmstadt.de.
<https://orcid.org/0000-0003-0574-890X>

Jasmin Haunschild and Christian Reuter, Science and Technology for Peace and Security (PEASEC), Technical University of Darmstadt, Darmstadt, Germany, E-mail: haunschild@peasec.de (J. Haunschild), reuter@peasec.tu-darmstadt.de (C. Reuter).
<https://orcid.org/0000-0002-6215-7973> (J. Haunschild)

2021 that killed over 200 people, demonstrate the devastating effects of insufficient crisis awareness. Due to the widespread use of mobile phones, cell broadcast and warning apps are increasingly used to alert the population. Warning apps combine multi-hazard information in a single platform,¹ which can be personalised with regard to personally relevant hazards and locations.² By contrast, cell broadcast (CB) distributes warnings to all mobile phones connected to a cell tower, offering resilience through low bandwidth and independence from the Internet. CB achieves almost perfect reach in countries with the necessary infrastructure,² but lacks personalisation. Whereas this is offered in warning apps, these have a low to medium reach,² as users need to actively download them. This incurs costs in choosing an app, battery drainage and by taking up data space,³ reducing the intention to use and to continue to use a warning app.⁴ Additionally, some countries' risk cultures lead to a lower perceived individual responsibility to prepare,^{5–7} which may further reduce the motivation to download a warning app. To address these concerns, additional warning channels should be explored. Instant messaging (IM) apps appear particularly viable due to their widespread use.

This idea is supported by current trends in news delivery, which can inform the spread of crisis information. In some countries, IM apps, which are widely and increasingly used around the globe,⁸ have become the main avenue of news consumption.⁹ Major news outlets have developed distribution bots for social media platforms and messaging apps.¹⁰ Following this trend, agencies have begun circulating crisis related information using messenger apps, such as the German Ministry of Health, informing about the COVID-19 pandemic via Telegram.¹¹

We stipulate that a warning channel within an IM app could lower adoption barriers and reduce the maladaptive rewards associated with downloading and using warning apps that appear to deter users with a low protection motivation.^{3,4} Indeed, a representative study has found that one third (34 %), and especially current non-users of warning apps, would prefer such a channel to a warning app.¹² However, warning apps have specific design requirements

resulting from the use in critical situations,¹³ and users want to personalise warnings and locations.¹⁴

An analysis of chatbot design guidelines identifies three studies that offer moderate re-usability, however, they are in the domains of health, education, and customer service rather than the public service or crisis management domains.¹⁵ One study compared the effect of using an app or a chatbot for reducing problematic drinking.¹⁶ As the bot increased the intention to change the behaviour but not the action itself, app and bot should be used complementarily. As human-bot interactions are highly context dependent¹⁵ and no studies exist related to public crisis warnings via IM apps made for that purpose, we design a chatbot that mimics warning apps by helping users set up warning subscriptions for relevant locations, warning types and severity levels. Due to the particular design requirements for warning channels,¹⁷ existing guidelines for designing UI and conversational interfaces^{18,19} must be adapted to the warning domain.

We follow the design science research (DSR) approach^{20,21} and its six steps: (1) identifying a problem, (2) defining objectives for a solution, (3) designing and developing an artifact, (4) demonstrating and (5) evaluating it, before (6) communicating the findings.²² Figure 1 shows the iterative development informed by user evaluations and the sections that describe each step. The DSR process closely aligns with human-centered design principles (ISO 9241-210), as both emphasise understanding user needs, iteratively designing solutions, and evaluating them. Specifically, our work adheres to these principles by

analysing usability requirements and assessing the roles of usability and usefulness for IM apps as warning channels.

After identifying the problem and motivating the topic in Section 1, we analyse usability requirements for warning apps and bots to define our solution's objectives in Section 2. Section 3 describes a conversational design workshop conducted to design the user journeys, personas and conversational interface. Next, we present the iteratively developed prototype, a *Telegram* warning bot, which uses an official warning app's API to send alerts and allows users to personalise warning subscriptions (Section 4). Section 5 presents two evaluation studies: First, we conducted interviews focusing on task fulfilment and usability. Secondly, participants performed tasks and evaluated usability, reward and, usage intention through an online survey. The results and limitations of our work are discussed in Section 6. Section 7 summarises and concludes the paper.

This study makes several contributions: through the DSR process, we contribute findings based on exaptation,²⁴ meaning the application of existing findings to a new domain based on the design and evaluation of an artifact. We use existing findings on the design of warning apps and chatbots and apply them to the domain of mobile public warning systems. The study results in the development of an artifact, a *Telegram* warning bot, which is available for further studies and development. Furthermore, by conducting user studies with a warning channel that aims to lower adoption barriers, our work offers insights into the effects of complementing smartphone apps with IM "botplications" and into the perceived advantages and disadvantages, utility and usability of such tools.²⁵

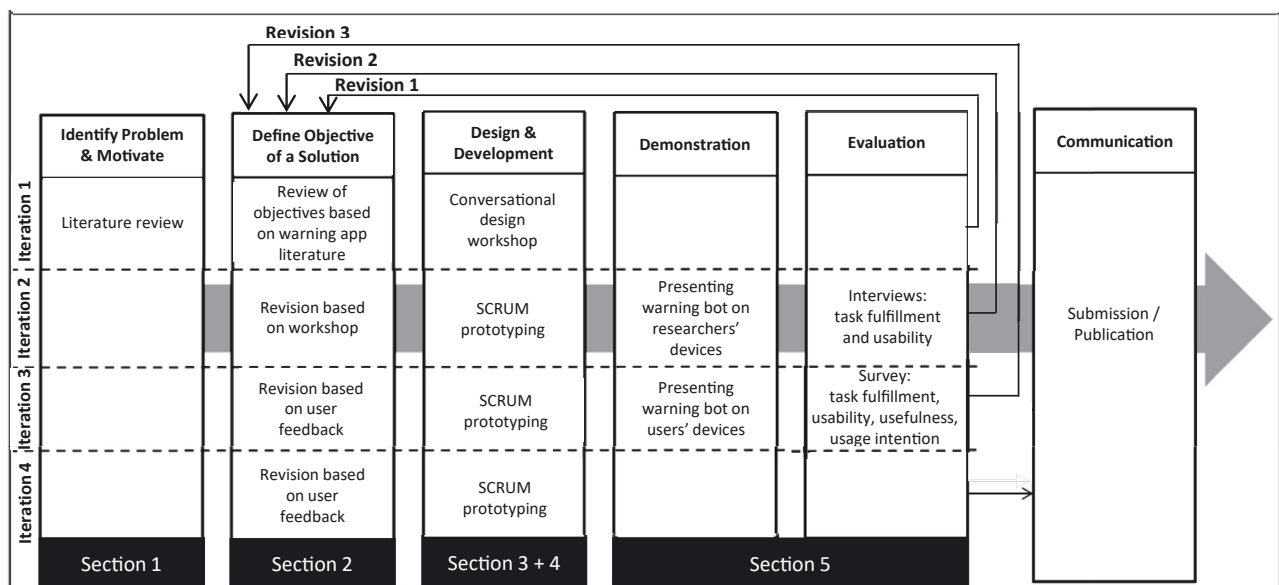


Figure 1: Research approach of our paper using the design science research (DSR) process,^{20,21} own figure adapted from Ref. ²³.

2 Objectives of the solution

The following section summarises the findings of the literature that describe the user requirements for warning apps, followed by a review of requirements for chatbots, from which we derive the objectives for the warning bot.

2.1 Mobile warnings and warning app requirements

Warning apps provide a mobile channel that allows to personalise warnings based on locations, by selecting relevant areas or via GPS-tracking.¹ In addition, relevant hazard types can be selected, e.g. severe weather warnings, bomb disposal, school closure or policing information.² As such, they are distinct from cell broadcast, which is used only for severe civil protection warnings and sends alerts only based on the current radio cell of the mobile device. We analysed the literature on warning apps to identify warning and design requirements that might also be relevant for other warning channels. The literature has analysed user preferences regarding warning apps,^{12,14,26} and reasons for warning app (non-)adoption and (dis-)continuance of use.^{4,27} Other studies have contributed warning app design requirements²⁸ and usability guidelines.¹³ Compared to social media, people value warning apps for their fast and reliable information.²⁹ For those who have used a warning app in a crisis, this information channel is perceived as more helpful than other online sources.¹² Users prefer an app that centralises various types of warnings and allows personalisation based on personally relevant warnings.^{12,14} In line with these findings, a study shows that the key factor that affects the duration of use of a warning app is its perceived usefulness. Users are more likely to continue using the app if they feel that it effectively serves its intended purpose – reliably providing critical information in an easy-to-use manner – and is free of errors.²⁷

A study of app reviews suggests that users of warning apps discontinue using warning apps due to reliability issues, with users not receiving relevant warnings or receiving false alarms at other times.³⁰ Usability factors also play a role: users discontinue using warning apps that have complex user interface graphics or require too much user input.²⁷ In addition, warning apps have specific design requirements that differ from the requirements of other apps, which relate to simplicity and ease of use, as well as demanding little energy, bandwidth, and storage space, to account for the specific use case in crises.¹⁷

Looking at users' motivations regarding the use of warning apps, Fischer-Preßler et al.⁴ find that maladaptive

rewards – advantages derived from not using a warning app, such as saving battery and data space – have a negative influence on app adoption and continued use. Similarly, inhibition theory posits that potential losses are perceived more strongly than potential gains,³¹ stressing the importance of effort expectancy and other perceived costs for digital service adoption.³² Finally, citizens in state-oriented risk cultures have a low motivation to engage in crisis management and preparedness,^{6,7} stressing the importance of low adoption costs.

Bonaretti et al.²⁸ point towards the importance of timeliness and representational fidelity of warning apps, meaning that warnings need to be easily and quickly noticed and understood by users, e.g. through good visual representations of the crisis and a noticeable auditory and visual alarm signal. Furthermore, the app must be dependable with regard to its warnings, sending an adequate alert level for each relevant and only for relevant emergencies.

Finally, Tan et al.¹³ present seven higher-level constructs with usability requirements, some specific to warning apps. They suggest (1) app dependability, consisting of instant start, data preservation, error-free operation and clean exit; (2) app design, including branding, orientation, minimal advertising, phone resource usage; (3) app utility, consisting of collaboration, content relevance and search; (4) UI graphics, with the requirements aesthetic graphics, realism and subtle animation; (5) UI input, including control obviousness, de-emphasis on settings, effort minimisation, fingertip size controls; (6) UI output, requiring concise language, standardised UI elements, user-centric terminology and audio output; and (7) UI structure with the elements logical path, top-to-bottom structure and few external links. In addition, Reuter et al.³³ show that users are reluctant to share private information with state agencies, highlighting the need for privacy by design.

We identify a set of objectives for mobile warnings:

- *Usefulness*:
 - *Adaptability*: offer many relevant instances of interaction through personalisation^{12,14,27}
 - *Representational fidelity*: reliably represent the warning status, is free of errors²⁸
 - *Timeliness*: noticeable warnings, fast delivery²⁸
 - *Criticality*: reliably present critical warnings³⁰
- *Resource efficiency*: minimise used data space, energy resources, learning and maintenance effort^{4,17}
- *Simplicity*: little user input, simple UI graphics, fast information extraction²⁷
- *Usability*: optimised for critical situations and privacy^{13,33}

2.2 Bot requirements

Chatbots are used in a large variety of fields, particularly in customer service support and e-commerce, but state agencies are also increasingly using them to offer information about regulations and public services.³⁴ A study analysing the interactions with a public service bot has shown that users largely use very short, utilitarian rather than social inputs and that this brevity is technically justified, as it results in a vastly higher rate of adequate responses by the bot.³⁴ Another study looking at older adults found that Internet competency had a great effect on whether social-oriented or task-oriented interaction styles were more beneficial in e-commerce, with task-orientation performing better for people with lower Internet competency scores.³⁵

Within the public service category, a warning bot pertains to the category of content curation chatbots, which “have a chatbot-driven Locus of Control where the user initiative is limited to accepting or rejecting content offers, or requesting specific content types, serving to filter the presented content selection”.³⁶ These are increasingly developing from one-off interactions to longer-term relationships with regular content provision. Because browsing and searching are less well-suited to dialogue interfaces, these types of chatbots “actively guide users to recommended content” by presenting options, often by using visual elements.³⁶

Klopfenstein et al.²⁵ identify that bots are increasingly run within IM apps to serve as functional replacements of mobile applications, e.g. as news bots instead of news apps. Such “botplications” offer a set of advantages for users and programmers: They are instantly available inside IM apps, require little data traffic, their push notifications appear as messages, the messages can be easily shared within the app and beyond, most messaging apps are independent of the operating system and the hosting platform offers user authentication.²⁵

However, bots have some significant differences compared to apps, especially regarding the User Interface (UI). Conversational user interfaces (CUIs) are sequential, characterised by turn-taking.^{18,37} In addition to text, they support pictures, stickers, videos, audio, generic data files and packaged data, such as geolocations or contact information.²⁵ In messaging bots, services and information are provided as streams of messages and notifications. Dialogues can take the form of text, voice, video input, or images.³⁸ They can be deployed on various platforms, including websites, messaging applications or social media platforms. In a bot, each message “should be conceptually seen as a micro application, while the conversation is a timeline of past application screens” 25, p. 559. A challenge also lies in

guiding the conversation: “because of the free-form nature of the medium, it is easy for bot users to get lost and not to be certain of what commands or what exact syntax is required to perform the desired action” 25, p. 560. Therefore, the bot should suggest next actions, use buttons or suggested syntax as guardrails.²⁵ To ensure guidance, simplicity, effectiveness, and to avoid errors, research also suggests avoiding natural language processing 25, p. 559. This rules out unexpected user input³⁹ and avoids misunderstandings by the bot,⁴⁰ which would lead to incorrect answers or a lack of responses, and in turn decrease trust in the bot, and thus negatively affect the user experience.⁴¹ Users also often prefer buttons to entering text as this provides them with clear input options enhancing control 42, p. 257.

Conducting a literature review of task-oriented conversational agents, Göbel et al.⁴³ identify five design requirements (goal orientation, usability, communication efficiency, human-likeness, user-centric integrity) and eleven design principles (goal extraction, conversation recovery, accessibility, adaptability, feedback and responsiveness, cultural sensitivity, natural communication, adaptive personality, emotional engagement, transparency, ethics and compliance).

The literature points to the following challenges and requirements for bots and CUIs:

- *Onboarding*: introducing function and interactions^{38,39,44}
- *User support and learning*: providing operating aids and allowing users to learn about functions on the fly³⁹
- *Context maintenance*: saving previous inputs and being able to refer to them at a later point in time^{38,39,41,44,45}
- *Guided conversation*: proactively suggesting follow-up actions, offering alternative choices, offering UI enhancements such as buttons or syntax commands²⁵
- *Social intelligence*: dealing with inappropriate and unspecified input³⁹
- *Manage conversation breakdown*: offering options of potential intentions after unfamiliar input⁴⁶
- *Transparency*: being clear about the process, especially for complex actions³⁹
- *Autonomy*: ensuring control over the interactions⁴⁷
- *Efficiency*: satisfying results in short time,⁴⁸ quick responses to minimise user input^{44,49}
- *Effort minimisation*: reducing cognitive effort by structuring process intuitively^{50,51}
- *Privacy*: following data protection laws and collecting minimal user data³⁸

3 Warning bot design

To decide how to translate the collected requirements into a design, we conducted a conversational design workshop with the involved researchers and programmers ($N = 5$), lasting 2 h. The decisions were based on the researchers' understanding of the knowledge base related to user preferences and user needs (see Section 2.1) as well as bot requirements described in the research literature (see Section 2.2). We applied the framework of conversational design,⁵² using workshop material for conversational interface design from a UX consulting agency, to translate the knowledge base into design decisions. The workshop focused on understanding user needs, motivations and behaviours; exploring the usage context and the expected user group(s); crafting language and tone by determining the appropriate language style, tone of voice, and the bot's personality to align with the context of warnings; defining conversation flows by creating user journeys and ideating the graphical elements of the user interface. The implemented design is discussed in more detail in Section 4. The literature review and the design workshop culminated in a collection of issues, derived meta-requirements and design implications, and settled on design solutions, which are presented in Figure 2. The key steps and exemplary outcomes are described in the following.

First, we discussed the key problem to be solved for users to achieve crisis awareness based on crisis informatics literature and user-centered studies on public warnings. For example, since people prefer to combine all relevant crisis functions in one tool, we chose to offer both warnings and preparedness information to increase utility (MR1).

Then, we described our targeted user group(s) and the usage context. We aim for situational awareness for the general population, i.e. people who have varying levels of technology affinity and who are moderately interested in receiving crisis warnings but not motivated enough to download a dedicated warning app (I1). By allowing users to subscribe to alerts via a messenger channel (DS1), we lower the barrier to adoption (MR3) and avoid maladaptive rewards, as users do not need to give up storage space for the installation of a standalone warning app (I3). In addition, predominantly, people have little experience with warning apps and personalised messenger bots, and we expect them to be unfamiliar with the set-up process (I5). Related to the conversational agent, this leads to the requirement of onboarding (MR6) and learnability (MR7) as well as a first message that introduces the bot and suggests the first next step (DS4) and the inclusion of an FAQ section (DS5). Related to the context, we expect the set-up of the bot to take place in a calm environment, whereas queries, warnings

and looking up of advice may occur in stressful situations, when users may be worried about a crisis (I4), leading, for example, to the requirement of reliability (MR4) and the offering of offline crisis management advice (DS2).

Then, we discussed the persona of the bot, focusing on its personality, its identity, role, and relationship with the user, and the expected communication style. Due to the serious nature of the topic, we orient the bot towards providing a *service conversation*, which is “constrained by rules and roles (e.g. customer service agent/customer)”.³⁷ Extrapolating from user requirements for warning apps, we assume that users want an error free (I6) and efficient process (I8) and focus on a bot-guided conversation (MR8) that can be conducted through buttons (DI8) and little text input (DS6), which is supported by this bot persona.

Finally, we devised the conversational flow based on the expected user journey (MR12), assuming that users would follow specific paths in their interactions with the system. We assume that users want to set up personalised (MR2) subscriptions for proactive warnings at specific locations (*push mode*); that they will request information about the warning status at a specific location (*pull mode*) to familiarise themselves with the bot and its output (I2); that they will want to look up crisis management and preparedness advice; that they will want to manage, adjust, or delete their subscriptions and settings; and that they will inquire about the bot, its functions and data (DI11). As the bot immediately starts (I10), we made sure to allow users to query the warning status as soon as the chat window opens (DS11).

By simulating the user journeys, we defined the conversation flow, seeking to increase intuitiveness, control, and transparency (MR17) to establish trustworthiness (I11). We also designed the graphical user interface, deciding on texts that would best describe the buttons' functionalities (see Section 4). As typing and reading is perceived as effort (I9), we decided to use a visual-centric conversation style (DI13) to reduce error rate, in which “graphical elements, such as buttons, images, emojis, and other visual elements, are mixed into the interaction alongside natural language inputs”⁵², p. 12. Therefore, buttons are used as the primary input channel to steer the conversation and make suggestions for how to proceed, reducing error, recovery time, and typing effort (MR14, MR15).^{25,42} The use of buttons, for example, for deletion of data or navigation back to the menu, also allowed us to steer the conversation and take autonomy (MR18), dealing with the issue of limited clickable options in messengers compared to warning apps (I12). Natural language is only used to input locations. Users can use the command “/help” to request assistance related to their previous interactions at any point (DS8). While natural

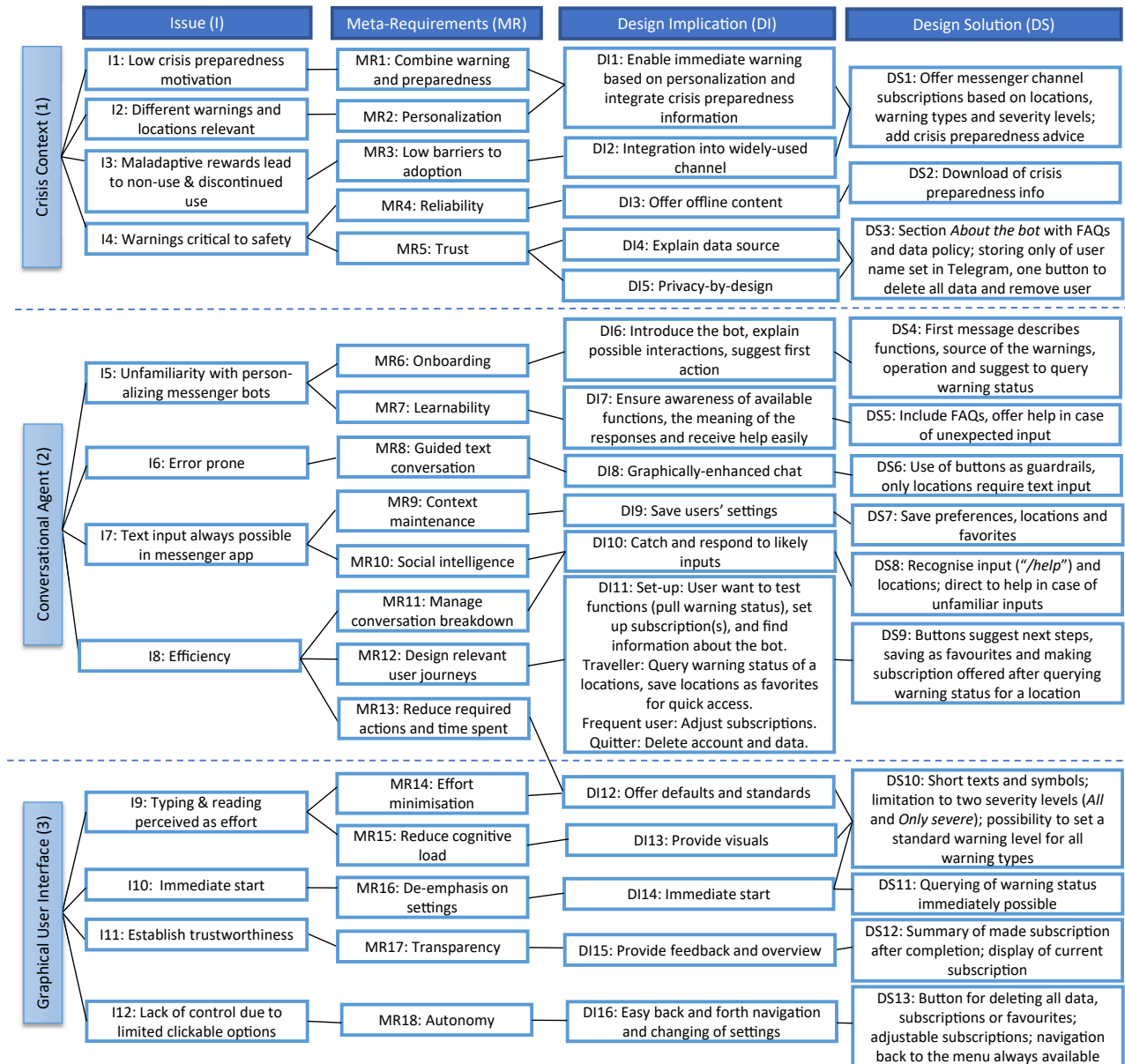


Figure 2: Issues, meta-requirements, design implications and design solutions for the warning messenger channel prototype. Categories inspired by Ref. ²³.

language is not necessary very often, users always have the ability to enter text (I7), requiring a level of social intelligence (MR10). Unfamiliar inputs were, therefore, also dealt with by directing users to the “/help” command.

4 Warning bot implementation

The following will describe how the challenges, requirements and defined objectives are implemented in the bot.

4.1 Crisis context

To ensure reliability and trust, we integrate the API of the official and, in Germany, most widely used warning app *NINA*, which combines information from several relevant safety and security sources, such as the weather service and the modular warning system (MoWaS), the unified warning system of the German civil protection authorities.⁵³ The requirements also implied that an already installed application would reduce the resources and maladaptive reward. In addition, the chosen platform should enable auditory and

visual alerts. For visibility and quick access, it should be able to send push alerts. Due to the requirement of personalisation, users also need to be able to specify and save their preferences through a transparent process. Analysing the rise of conversational interfaces,²⁵ finds that bots within messaging platforms fulfil many of these functions, making them attractive platforms for delivering news.⁹ Additionally, Telegram is widely used as a news outlet in some countries, such as Singapore, with users particularly describing efficiency, convenience and effortless access as key benefits.⁵⁴ IM apps are widely used. Globally, *WhatsApp* has 2 billion monthly users, *WeChat* has 1.3 billion, Facebook 1 billion, *Telegram* 0.9 billion, *SnapChat* 0.8 billion, and *QQ* 0.54 billion.⁵⁵ While *Telegram* has fewer users,⁵⁶ showed that it offers an open-source API⁵⁷ as well as many features that are important for the usability of bots,^{25,38} such as the possibility to create buttons, define quick replies, customise text messages, and publish media for groups or channels for a large number of members (see Table 7 in Appendix C for a comparison of IM apps). Despite the lower usage, we chose *Telegram* as the IM to implement our prototype in, as it offers most relevant affordances and thus allows us to build a solution that is more comparable to a warning app. Within the IM, a bot manages the subscriptions and message delivery based on users' settings.

To implement the system, we selected Python 3 and the `pyTelegramBotAPI` library to facilitate communication with Telegram. We also use the *NINA* API to retrieve relevant information on warnings and disasters. The bot's logic comprises these main tasks: managing the database, generating meaningful responses, providing a user interface, and forwarding slightly edited information from the *NINA* API. Users can proactively request the warning status of locations or set subscriptions to be automatically alerted by the bot (see Figure 3). The bot guides users through multi-step processes, such as creating subscriptions, by providing appropriate prompts. The bot outputs text messages in the *Telegram* chat. The bot's start is managed through the `bot_runner`, which creates

three threads. In the first, the subscription mechanism checks for new warnings every 2 min. In the second thread, the receiver waits for user input in the chat and then calls the appropriate methods in the controller. In the third thread, the `warning_handler` runs. It processes all active warnings upon the initial start of the bot and then scans for new warnings every 2 min by default. The `controller` then accesses various other modules, such as `place_converter`, `nina_service`, `data_service`, `text_templates` and `sender`. The sender then sends the chat message to the user. In the `place_converter`, suggestions for requested cities are generated. The `nina_service` serves as the interface to the *NINA* API, and the `data_service` represents the interface with our database. The appropriate text outputs are created by `text_templates`. For the information architecture, the system employs a database consisting of three JSON files. The first file contains personal information, including the chat ID from Telegram, completed subscriptions, locations saved as favourites, any muted subscriptions and default warning level. The second JSON file stores the postal codes for which active warnings apply, while the third saves which alerts have already been sent to a specific chat ID to prevent multiple warnings.

4.2 Conversational agent

Figure 4 describes the expected user inputs and resulting states based on the results of the workshop. It shows that warnings can be queried by typing the name or postal code of a location or submitting the GPS location. The query can be turned into a subscription for the queried location, or it can be saved as a favourite for future quick access to the location's warning status. In addition, subscriptions can also be made for specific or all warning types, and all or only severe warnings. The subscriptions can be managed or muted. In addition, data can be managed and deleted and information about the bot can be found. Finally, emergency tips can be downloaded.

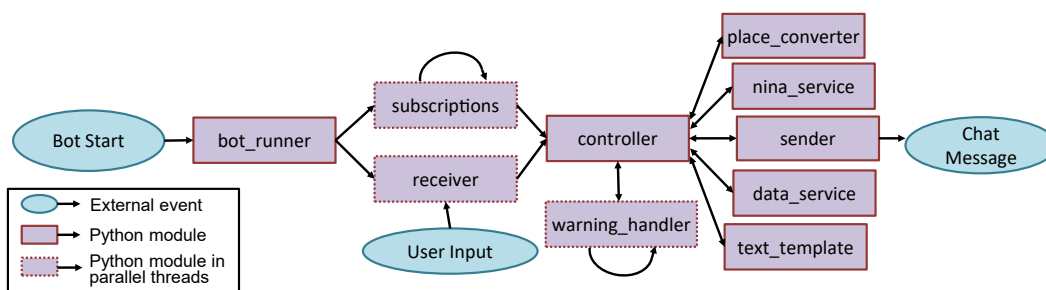


Figure 3: Model of the bot's functional logic.

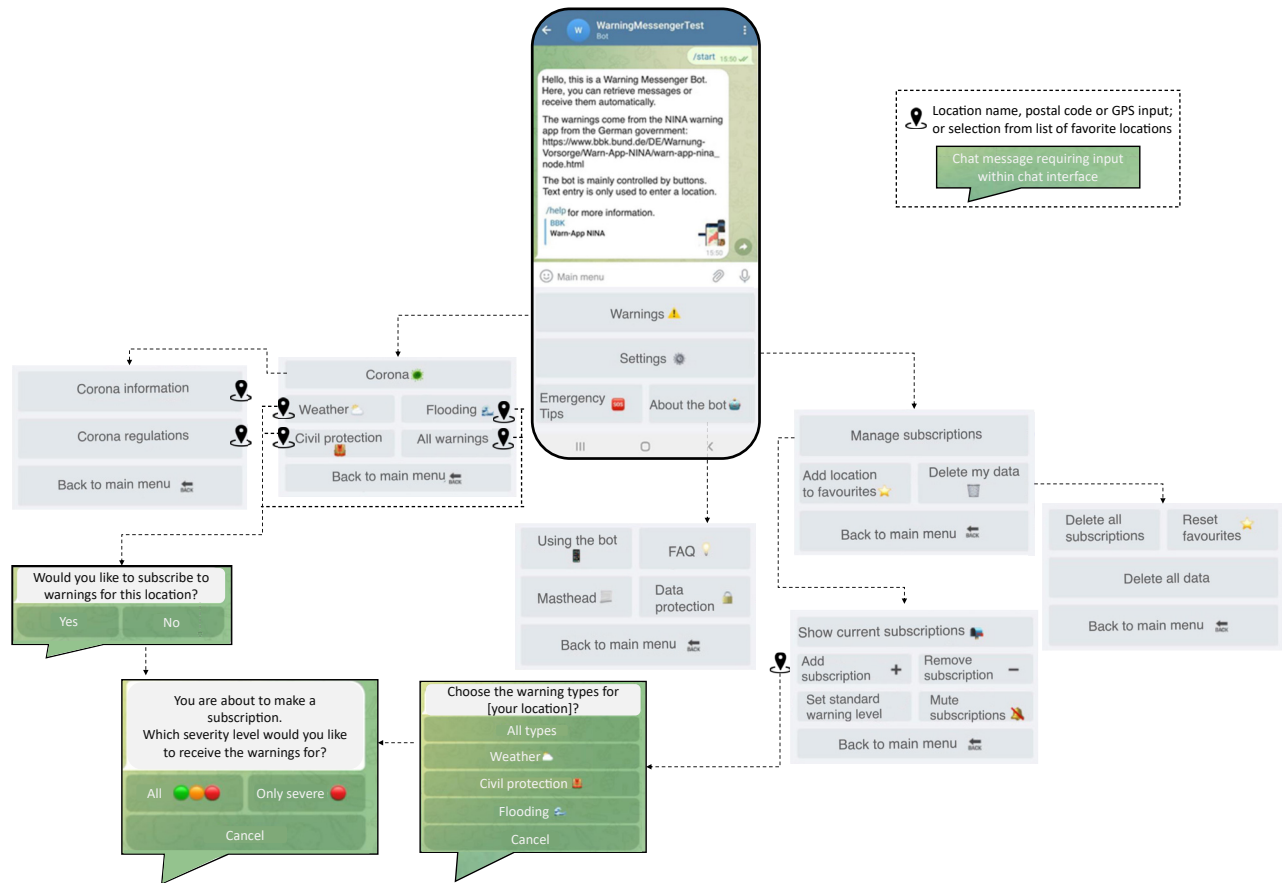


Figure 4: Model of guardrails used to suggest further actions, with button inputs expected for navigation, and text, GPS or postal code for locations.

4.3 Graphical user interface

While offering wide usage and many useful features, some design options within *Telegram* are limited. The messages have a fixed design, which can be personalised in the app's settings. The notification sounds can be set individually within the app and specified for different chats, including by uploading ringtones. Contrary to the requirements of warning apps, the latest messages appear at the bottom of the application and not at the top. Each message is automatically timestamped, and the day is also shown. Regarding the text, some options are available to format the text and highlight keywords via the API. Colouring the text is not possible, but messages can be highlighted by using visualisations such as coloured images or emojis. The official icons of the *NINA* API⁵³ are used to represent the event code. The warning message contains the title and a description of the warning, including the affected areas. If available, the alert also includes the date and time, severity, and version. At the end of the message, there is a link for more information that leads to the source of the message on the website of

the German Federal Office for Civil Protection. When the warnings are transmitted by another agency, such as the weather service, no link is available. After the final round of evaluation and implementation of resulting changes, the bot offered the main features depicted in Figure 5.

5 Demonstration and user evaluation

Following the principles of human centered design, we conducted two rounds of user evaluation, each leading to the implementation of changes to improve our solution. Our iterations focus on the set-up process, as this provides targeted insights into central usability aspects, such as learnability, which were goals that we aimed to achieve with our prototype. In addition, technology adoption is complex and, ideally, requires several instances of data collection, which allows incorporating aspects such as the impact of action on judgement or habit.⁵⁸ Such a longitudinal study was not conducted due to ethical concerns: we could not ensure the

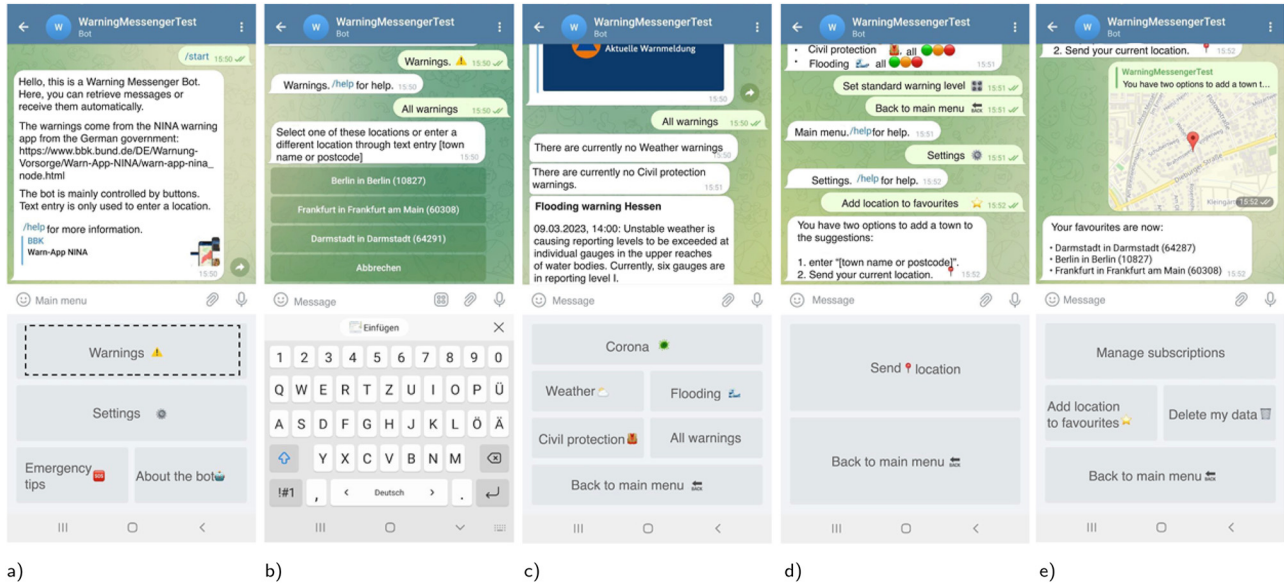


Figure 5: Screenshots of the design and key features of the warning bot, including pulling the warning status for locations and receiving push messages with warnings for saved locations: (a) starting the bot; (b + c) requesting and receiving the warning status for a city or GPS location, (d) saving the city or GPS location into favourites for frequent future warning pulls, (e) setting up locations for immediate warning updates as push messages.

reliable delivery of warnings through the prototype and thus wanted to avoid a false sense of safety in participants. While this means that we cannot judge the usefulness of the warnings in a real setting, the focus on the set-up is justified because it is the instance that requires the most intense user interaction with the bot. By contrast, at a later stage, users are expected to mainly receive alerts passively based on their subscriptions and to make only minor adjustments to their subscriptions. Participants had the right to leave the evaluation at any point.

For the evaluation, we first conducted interviews ($N = 14$) while observing users who fulfilled a given task related to setting up the bot. Interviewers noted users' comments on the tasks and asked follow-up questions where feasible. Finally, users evaluated the systems' usability through a short SUS questionnaire. The second study consists of an online survey ($N = 26$), in which users reported their experiences of performing similar tasks, using the bot on their own devices. Some questions were added to receive feedback on aspects that had been changed after the first evaluation, such as the inclusion of emojis. The survey also focused more on aspects of utility in addition to usability.

5.1 Sample

The sample of the first evaluation study was based on the researchers' personal network, with an attempt to recruit a diverse group. Participants' age, gender, and their affinity

Table 1: Research sample details, where N represents the sample size and ATI stands for affinity for technology interaction.

Eval.	Recruitment	N	N /Age	Gender	ATI M (SD)
1	Personal network	14	< 20: 3 20 – 29: 6 30 – 39: 1 40 – 49: 2 50 – 59: 2	5 women, 9 men	3.44 (1.36)
2	Online ads	26	< 20: 3 20 – 29: 17 30 – 39: 5	15 women, 10 men, 1 diverse	3.94 (0.63)

for technology interaction (assessed via the ATI, 9 items measured on a 6-point Likert scale,⁵⁹ German version⁶⁰) are presented in Table 1. It shows that sample 1 represents a large range of age groups and levels of affinity for technology. In the second sample, possibly due to it being an online survey, the sample was younger – with no one older than 39 years – and the median ATI was higher and less varied. While in sample 1, the average ATI score is somewhat smaller than the mean of $M = 3.61$ found in the German population, that of sample 2 is somewhat higher.⁵⁹

5.2 Measurement of utility and perceived usefulness

To test the success rate of users performing specific functions with the bot, in both evaluations, users were asked

Table 2: Evaluations and tasks used for evaluating the warning bot.

Task	Evaluation	Description
1	1 & 2	Add the warning channel in Telegram.
2	1 & 2	Check the warning status and add a subscription.
3	1	Add your location to favourites and set up a subscription.
4	2	Check for a warning in the location “Testcity”.
	1	View your personal subscriptions.
5	2	Add your current location to your favourites and set up a subscription.
	1 & 2	Find and read the privacy policy, then delete your data from the bot.
6	2	Find and read the bot’s crisis preparedness information.

to perform a set of tasks (see Table 2). The selected tasks represented the main functions of the bot and expected challenges in the set-up process. They ensured that users had experienced the available functions before evaluating usability and usefulness. In addition, the tasks can reveal challenges related to specific interactions.

Several approaches exist for measuring whether the artifact is meeting users’ needs and offers a positive user experience. The most commonly used aspect of user experience is usability, defined as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [61, p. 4]. The most commonly used standardised measurement of usability in research and practice is the 10-item System Usability Scale (SUS).^[62–64] By exploring usability, we can judge whether bot-driven personalisation within messenger channels, despite the existing UI limitations, can be designed in a way that allows for a positive set-up experience.

However, research indicates that usability explains only 39 % of users’ recommendations of a system,^[65] implying that additional aspects are important. The Technology Acceptance Model (TAM) suggests perceived usefulness as the second main factor.^[66] It is necessary to account for as it describes whether an artifact fulfils a need related to users’ contexts and environments.^[67,68] In this study, perceived usefulness is measured through elements of the User Engagement Scale (UES),^[69] which consists of the dimensions *reward*, *perceived usability*, *aesthetic appeal*, and *focused attention*.^[70] The scale offers a good addition to the SUS by focusing on aspects beyond usability.^[70] For example, *reward* includes the perception that the experience was

worthwhile, or the willingness to recommend an application, which is related to perceived usefulness (see Q25 in the Appendix B). Since user experience has some relevance in safety-critical systems but may be less important than usability,^[71] we put usefulness and usability at the centre of our inquiry, including more items on perceived usability (4 items) and reward factors (5 items) and fewer covering aesthetic appeal (3 items) and focused attention (2 items). In addition, we included open questions concerning the reasoning for (not) intending to use the system, as well as perceived advantages and disadvantages compared with other mobile warning channels. Through open questions, we also assessed participants’ feedback about tasks and the artifact generally, as well as their reasoning for their adoption intention and perceived advantages and disadvantages compared to other warning channels. Two quality control questions were added to ensure that the participants were really completing the tasks. The full questions and tasks can be found in the Appendices A and B.

5.3 Data analysis and results

We followed the standard for measuring ATI by offering a 6-point scale, all other scales are 5-point Likert scales. In addition, we used thematic analysis to analyse the qualitative answers^[72] and deduce any challenges related to handling the bot, its structure, visual interface and conversational interface. To analyse the SUS, following common practice, the negatively framed items are inverted, and the data is normalised to represent a scale from 0 to 100 for the whole score consisting of 10 questions, as this allows for a better comparison of the results with other systems.^[63] The SUS results are analysed with non-parametric tests, as the QQ-plot indicates a violation of the assumption of normal distribution, which is corroborated by the positive skewness of the histogram. Despite the Shapiro-Wilk test resulting in a significance of $p < 0.01$, we decided to use non-parametric tests due to the relatively small sample size, which limits the robustness of analyses against non-normally distributed data.

The Likert-scale data is thus analysed using Kendall’s tau-b coefficient, which is robust to skewed data and suitable for analysing ordinal data on 5-point Likert scales, as it accounts for tied ranks, which are common in such data, and provides a measure of association that aligns with the ordinal nature of Likert responses.^[73] The qualitative answers were coded independently by two researchers through thematic analysis,^[72] with the coding scheme developed inductively from the text.

5.3.1 Interview evaluation

In the interview, users interacted with the bot on the researchers' devices. They were encouraged to act as they would using their own devices and to try to use the bot on their own. However, in case they could not perform the task on their own, they could seek out the help of the interviewer who was monitoring the process and taking notes about any unexpected behaviour or user comments. The task-based interviews showed that three out of $N = 14$ participants needed support in order to accomplish the first task, one person needed assistance with task 3. Even the people with a low ATI score felt very or rather well-informed about the functions ($M = 4.43, SD = 0.51$) and operation of the bot ($M = 4.43, SD = 0.65$). Users predominantly stated that the tasks were very easy or easy to accomplish. In their qualitative answers, they showed a good understanding of the core functions. The overall usability rating, using the common 10-point scale for each item, is 90, with all items rated above 8, indicating an excellent usability,⁶³ which, however, can only be used as a rough indication as participants had the assistance of the interviewers. Asked about the likelihood of using a similar tool in the future, the mean rating was $M = 3.5$ ($SD = 1.02$). When asked specifically whether they would use this tool, just over one third ($n = 5$) stated that they would. These prospective users were all in the two youngest age categories (<29 years) but had varying experiences with warning apps, ranging from not being aware of their existence to being current users. The logistic regression analysis – to be taken with caution due to the small sample size – revealed that neither participants' mean ATI ($p = 0.83$) nor SUS ($p = 0.19$) significantly predicted the binary outcome of planned future use, indicating the relevance of other factors. Looking at users' reasoning,

arguments against using the tool were mainly disinterest in warnings generally ($n = 4$), lack of need due to use of a warning app ($n = 2$), not using *Telegram* ($n = 2$) or preferring the design of an app rather than the messenger ($n = 1$). Positive reasoning related to the advantage of receiving automatic warnings was not requiring a separate app and the intuitive design. With regard to design improvements, users requested shorter texts by using symbols and emojis, and a short confirmation upon successfully setting up a warning subscription. For the next iteration, we made the suggested changes.

5.3.2 Online survey evaluation

The survey confirmed that all tasks are rather or very easy to accomplish ($M = 4.32$). The first task of finding and starting the bot was perceived as the most difficult ($M = 4.08$). Only one person failed to retrieve the warning status for a given and current location. The mean usability rating, again using the 10-point scale, is 76.88 ($SD = 17.58$), indicating good to excellent usability.⁶³ Figure 6 depicts the evaluation of each item on a 5-point scale for easier comparison with the other evaluations. Aspects that received a particularly positive rating concerned being able to use the system right away and without needing expertise or training. Due to the conversational design, which results in previous inputs becoming less visible, we introduced an additional item asking about the transparency of the system's status, resulting in a relatively good rating ($M = 3.88, SD = 1.10$).

Regarding privacy, users perceived that the bot's handling of private information seemed trustworthy ($M = 3.69, SD = 1.30$), that they did not have qualms using the bot with regard to data security ($M = 3.73, SD = 1.31$) or privacy ($M = 3.73, SD = 1.06$). The download of the official

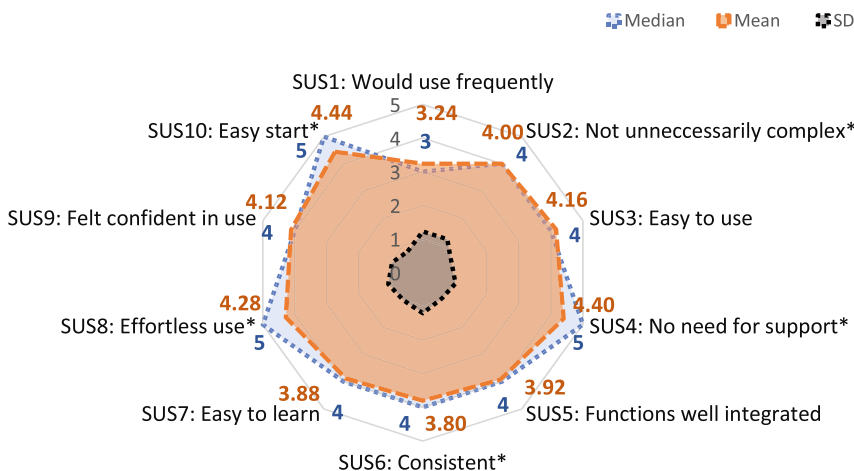


Figure 6: Mean (orange, middle line), median (blue, outer line) and standard deviation (black, inner line) of each SUS item's 5-point scale rating, *inverted items.

Table 3: Usage and recommendation intentions, rated on a scale from 1 (fully disagree) to 5 (fully agree).

Intention to	M	SD	Fully agree	Agree or fully agree
...use frequently	3.25	1.22	20 %	44 %
...use on favourite messenger				
...for subscriptions	3.83	1.37	40 %	68 %
...for inquiries	3.25	1.15	12 %	44 %
...recommend system	3.50	1.10	20 %	48 %

prevention brochure as crisis prevention advice was perceived as helpful by all but two participants ($M = 3.88$, $SD = 1.41$). Most users appreciated the offline availability, although they would have preferred a more concise summary (the used one has seven pages). The new design iteration included emojis to denote the warning topics and severity. These were largely seen as helpful ($M = 3.88$, $SD = 1.08$). Asking about improvements, 10 users made comments which related to small design improvements (such as in the way the names of locations were presented) and to a technical issue that resulted in longer loading times (which was fixed in the course of the study).

Regarding the adoption intention, 20 % fully agreed that they would like to use the bot frequently and almost half rather or fully agreed. If the bot were available on users' favourite messenger app, 20 % more planned on using the bot to make subscriptions. Fewer people would use the bot to query the warning status of specific locations. Finally, another important indicator of the overall usefulness of a system is the likelihood of recommending a system to friends or family. The results are depicted in Table 3 and show that almost half of the participants would recommend the bot to friends and family.

In order to explain these preferences, we correlated the intentions to use the system with the individual technology affinity, the usability rating and the dimensions of the user engagement rating through a Kendall tau-b test. The results of the correlation analysis are depicted in Table 4. It shows

that the intention to use the bot (item 1 of the SUS) correlates with the ATI, SUS and the reward dimension of the user engagement scale. It also correlates with the intention to use the bot in the future if available on users' favourite messenger app.

People who would be willing to use the bot for its core function, warning subscriptions, rate it high in the reward dimension and would recommend the bot. People who would use the bot to inquire about the warning status, a secondary function introduced mainly to familiarise users with the bot, does not correlate with the perceived usefulness or the intention to recommend the bot. However, it correlated with technology affinity, further stressing that this is a feature for those who like to engage with technology. Affinity for technology, usability, and all dimensions of the user engagement scale (UES Perceived Usability $\tau_b = 0.410$, $p = 0.005$, UES Aesthetic Appeal $\tau_b = 0.531$, $p < 0.001$, UES Focused Attention $\tau_b = 0.615$, $p < 0.001$) all affect the intention to recommend the system.

Looking at users' reasoning, those planning to use the bot in the future mentioned the ease of use and not requiring the download of an additional app. People who do not plan on using the bot mentioned not needing a mobile phone to be made aware of severe crises and rather using different sources, such as browser searches, indicating a perceived low need to be informed in a timely manner, or a general disinterest in crisis warnings, e.g.: "I think that more serious matters would somehow reach me, and I wouldn't need a phone for that. I have never thought about it". Another group who are less interested are those people who are already using a warning app, who described little added value and some disadvantages regarding the user interface, particularly the display of the preparedness information.

Asked about the advantages and disadvantages of the bot compared to a warning app, users mentioned specifically the flexibility offered by the bot and not requiring another app compared to warning apps. One participant stated: "If I'm honest, I haven't used any of the apps yet because I didn't want to install another app and familiarise

Table 4: Kendall tau-b correlation coefficients between usage intentions and evaluations.

Intention to	...use for subscriptions	...use for inquiries	...recommend	ATI	SUS	UES	UES RW
...use frequently	0.366*	0.359*	0.233	0.305*	0.356*	0.131	0.404*
...use on favourite messenger							
...for subscriptions	1	0.285	0.412*	0.259	0.297	0.351*	0.470**
...for inquiries		1	0.139	0.368*	0.116	0.236	0.251
...recommend			1	0.400*	0.510**	0.664**	0.724**

ATI = affinity for technology interaction, SUS = system usability scale, UES = user engagement scale, UES RW = reward dimension of the UES. * $p < 0.05$; ** $p < 0.01$.

myself with it. Lack of storage space was also a thing a while ago... A Telegram bot is more useful for me than another app” (P24). However, other users mentioned that the bot required a messenger app, particularly Telegram. Another perceived disadvantage is the lack of flexibility with regard to the user interface design in the messenger bot. While the handling of the bot received positive evaluations and was perceived as intuitive, one user mentioned that it might pose a challenge for older people: “[...] if the buttons are gone at the bottom, you have to press on a quite small icon to call them up again” (P16). Another disadvantage of the bot compared to apps or CB is that the latter require less or no set-up effort, but they can operate simply based on the live location, which is not the case for the bot (see Table 5 for the full list of themes).

6 Discussion

In the following, we discuss the implications of the study for the design of warning channels and for bot-driven personalisation.

6.1 Warning “botplications” requirements

Based on the requirements for warning apps, the study has shown that messaging channels offer some, if not all, relevant functions that are central to delivering timely, personalised warnings.

Bonaretti et al.²⁸ analyse warning apps with regard to different layers. Looking at these layers (see Table 6), we find that warning apps and warning messaging channels are identical in that they run on the same physical layer on smartphones. However, if an IM app has already been installed, users do not need to invest in additional storage space, and messages typically require a small amount of

Internet data. In addition, although not widely available, messenger apps can use peer-to-peer communication, making them, to some extent, resilient to Internet outages. On the deep layer, concerning the alerting object, the bot can send the same alerts as the app, due to the availability of an API of the official alerting app. The crucial difference is at the software level, where one uses a warning app and the other an IM app, whereby the choice of messaging app should depend on the number of users. The surface layer, too, is highly dependent on the software used. Generally, options for sound and visuals are limited by the IM apps’ functions and design.

The software layer also influences the provided warning and crisis functions. Regarding the functions, an important difference is that currently, location tracking is not possible within the chosen IM app. Therefore, warnings cannot be made based on the current location of users. However, users can allow short-term access to the location and the bot can use this information to suggest a warning subscription for the respective location. Due to this process, the set-up takes longer than in some warning apps, where the location-based warning is enabled as soon as users grant access to their location information. In addition, the bot cannot send warnings based on the current location, only based on predefined locations of interest. This leads to the bot requiring more set-up than a well-implemented warning app, which in its initial start can ask for the permission to use GPS, enabling warnings with only one click. Another difference is that it is currently not possible to show an interactive map within the IM app. Therefore, it is more difficult to show users that the app is active and to give them some insights into the warning status. We, therefore, implemented a feature that allows users to query the warning status of locations, which offers an immediate way for users to interact with the bot and verify its functionality.

Table 5: Users’ reported advantages and disadvantages (*n*) of the warning bot compared to other warning channels, *N* = 20.

	Advantages	Disadvantages
App requirement	No additional app needed (6)	Requires <i>Telegram</i> app (8)
Interface	Easy to use and to learn (4) Topic engagement through interactive design (1)	More difficult set-up compared with one-click GPS-based warnings in warning apps (2) <ul style="list-style-type: none"> – Less attractive and flexible interface (2) – Better presentation of prevention information (1) Display of maps (1)
Perceived usefulness	Subscriptions for different locations (4) Warning status queries (6)	Disinterest in warnings (4) Use of other channels, such as Internet search (4)
Resilience	–	Less resilient against Internet outages than CB (2)
Privacy	Better data protection (1)	No warnings when in silent mode (1)
		–

Table 6: Layered nature of a warning app, layers based on Ref. 28.

Structure	Warning app	Bot
Deep layer (warning object)		
Data	Send official warnings through MoWaS	Send official warnings via API
Software application	Different apps and features offered, in Germany: <i>NINA</i> , <i>KATWARN</i> , <i>hessenWARN</i>	Popular messengers: <i>WhatsApp</i> , <i>Telegram</i> , <i>Signal</i>
Surface layer (notification)		
Sound	Individually adjustable; can override silent mode, often background activity deactivated due to infrequent use	Can be set for the “channels”, but no specific override permissions
Visual	Can be freely designed	Limited design options, depending on messenger
Physical layer (hardware and network)		
Device	Requires data capacity and battery, e.g. for GPS tracking	Requires data capacity for messenger app, though often already installed; no additional battery drainage
Internet connection	Often background activity deactivated due to infrequent use	Requires Internet, possibility of peer-to-peer communication

This feature is not common in warning apps but may provide additional benefit for people who enjoy engaging with technology – especially as “site-seeing”, i.e. digitally visiting locations affected by crises, is a common online activity.⁷⁴ Another important function is offering additional advice. However, the conversational interface appears inadequate for displaying longer texts. Therefore, we opted to offer advice through a file that can be downloaded. While users prefer to have all information in the warning app rather than through external links,¹³ this allows saving the file on the device, making it available offline.

With regard to the conversational design, the implementation process has shown that many challenges arise from the onboarding process. This is in line with research on voice-based interfaces for visually impaired people, which finds that discoverability of the available functions is a challenge.⁷⁵ For onboarding, we have chosen to let users query locations’ warning status, as this provides an immediate insight into the capabilities of the bot and the type of output to be expected. In this way, users can experience the bot’s warning function before setting up subscriptions. This interaction is used to invite users to set up warning subscriptions for the queried location, which the study confirms to be the key function of the bot. For the set-up process, the chosen settings need to be transparent. In addition, buttons can serve as guardrails to give guidance about the available options. The evaluation has shown that efficiency is a key element with the warning bot. Efficiency was improved by displaying colours and emojis to give immediate visual cues, reducing time spent reading. In addition, texts were shortened substantially. Providing sufficient information to navigate the bot while keeping text to a minimum is another

challenge that is intensified in the messaging platform. Users expressly stated that the bot was easy to use. This indicates that the recommendation of buttons to guide the conversation²⁵ can be confirmed in this context.

While many studies indicate that human likeness increases positive evaluations of bots, this was not mentioned by users in our study. In a study of a municipal service bot, the relevance of human likeness was regarded very differently between users. This may be due to the character of the bot. While the municipal service navigation bot used natural language to support users, the bot developed here is closely related to an application. This supports the differentiation of botplication from other types of conversational bots, with botplings aiming for simplicity and effectiveness, whereas human likeness is irrelevant or secondary.²⁵ The analysis, discussion, and evaluation of chatbots’ maturity is mainly based on the use of natural language processing. A typology by⁷⁶ differentiated between first level bots that facilitate service triage (matching user request with rules and linked data sources); second level bots for service information gathering and analysis that access data from multiple sources to personalise the service; and third level bots for service negotiation, which entail predictive analytics and neural networks to evaluate customer circumstances based on various sources such as user profiles and service profiles and providing multiple personalised service options for the customer in their natural language. However, the task-oriented botplication described in this study does not fit into the typology. For example, in the typology, accessing user profiles or continuing service interaction occurs along with NLP components such as sentiment analysis. The bot used here shows that personalised service delivery does

not require large language processing capabilities. Indeed, the design recommendations for CUIs within chatbots suggest a limited use of text and an extensive use of graphical UI elements to steer the conversation in an efficient manner.²⁵

6.2 The role of usability and perceived usefulness for intended future use

Looking at the utility of the warning bot, we find that 68 % of study participants stated that they would use the system if it was provided through their favourite IM app. In addition, almost half stated that they would recommend the bot.

The reward dimension of the user engagement scale describes aspects such as whether the time spent with the system was worthwhile and emerged as the most significant, impacting also the intentions to use the bot generally and specifically for warning subscriptions. The intention to recommend the warning bot only correlates with the intention to use it for warning subscriptions in a preferred IM app, indicating that this aspect is perceived as the most useful feature for other people. Thus, this underlines the added value of bots when coupled with proactive notifications, as this is not achieved through other sources, such as Internet searches.

The study suggests that usability, in the context of warning channels, is less relevant for predicting the intention to adopt a system. Instead, the intention to use the bot is more closely related to the utility of a system – aspects that are described in the reward dimension of the user engagement scale.⁷⁰ This supports the value of the user engagement scale and its different dimensions, which has also proven useful in the health and news domains.^{77–79} Therefore, future representative user studies that evaluate chatbot designs or compare them to alternative warning channels should include such aspects in addition to usability aspects. This result is in line with previous findings on e-government non-adoption. One study found that users preferred to remain with their status quo, interacting with humans: “The less need citizens feel for public services or the more convenient they find conventional ways of contacting administrations, the less they are inclined to use public e-services”.⁸⁰ Adapted to the use of warning apps, we also find that many respondents, even when indicating that the tool is useful and has high usability, state that they would continue using public media for warnings. Similar to another qualitative study on e-government, we find that a perceived lack of need or infrequent usage is one of the main barriers.⁸¹ In the open statements, it also became evident that, at least for some of the participants, it had not become clear that the subscriptions would automatically alert them in case of a warning. As a result, two participants stated that they might

not be querying information just as a crisis was impending, likely due to the unfamiliarity with bot-based subscriptions. Still, more attention should be paid to ensuring that users understand the scope of the offered function, e.g. through a short introductory video.^{82,83}

In this study, we sought to investigate the usability of a warning messaging bot as an alternative warning channel to warning apps. Therefore, we attempted to emulate features that are offered by the official German warning apps. However, warning apps could conceivably offer additional safety and security functions or include more accessibility features, such as haptic instead of auditory signals or video emergency calls for people with hearing impairments. The structure of the conversational interface introduced in this study can be used to design a fully voice-based process for people with impaired vision, allowing them to query information through natural speech or to set up automatic warning subscriptions. A study shows that conversational assistants provide a range of useful functions for blind people.⁷⁵ With added natural language processing functions, the developed conversational design could be used to make personalised warnings accessible to further user groups, who tend to prefer to be able to use mainstream channels rather than specifically designed ones.⁸⁴ However, moving towards accessibility would require extensive user studies and changes in the way that information and alert signals are presented.⁸⁵

While the additional steps needed to set up a subscription may be perceived as a disadvantage, one user reported that the process increased engagement with crisis management, possibly by confronting users with the available selection of warning types. This finding may echo an increased engagement with municipal services that was achieved through the use of a chatbot.⁵¹

6.3 Limitations and future work

The paper has several limitations regarding the interactions that were analysed and the design of the warning bot. Firstly, due to ethical concerns related to ensuring the reliability of the bot “in the wild”, the evaluation did not encompass real-world use, restricting the measurement to participants’ experiences in the set-up and their intentions for future use. Future work should create an experimental setting that allows the usability of the alerts to be tested. One question addressed the advantages of a messenger bot compared to other warning channels. This question was answered by both current users and non-users of warning apps. Future work should explore the affordances of a messenger warning bot with the affordances of warning apps and cell broadcast, similar to a study which has analysed

affordances of news consumption via messenger bots.⁵⁴ In addition, the evaluation is biased towards younger users, who appear to be more willing to use the bot in the future and who are likely to be experienced messenger app users.⁸⁶

With regard to using messenger apps as the containers of bots, we find that some messenger apps offer significant features for open source, low-cost implementations. However, a disadvantage is that certain aspects cannot be adapted. In this prototype, a significant usability limitation is that the text input fields cannot be hidden or disabled. This suggests to users that text inputs may be required. In addition, activation of the text input field can lead to the disappearing of the buttons, requiring some familiarity to recuperate the buttons. In addition, other research has also explored the relevance of humanness of bots with inconclusive results and suggested a dependence on the task and domain⁸⁷. Inspired by the usability requirement of simplicity, we did not implement any humanness in the bot in the warning context. Future research could explore the relevance of this attribute through a comparative experiment.

Regarding the conversational design of the bot, due to the importance of achieving correct results in a safety-critical domain, the current prototype does not use natural language processing (NLP). While this does justice to the requirement of minimal input, especially for mobile use, and reduces errors, NLP could enable users to interact more freely with the bot like⁸⁸ have put forth. However, our system's backend, functions and conversational design structure could be used for future accessibility studies, where the messenger or smart speakers may be used to set up the warning bot preferences and where a smart home environment may be used for visual, haptic or auditory signals adapted to users' accessibility needs. This could enable users with visual and hearing impairments to set up and receive noticeable alerts. A previous study has already suggested different signals and intensities to communicate varying warning levels.⁸⁹ The use of natural language for output would also allow conveying more extensive information regarding crisis preparedness. Using NLP, the bot could answer specific questions related to crisis preparedness upon request. Previous research suggests that such information should include statistical information that stresses the benefits of crisis preparedness.⁹⁰

7 Conclusions

The use of *botplications*, i.e. bots that fulfil the tasks currently carried out by apps, holds the promise of lowering adoption barriers by not requiring the download of specific apps for specific functions. Therefore, this study explores

the use of a bot run within an IM app as an alternative to warning apps. However, the design of messaging bots, particularly for conversational interfaces, presents distinct challenges and necessitates specific considerations, as these interfaces are characterised by sequential interactions.

Using the design science research approach, this study investigated how functions and usability requirements of warning apps can be transferred to a messaging warning channel. A comprehensive set of requirements for graphically enhanced CUIs in safety-critical contexts was developed. After designing and implementing a warning bot using the API of an official warning app, two rounds of user evaluation provided valuable insights into the bot's usability and user perceptions. The usability analysis shows a positive evaluation, indicating that user-friendly bot-based personalisation, while currently still uncommon in news delivery, can be accomplished. The task-based interviews indicated positive user feedback in general, with only a small percentage requiring some support. Suggestions for design improvements included shorter texts, increased use of symbols and emojis, and the addition of short affirmative messages upon successful actions. These insights guided refinements for the subsequent iteration. Despite the generally high usability rating, the intention to use the system in the future varied. The second round of evaluation, which explored additional elements related to user engagement, showed that particularly elements related to perceived usefulness were relevant to the intention to keep using the system and to recommend it. This indicative finding underlines the relevance of perceived usefulness in addition to usability for technology adoption and should be further verified for warning bots in future studies. The qualitative findings show that a general disinterest in crisis warnings and a low perceived personal responsibility for crisis engagement are the main factors for not wanting to use the tool.

A second strand of reasoning is related to the use of other warning channels. The warning bot prototype and user evaluation indicate that reliable warnings tailored to personal preferences can be provided through IM apps as a viable addition to warning apps or cell broadcast. Though current IM apps offer clear technical constraints like the unavailability of live location tracking or having to adjust auditory alert levels manually. This limits the promise of warning bots and points to the app developers, who could introduce warning-related functions and thus ease the bot set-up process and lower adoption barriers.

In summary, the research demonstrates the significance of tailored conversational design for messaging bots in disaster contexts. The iterative design process, supported by user evaluation, led to a refined interface capable of

efficiently delivering trustworthy crisis warnings. The final graphically enhanced CUI shows that a suitable warning bot does not necessarily require complex NLP capabilities. The study hence contributes valuable insights for the development of effective and user-friendly messaging bots in safety-critical applications. Future research should compare the use of a warning bot and a warning app “in the wild” to expand the research focus from set-up to warning perception. The available prototype can furthermore be used for conversational personalisation of warning channels that can increase the accessibility of mobile warnings or warnings in a smart home context.

Acknowledgments: We thank all study participants for their time and insights.

Research ethics: Not applicable.

Informed consent: Informed consent was obtained from all individuals included in this study, or their legal guardians or wards.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Use of Large Language Models, AI and Machine Learning Tools: None declared.

Conflict of interest: The authors state no conflict of interest.

Research funding: This research was funded by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE and by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72].

Data availability: Not applicable.

Appendix A: Questions and tasks of interview evaluation

- Q1: **Enter the number you will receive from the interviewer here** [Number]
 - Q2: **How old are you?** (12–19; 20–29; 30–39; 40–49; 50–59; 60–69; 70–79; not specified)
 - Q3: **Which gender do you identify with?** (Female; Male; Other; Not specified)
 - Q4: **The following questions are about your interaction with technical systems. By ‘technical systems’ we mean apps and other software applications as well as complete digital devices (e.g. cell phone, computer, TV, car navigation).** Please select the answer that applies: (Strongly disagree; Rather disagree; Moderately disagree; Rather agree; Strongly agree)
 - I like to occupy myself in greater detail with technical systems.
 - I like testing the functions of new technical systems.
 - I predominantly deal with technical systems because I have to.
 - When I have a new technical system in front of me, I try it out intensively.
 - I enjoy spending time becoming acquainted with a new technical system.
 - It is enough for me that a technical system works, I don’t care how or why.
 - I try to understand how a technical system exactly works.
 - It is enough for me to know the basic functions of a technical system.
 - I try to make full use of the capabilities of a technical system.
- Task 1: Add the bot to the channels in Telegram.
- Q5: **Did you complete the task?** (Yes; Yes, with help; No)
 - Q6: **How well do you feel informed about:** (1 (very good) – 5 (very poor))
 - the function of the bot?
 - the operation of the bot?
 - Q7: **Describe briefly what the purpose of the bot is.** [Comment]
 - Q8: **Do you have any further comments/suggestions for improvement?** (misleading terms/descriptions, spelling, menu design...) [Comment]
- Task 2: Use the bot to check the warning status and add a subscription.
- (The type of warning and the name of the place will be given to you by the interviewer).
- Q9: **Did you complete the task?** (Yes; Yes, with help; No)
 - Q10: **Explain to the interviewer what settings you have just made.** [Comment]
 - Q11: **How informative do you find the alerts?** (1 (very informative) – 5 (not informative at all))
 - Q12: **How easy was it to find the functions?** (1 (very easy) – 5 (very difficult))
 - Q13: **Do you have any further comments/suggestions for improvement?** (unclear terms/descriptions, spelling, menu design...) [Comment]
 - Q14: **If the/help function was used, was it helpful?** [Comment]

Task 3: Add location to favourites and receive warnings accordingly.

- Q15: **Did you complete the task?** (Yes; Yes, with help; No)

Favourites allow you to quickly specify the desired location. For example, you don't have to enter "CITYNAME" in the text field for your next warning query, but the city will be suggested to you directly. Currently there are three possible city suggestions that you can personalise.

- Q16: **How many favourites do you find useful?** (1; 2; 3; 4; 5; 6; 7; 8)
- Q17: **How easy was it to find the functions?** (1 (very easy) – 5 (very difficult))
- Q18: **How would you rate the difficulty of the task?** (1 (very easy) – 5 (very difficult))
- Q19: **Do you have any further comments/suggestions for improvement?** (unclear terms/descriptions, spelling, menu design...) [Comment]
- Q20: **Was the/help function used?** (Yes; No)
- Q21: **If the/help function was used, was it helpful?** [Comment]

Task 4: View your personal subscriptions.

- Q22: **Did you complete the task?** (Yes; Yes, with help; No)
- Q23: **How transparent do you find the app in terms of knowing your subscription status?** (Very transparent; Rather transparent; Neither; Rather not transparent; Not transparent at all)
- Q24: **How easy was it to find the functions?** (1 (very easy) – 5 (very difficult))
- Q25: **How would you rate the difficulty of the task?** (1 (very easy) – 5 (very difficult))
- Q26: **Do you have any further comments/suggestions for improvement?** (unclear terms/descriptions, spelling, menu design...) [Comment]
- Q27: **Was the/help function used?** (Yes; No)
- Q28: **If the/help function was used, was it helpful?** [Comment]

Task 5: Query the privacy information of the channel and delete your data.

- Q29: **Did you complete the task?** (Yes; Yes, with help; No)
- Q30: **How trustworthy do you find the data collection and data processing within the bot?** (1 (very trustworthy) – 5 (not at all trustworthy))
- Q31: **Are there any privacy-related issues that would prevent you from using the bot?** [Comment]

- Q32: **How satisfied are you with the bot's privacy settings?** (1 (very satisfied) – 5 (not at all satisfied))
- Q33: **How easy was it to locate the features?** (1 (very easy) – 5 (very difficult))
- Q34: **How would you rate the difficulty of the task?** (1 (very easy) – 5 (very difficult))
- Q35: **Do you have any other comments/suggestions for improvement?** (unclear terms/descriptions, spelling, menu design...) [Comment]
- Q36: **Was the/help function used?** (Yes; No)
- Q37: **If the/help function was used, was it helpful?** [Comment]

Independent of specific tasks:

- Q38: **How long did you spend using the bot outside of the previous tasks?** (Less than 5 min; 5–10 min; 10–15 min; 15–20 min; More than 20 min)
- Q39: **How much would you want to use the bot?** (1 (very much) – 5 (not at all))
- Q40: **Please explain why?** [Comment]
- Q41: **When answering these questions, please refer to the entire time you have been involved with the bot.** (1 (strongly agree) – 5 (strongly disagree))
 - I find the bot unnecessarily complex.
 - I think the bot is easy to use.
 - I think that I would need the support of a technical person to be able to use this bot.
 - I found the various functions of this bot were well integrated.
 - I thought there was too much inconsistency in this bot.
 - I would imagine that most people would learn to use this bot very quickly.
 - I found the bot very cumbersome to use.
 - I felt very confident using the bot.
 - I needed to learn a lot of things before I could get going with the bot.
- Q42: **If you want to justify your choice in the previous multiple choice questions, you can do so here.** Improving our bot will be easier if you leave us concrete praise or criticism! [Open question]
- Q43: **Are you currently using or have you used a disaster warning app in the past, like NINA or KatWarn?** (I did not know about such apps until now/I know them, but have never used them/I currently use such an app/I have used such an app in the past, but not anymore)
- Q44: **Do you have any other comments/suggestions for improvement?** (unclear terms/descriptions, spelling, menu design...) [Comment]

- Q45: **Last but not least: Would you use this bot? Why?** (Yes, because... [Comment]; No, because... [Comment])
- Q46: **Do you have any feedback about the study?** [Comment]

Appendix B: Questions of online survey evaluation

- Q1: **Please enter your age group.** (12–19; 20–29; 30–39; 40–49; 50–59; 60–69; 70–79; not specified)
- Q2: **Which gender do you identify with?** (Female; Male; Other; Prefer not to say)
- Q3: **The following questions are about your interaction with technical systems. By “technical systems” we mean apps and other software applications as well as complete digital devices (e.g. cell phone, computer, TV, car navigation).**
Please select the answer that applies: (Strongly disagree; rather disagree; moderately disagree; rather agree; strongly agree)
 - I like to occupy myself in greater detail with technical systems.
 - I like testing the functions of new technical systems.
 - I predominantly deal with technical systems because I have to.
 - When I have a new technical system in front of me, I try it out intensively.
 - I enjoy spending time becoming acquainted with a new technical system.
 - It is enough for me that a technical system works, I don't care how or why.
 - I try to understand how a technical system exactly works.
 - It is enough for me to know the basic functions of a technical system.
 - I try to make full use of the capabilities of a technical system.

Task 1: Find and start bot.

- Q4: **How easy or difficult did you find it to find and start the bot?** (1 (very easy) – 5 (very difficult); I could not complete the task)
- Q5: **What do you think of the length of the welcome text? Do you have any further comments/suggestions for improvement?** (misleading terms/descriptions, spelling, menu design...) (Too long; just right; too short; [Comment])

- Q6: **After reading the introductory text, how well do you feel informed about ...** (1 (very good) – 5 (very bad))
 - the function of the bot?
 - the operation of the bot?

Task 2: Check for a current warning and set up automatic alerts for your location.

- Q7: **How easy or difficult did you find it to...** (1 (very easy) – 5 (very difficult); I could not complete the task)
 - retrieve the warnings for your place of residence?
 - subscribe to alerts for your place of residence?
- Q8: **This question is about the presentation of the text at the step “You are about to add a subscription. What alerts do you want to receive?”** Please select the answer that applies: (1 (I strongly agree) – 5 (I strongly disagree))
 - The answer options are intuitive.
 - The number of possible answers is sufficient.
 - I find the pictures next to “all warnings” appropriate.
 - The pictures next to “only serious warnings” are appropriate.
 - You can get to the questioned step via “Add subscription”.

Task 3: Check for a warning for the location “Testhausen”.

- Q9: **How easy or difficult did you find it to retrieve the warning status for the given location?** Please select the answer that applies: (1 (very easy) – 5 (very difficult); I could not complete the task)
- Q10: **How informative do you find the warnings?** Please select the answer that applies: (1 (very informative) – 5 (not informative at all))
- Q11: **Is there anything you would improve about the presentation of the warnings?** [Comment]
- Q12: **What do you think of the images used? If you find any of the images inappropriate, please let us know in the comments box.** (1 (very inappropriate) – 5 (very appropriate); [Comment])
- Q13: **What was the warning level of the warning you received?** (Extreme ●●●; Moderate ●●; All ●●●; Medium ●●; Extreme ●●●; Serious only ●●; Serious ●)

Task 4: Add your current location to your favourites and get according warnings.

- Q14: **How easy or difficult did you find it to retrieve the alerts for your current location?** Please select the

answer that applies: (1 (very easy) – 5 (very difficult); I could not complete the task)

Task 5: Find and read the privacy policy, then delete your data from the bot.

- Q15: **How easy or difficult did you find it to retrieve the privacy information?** Please select the answer that applies: (1 (very easy) – 5 (very difficult); I could not complete the task)
- Q16: **Please indicate how much you agree with the following statements.** (1 (strongly disagree) – 5 (strongly agree); I could not complete the task)
 - The use of personal data in the bot is trustworthy.
 - I would have concerns about data protection when using the bot.
 - I would have privacy concerns when using the bot.

Task 6: Navigate to the bot's emergency information section.

- Q17: **How easy or difficult did you find it to find tips for an emergency and emergency preparedness?** (1 (very easy) – 5 (very difficult))
- Q18: **What do you think about the possibility to download the emergency tips as a PDF?** Please describe briefly what you find helpful or how emergency tips could be improved. (1 (not helpful at all) – 5 (very helpful)); [Comment]
- Q19: **What format do you prefer for the emergency tips?** (Emergency tips as link; Emergency tips as PDF; I would like to have both options available)

Closing questions

- Q20: **I think I would like to use this system regularly.** Please give reasons for your answer. (1 (do not agree at all) – 5 (fully agree)); [Comment]
- Q21: **When answering these questions, please refer to the entire time you have been involved with the bot. Don't think too long about your answers, as what matters to us is your subjective opinion.** (1 (do not agree at all) – 5 (fully agree))
 - I find the bot unnecessarily complex.
 - I think the bot is easy to use.
 - I think that I would need the support of a technical person to be able to use this bot.
 - I found the various functions of this bot were well integrated.
 - I thought there was too much inconsistency in this bot.
 - I would imagine that most people would learn to use this bot very quickly.

- I found the bot very cumbersome to use.
- I felt very confident using the bot.
- I needed to learn a lot of things before I could get going with the bot.
- I am observant and therefore chose 2.
- The images used in the bot are fitting.
- The bot should use fewer images.
- It was clear at all times which alerts I subscribed to and for which locations.

There are already several mobile warning channels in Germany. In the following, we are interested in your assessment of the bot in comparison to warning apps and cell broadcast.

- Q22: **In your opinion, what are the advantages or disadvantages of the bot compared to warning apps or Cell Broadcast?** (Info: *NINA* or *KatWarn* are apps that can send warnings to mobile phones in crises and offer a variety of other functions that can be helpful in crises. For example, they can call up warnings for the current location. Cell Broadcast can send warnings as push messages without having to install an app.) [Comment]
- Q23: **What would you change about the bot?** (Nothing at all; The following: [Answer])
- Q24: **If the bot was available for my favourite messenger (*WhatsApp*, *Telegram*, *Signal* etc.), I would use it to...** (1 (very unlikely) – 5 (very likely))
 - ...be alerted automatically for certain locations (favourites).
 - ...query the warning status for individual locations.
- Q25 [*User Engagement Scale*]: **Please indicate how much you agree with the following statements about your subjective experience of the bot.** (1 (strongly disagree) – 5 (strongly agree))
 - Time flew by when I was using the warning bot. [*Focused Attention*]
 - I was frustrated while using the warning bot. [*Perceived Usability*]
 - I found the warning bot confusing. [*Perceived Usability*]
 - Using the warning bot was exhausting. [*Perceived Usability*]
 - I could not run all the applications I wanted to with the warning bot. [*Perceived Usability*]
 - The warning bot was appealing. [*Aesthetic Appeal*]
 - The warning bot was aesthetically pleasing. [*Aesthetic Appeal*]
 - The warning bot appeals to my visual senses. [*Aesthetic Appeal*]

- The warning bot is worth using
- My experience with the warning bot was worthwhile. [Reward]
- I would recommend the warning bot to my family and friends. [Reward]
- Out of curiosity, I would continue to use the warning bot. [Reward]
- The warning bot has captivated me. [Reward]
- Q26: **Do you have feedback regarding the study?** [Comment]

Appendix C: Comparison of messenger app characteristics

Table 7: Comparison of mobile instant messaging apps.

	WhatsApp	Viber	Telegram
Publisher	WhatsApp Inc./Meta Platform	Rakuten Viber	Telegram FZ-LLC
Website	www.WhatsApp.com	www.Viber.com	www.Telegram.org
Price	Free	Free	Free
Backups	Cloud	Cloud	Cloud, JSON, HTML
Broadcast	✓	✓	–
Online status	✓	✓	✓
Text formatting	✓	✓	✓
Group chats	1,024 members	250 members	200.000 members
Channels	–	∞ members	∞ members
Multi-device	✓	✓	✓
Online status	✓	✓	✓
Platform	Android, IOS, Windows, Mac, Linux	Android, IOS, Windows, Mac, Linux	Android, IOS, Windows, Mac, Linux
Photo editor	✓	✓	✓
Read confirmation	✓	✓	✓
Registration	Mobile phone number	Mobile phone number	Mobile phone number
Chat safety	End-to-end	End-to-end	End-to-end, MTProto
Send pictures	✓	✓	✓
Send videos	✓	✓	✓
Send files 16MB	200MB	2GB (premium 4GB)	✓
Share location	✓	✓	✓
Sticker	✓	✓	✓
Users	Approx. 2 bil.	Approx. 1.39 bil.	Approx. 700 mil.
Share username instead of phone number	–	–	✓

References

1. Tan, M. L.; Prasanna, R.; Stock, K.; Hudson-Doyle, E.; Leonard, G.; Johnston, D. Mobile Applications in Crisis Informatics Literature: A Systematic Review. *Int. J. Disaster Risk Reduct.* **2017**, *24*, 297–311. .
2. Hauri, A.; Kohler, K.; Scharte, B. *A Comparative Assessment of Mobile Device-Based Multi-Hazard Warnings: Saving Lives Through Public Alerts in Europe*; Center for Security Studies (CSS), ETH Zürich, 2022.
3. Li, T.; Cobb, C.; Yang, J. J.; Baviskar, S.; Agarwal, Y.; Li, B.; Bauer, L.; Hong, J. I. What Makes People Install a COVID-19 Contact-Tracing App? Understanding the Influence of App Design and Individual Difference on Contact-Tracing App Adoption Intention. *Pervasive Mob. Comput.* **2021**, *75*, 101439. .
4. Fischer-Preßler, D.; Bonaretti, D.; Fischbach, K. A. Protection-Motivation Perspective to Explain Intention to Use and Continue to Use Mobile Warning Systems. *Bus. Inf. Syst. Eng.* **2021**, *64*, 167–182.

5. Cornia, A.; Dressel, K.; Pfeil, P. Risk Cultures and Dominant Approaches towards Disasters in Seven European Countries. *J. Risk Res.* **2016**, *19*, 288–304. .
6. Reuter, C.; Kaufhold, M.-A.; Schmid, S.; Spielhofer, T.; Hahne, A. S. The Impact of Risk Cultures: Citizens' Perception of Social Media Use in Emergencies Across Europe. *Technol. Forecast. Soc. Change* **2019**, *148*, 119724. .
7. Appleby-Arnold, S.; Brockdorff, N. *Culture and Disaster Risk Management — Synthesis of Citizens' Reactions and Opinions During 6 Citizen Summits: Romania, Malta, Italy, Germany, Portugal and the Netherlands*; Culture And RiSk management in Man-made And Natural Disasters: Malta, 2018.
8. Statista. eMarketer Anzahl der Nutzer von Messenger-Apps Weltweit in den Jahren 2018 bis 2020 sowie eine Prognose bis 2025, 2021. <https://de.statista.com/statistik/daten/studie/1073385/umfrage/anzahl-der-nutzer-von-messenger-apps-weltweit/>.
9. Newman, N.; Fletcher, R.; Eddy, K.; Robertson, C. T.; Nielsen, R. K. *Digital News Report 2023*; Reuters Institute for the Study of Journalism: Oxford, UK, 2023.
10. Jones, B.; Jones, R. Public Service Chatbots: Automating Conversation with BBC News. *Digit. Journal.* **2019**, *7*, 1032–1053. .
11. Bundesministerium für Gesundheit. Bundesministerium für Gesundheit Telegram Infokanal des Bundesministerium für Gesundheit, 2020. https://t.me/s/corona_infokanal_bmg?before=17.
12. Haunschild, J.; Kaufhold, M.-A.; Reuter, C. Perceptions and Use of Warning Apps — Did Recent Crises Lead to Changes in Germany? In *Mensch und Computer 2022 — Tagungsband*: New York, 2022; pp. 25–40.
13. Tan, M. L.; Prasanna, R.; Stock, K.; Hudson-Doyle, E.; Leonard, G.; Johnston, D. Modified Usability Framework for Disaster Apps: A Qualitative Thematic Analysis of User Reviews. *Int. J. Disaster Risk Sci.* **2020**, *11*, 615–629. .
14. Kaufhold, M.-A.; Haunschild, J.; Reuter, C. Warning the Public: A Survey on Attitudes, Expectations and Use of Mobile Crisis Apps in Germany. In *Proceedings of the 28th European Conference on Information Systems (ECIS)*: Marrakech, Morocco, 2020.
15. Elshan, E.; Engel, C.; Ebel, P.; Siemon, D. Assessing the Reusability of Design Principles in the Realm of Conversational Agents. In *The Transdisciplinary Reach of Design Science Research*; Springer: Cham, 2022; pp 128–141.
16. Dulin, P.; Mertz, R.; Edwards, A.; King, D. Contrasting a Mobile App with a Conversational Chatbot for Reducing Alcohol Consumption: Randomized Controlled Pilot Trial. *JMIR Form. Res.* **2022**, *6*, e33037. .
17. Tan, M. L.; Prasanna, R.; Stock, K.; Hudson-Doyle, E.; Leonard, G.; Johnston, D. Enhancing the Usability of a Disaster App: Exploring the Perspectives of the Public as Users. In *16th ISCRAM Conference*, 2019; pp. 876–886.
18. Sugisaki, K.; Bleiker, A. Usability Guidelines and Evaluation Criteria for Conversational User Interfaces: A Heuristic and Linguistic Approach. In *Proceedings of the Conference on Mensch Und Computer*, 2020; pp. 309–319.
19. Shneiderman, B.; Plaisant, C.; Cohen, M.; Jacobs, S.; Elmqvist, N.; Diakopoulos, N. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 6th ed; Pearson: London, England, 2016.
20. Hevner, A. R.; March, S. T.; Park, J.; Ram, S. Design Science in Information Systems Research. *MIS Q.* **2004**, *28*, 75–105. .
21. Peffers, K.; Tuunanen, T.; Rothenberger, M. A.; Chatterjee, S. A Design Science Research Methodology for Information Systems Research. *J. Manag. Inf. Syst.* **2007**, *24*, 45–77. .
22. Wobbrock, J. O.; Kientz, J. A. Research Contributions in Human-Computer Interaction. *Interactions* **2016**, *23*, 38–44. .
23. Meier, P.; Beinke, J. H.; Fitte, C.; Behne, A.; Teuteberg, F. FeelFit — Design and Evaluation of a Conversational Agent to Enhance Health Awareness. In *International Conference on Information Systems*, 2019.
24. Gregor, S.; Hevner, A. R. Positioning and Presenting Design Science Research for Maximum Impact. *MIS Q.* **2013**, *37*, 337–355. .
25. Klopfenstein, L. C.; Delpriori, S.; Malatini, S.; Bogliolo, A. The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms. In *Proceedings of the 2017 Conference on Designing Interactive Systems*, 2017; pp. 555–565.
26. Dallo, I.; Martí, M. Why Should I Use a Multi-Hazard App? Assessing the Public's Information Needs and App Feature Preferences in a Participatory Process. *Int. J. Disaster Risk Reduct.* **2021**, *57*, 102197. .
27. Tan, M. L.; Prasanna, R.; Stock, K.; Doyle, E. E. H.; Leonard, G.; Johnston, D. Usability Factors Influencing the Continuance Intention of Disaster Apps: A Mixed-Methods Study. *Int. J. Disaster Risk Reduct.* **2020**, *50*, 101874. .
28. Bonaretti, D.; Fischer, D. Timeliness, Trustworthiness, and Situational Awareness: Three Design Goals for Warning with Emergency Apps. In *Forty-Second International Conference on Information Systems*, 2021; pp. 1–17.
29. Haunschild, J.; Kaufhold, M.-A.; Reuter, C. Sticking with Landlines? Citizens' and Police Social Media Use and Expectation during Emergencies. In *Proceedings of the International Conference on Wirtschaftsinformatik*, 2020; pp. 1–16.
30. Tan, M. L.; Prasanna, R.; Stock, K.; Doyle, E. E.; Leonard, G.; Johnston, D. Understanding End-Users' Perspectives: Towards Developing Usability Guidelines for Disaster Apps. *Prog. Disaster Sci.* **2020**, *7*, 100118. .
31. Kahneman, D.; Tversky, A. Choices, Values, and Frames. *Am. Psychol.* **1984**, *39*, 341–350. .
32. Distel, B.; Ogonek, N. To Adopt or Not to Adopt: A Literature Review on Barriers to Citizens' Adoption of E-Government Services. In *European Conference on Information Systems 2016 Proceedings*, 2016.
33. Reuter, C.; Kaufhold, M.-A.; Leopold, I.; Knipp, H. KATWARN, NINA, or FEMA? Multi-Method Study on Distribution, Use and Public Views on Crisis Apps. In *European Conference on Information Systems (ECIS)*, 2017; pp. 2187–2201.
34. Følstad, A.; Bjerkreim-Hanssen, N. User Interactions with a Municipality Chatbot—Lessons Learnt from Dialogue Analysis. *Int. J. Hum. Comput. Interact.* **2023**, *40*, 1–14. .
35. Chattaraman, V.; Kwon, W.-S.; Gilbert, J. E.; Ross, K. Should AI-Based, Conversational Digital Assistants Employ Social- or Task-Oriented Interaction Style? A Task-Competency and Reciprocity Perspective for Older Adults. *Comput. Hum. Behav.* **2019**, *90*, 315–330. .
36. Følstad, A.; Skjuve, M.; Brandtzaeg, P. B. Different Chatbots for Different Purposes: Towards a Typology of Chatbots to

- Understand Interaction Design. In *International Conference on Internet Science (INSCI)*, 2019; pp. 145–156.
37. Moore, R. J.; Arar, R.; Ren, G.-J.; Szymanski, M. H. Conversational UX Design. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017; pp. 492–497.
 38. Janssen, A.; Cardona, D. R.; Passlick, J.; Breitner, M. H. How to Make Chatbots Productive — A User-Oriented Implementation Framework. *Int. J. Hum. Comput. Stud.* **2022**, *168*, 102921. .
 39. Chaves, A. P.; Gerosa, M. A. How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design. *Int. J. Hum. Comput. Interact.* **2021**, *37*, 729–758. .
 40. Li, C.-H.; Yeh, S.-F.; Chang, T.-J.; Tsai, M.-H.; Chen, K.; Chang, Y.-J. A Conversation Analysis of Non-Progress and Coping Strategies with a Banking Task-Oriented Chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020; pp. 1–12.
 41. Følstad, A.; Brandtzaeg, P. B. Users' Experiences with Chatbots: Findings from a Questionnaire Study. *Qual. User Exp.* **2020**, *5*, 1–14. .
 42. Zamora, J. I'm Sorry, Dave, I'm Afraid I Can't Do That: Chatbot Perception and Expectations. In *Proceedings of the 5th International Conference on Human Agent Interaction*, 2017; pp. 253–260.
 43. Göbel, J.; Boldt, J.; Betke, H.; Kuehnel, S.; Damarowsky, J.; Böhmer, M.; Sackmann, S. Towards a Design Theory for Task-Oriented Conversational Agents. *European Conference on Information Systems (ECIS 2024)*, 2024.
 44. Jain, M.; Kumar, P.; Kota, R.; Patel, S. N. Evaluating and Informing the Design of Chatbots. In *Proceedings of the 2018 Designing Interactive Systems Conference*, 2018; pp. 895–906.
 45. Stieglitz, S.; Hofeditz, L.; Brünker, F.; Ehnis, C.; Mirbabaie, M.; Ross, B. Design Principles for Conversational Agents to Support Emergency Management Agencies. *Int. J. Inf. Manag.* **2022**, *63*, 1–11. .
 46. Ashktorab, Z.; Jain, M.; Liao, Q. V.; Weisz, J. D. Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019; pp. 1–12.
 47. Nguyen, Q. N.; Sidorova, A.; Torres, R. User Interactions with Chatbot Interfaces vs. Menu-Based Interfaces: An Empirical Study. *Comput. Hum. Behav.* **2022**, *128*, 1–13. .
 48. Moore, R. J.; Arar, R.; Ren, G.-J.; Szymanski, M. H., Eds. *Studies in Conversational UX Design*; Springer Nature: Cham, Switzerland, 2018.
 49. Piccolo, L. S. G.; Roberts, S.; Iosif, A.; Alani, H. Designing Chatbots for Crises: A Case Study Contrasting Potential and Reality. In *Proceedings of the 32nd International BCS Human Computer Interaction Conference*, 2018; pp. 1–10.
 50. Makasi, T.; Nili, A.; Souza, K.; Tate, M. Public Service Values and Chatbots in the Public Sector: Reconciling Designer Efforts and User Expectations. In *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022; pp. 2334–2343.
 51. Abbas, N.; Følstad, A.; Bjørkli, C. A. *Chatbot Research and Design*; Følstad, A., Araujo, T., Papadopoulos, S., Law, E. L.-C., Luger, E., Goodwin, M., Brandtzaeg, P. B., Eds.; Springer International Publishing: Cham, Switzerland, Vol. 13815, 2023; pp 66–82.
 52. Moore, R. J.; Arar, R. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework*; Association for Computing Machinery: New York, NY, 2019.
 53. Bundesstelle für Open Data. Bund.dev Bundesamt für Bevölkerungsschutz — NINA API — OpenAPI Documentation, 2022. <https://nina.api.bund.dev>.
 54. Lou, C.; Tandoc, E. C.; Hong, L. X.; Pong, X. Y.; Lye, W. X.; Sng, N. G. When Motivations Meet Affordances: News Consumption on Telegram. *Journal. Stud.* **2021**, *22*, 934–952. .
 55. Statista. Statista Most Popular Messenger Apps. 2024. <https://www.statista.com/statistics/258749/most-popular-global-mobile-messenger-apps/>.
 56. Sutikno, T.; Handayani, L.; Stiawan, D.; Riyadi, M. A.; Subroto, I. M. I. WhatsApp, Viber and Telegram: Which is the Best for Instant Messaging? *Int. J. Electr. Comput. Eng.* **2016**, *6*, 909–914. .
 57. Telegram Telegram Bot API, 2022. <https://core.telegram.org/bots/api>.
 58. Kim, S. Toward an Integrative Framework of Technology Use (IFTU): Alternative Three-Wave Panel Models and Empirical Tests. In *DIGIT 2006 Proceedings*, 2006; pp. 1–35.
 59. Franke, T.; Attig, C.; Wessel, D. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *Int. J. Hum. Comput. Interact.* **2019**, *35*, 456–467. .
 60. Wessel, D.; Attig, C.; Franke, T. ATI-S — An Ultra-Short Scale for Assessing Affinity for Technology Interaction in User Studies. In *Proceedings of Mensch Und Computer 2019*, 2019; pp. 147–154.
 61. International Organization for Standardization. Ergonomics of Human-System Interaction — Part 11: Usability: Definitions and Concepts. **2018**.
 62. Brooke, J. *Usability Evaluation in Industry*; Jordan, P. W., Thomas, B., Weerdmeester, B. A., McClelland, I. L., Eds.; Taylor & Francis: Abingdon-on-Thames, 1995.
 63. Bangor, A.; Kortum, P. T.; Miller, J. T. An Empirical Evaluation of the System Usability Scale. *Int. J. Hum. Comput. Interact.* **2008**, *24*, 574–594. .
 64. Lewis, J. R. The System Usability Scale: Past, Present, and Future. *Int. J. Hum. Comput. Interact.* **2018**, *34*, 577–590. .
 65. Sauro, J.; Lewis, J. R. *Quantifying the User Experience: Practical Statistics for User Research*, 2nd ed; Elsevier Publishing: Amsterdam, 2016.
 66. Davis, F. D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Q.* **1989**, *13*, 319. .
 67. MacDonald, C. M.; Atwood, M. E. What Does it Mean for a System to be Useful? An Exploratory Study of Usefulness. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, 2014; pp. 885–894.
 68. Toomim, M.; Kriplean, T.; Pörtner, C.; Landay, J. Utility of Human-Computer Interactions: Toward a Science of Preference Measurement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)*, 2011; pp. 2275–2284.
 69. O'Brien, H. L.; Toms, E. G. The Development and Evaluation of a Survey to Measure User Engagement. *J. Am. Soc. Inf. Technol.* **2010**, *61*, 50–69. .
 70. O'Brien, H. L.; Cairns, P.; Hall, M. A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form. *Int. J. Hum. Comput. Stud.* **2018**, *112*, 28–39. .
 71. Mentler, T.; Herczeg, M. On the Role of User Experience in Mission- or Safety-Critical Systems. In *Mensch und Computer 2016*, 2016.

72. Braun, V.; Clarke, V. Using Thematic Analysis in Psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. .
73. Rovai, A. P.; Baker, J. D.; Ponton, M. K. *Social Science Research Design and Statistics: A Practitioner's Guide to Research Methods and IBM SPSS*; Watertree Press LLC: Chesapeake, VA, 2013.
74. Hughes, A. L.; Palen, L.; Sutton, J.; Liu, S. B.; Vieweg, S. "Site-Seeing" in Disaster: An Examination of On-Line Social Convergence. In *5th International ISCRAM Conference*, 2008; pp. 1–10.
75. Pradhan, A.; Mehta, K.; Findlater, L. "Accessibility Came by Accident": Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018; pp. 1–13.
76. Makasi, T.; Nili, A.; Desouza, K. C.; Tate, M. A Typology of Chatbots in Public Service Delivery. *IEEE Softw.* **2022**, *39*, 58–66. .
77. O'Brien, H.; Cairns, P. An Empirical Evaluation of the User Engagement Scale (UES) in Online News Environments. *Inf. Process. Manag.* **2015**, *51*, 413–427. .
78. Haque, M. D. R.; Rubya, S. An Overview of Chatbot-Based Mobile Mental Health Apps: Insights from App Description and User Reviews. *JMIR mHealth uHealth* **2023**, *11*, 1–18. .
79. Holdener, M.; Gut, A.; Angerer, A. Applicability of the User Engagement Scale to Mobile Health: A Survey-Based Quantitative Study. *JMIR mHealth uHealth* **2020**, *8*, 1–9. .
80. Distel, B. Assessing Citizens' Non-Adoption of Public E-Services in Germany. *Inf. Polity* **2020**, *25*, 39–360. .
81. Distel, B. Bringing Light into the Shadows: A Qualitative Interview Study on Citizens' Non-Adoption of E-Government. *Electron. J. e-Gov.* **2018**, *16*, 98–205.
82. Yeh, S.-F.; Wu, M.-H.; Chen, T.-Y.; Lin, Y.-C.; Chang, X.; Chiang, Y.-H.; Chang, Y.-J. How to Guide Task-Oriented Chatbot Users, and when: A Mixed-Methods Study of Combinations of Chatbot Guidance Types and Timings. In *CHI Conference on Human Factors in Computing Systems*, 2022; pp. 1–16.
83. Furqan, A.; Myers, C.; Zhu, J. Learnability Through Adaptive Discovery Tools in Voice User Interfaces. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017; pp. 1617–1623.
84. Martiniello, N.; Eisenbarth, W.; Lehane, C.; Johnson, A.; Wittich, W. Exploring the Use of Smartphones and Tablets Among People with Visual Impairments: Are Mainstream Devices Replacing the Use of Traditional Visual Aids? *Assist. Technol.* **2022**, *34*, 34–45. .
85. Sherman-Morris, K.; Pechacek, T.; Griffin, D. J.; Senkbeil, J. Tornado Warning Awareness, Information Needs and the Barriers to Protective Action of Individuals Who Are Blind. *Int. J. Disaster Risk Reduct.* **2020**, *50*, 101709. .
86. Statista U.S. Top Messaging Services by Age, 2024. <https://www.statista.com/statistics/1310687/us-top-messaging-services-by-age/>.
87. Rapp, A.; Curti, L.; Boldi, A. The Human Side of Human-Chatbot Interaction: A Systematic Literature Review of Ten Years of Research on Text-Based Chatbots. *Int. J. Hum. Comput. Stud.* **2021**, *151*, 102630. .
88. Staegemann, D.; Volk, M.; Daase, C.; Pohl, M.; Turowski, K. A Concept for the Use of Chatbots to Provide the Public with Vital Information in Crisis Situations. In *Proceedings of Sixth International Congress on Information and Communication Technology: Singapore*, 2022; pp. 281–289.
89. Haesler, S.; Wendelborn, M.; Reuter, C. Getting the Residents' Attention: The Perception of Warning Channels in Smart Home Warning Systems. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, 2023; pp. 1114–1127.
90. Haunschild, J.; Pauli, S.; Reuter, C. Preparedness Nudging for Warning Apps? A Mixed-Method Study Investigating Popularity and Effects of Preparedness Alerts in Warning Apps. *Int. J. Hum. Comput. Stud.* **2023**, *172*, 102995. .

Supplementary Material: This article contains supplementary material (<https://doi.org/10.1515/icom-2024-0067>).