Towards a general methodology to assess the ease of use of public safety applications and crisis communications services

Eridy Lukau
Fraunhofer Institute FOKUS
Berlin, Germany
eridy.lukau@fokus.fraunhofer.de

Janine Hellriegel
Fraunhofer Institute FOKUS
Berlin, Germany
janine.hellriegel@fokus.fraunhofer.de

Michael Klafft
Fraunhofer Institute FOKUS
Jade University Wilhelmshaven
Berlin, Germany
michal.klafft@fokus.fraunhofer.de

ABSTRACT

Ever since Davis' Technology acceptance model for information technology has been introduced in the 1980ies, ease of use has been identified as a key determinant of people's willingness to use information systems. This factor is particularly relevant for information and communication technology (ICT) for disaster situations, where users are under severe stress, potentially in danger, and have to cope with adverse conditions such as partial or complete breakdowns of infrastructure, power blackouts, and network congestions. The impact of such adverse conditions show that ease of use of ICT for disaster management and crisis communication is not only determined by the characteristics of the user interface, but also substantially impacted by properties of the backend and the infrastructure, such as its resilience and availability. This paper proposes a generalized metric for a holistic assessment of public safety systems and services for disaster management as well as emergency and crisis communication. This metric can be used to compare different ICT alternatives and to extend existing qualitative and quantitative approaches such as thinking aloud, focus groups, surveys and field tests by specifically addressing the unique aspects of a disasters which are difficult to simulate in trials and difficult to assess in interviews and surveys, in particular if a newly developed system or technology has never been used in a real disaster setting so far.

KEYWORDS

Ease of use, ICT for disaster management, crisis communication, metric

1 INTRODUCTION

Our introduction motivates the need for a general ease of use assessment method for public safety applications and crisis communication services. So far, ease of use, usability, user experience assessments, and use of applications for citizens in crisis situations have been mostly analyzed on a case-by-case basis via taylor-made approaches for the specific case or application type, often based on a mix of semi-structured interviews and specifically designed questionnaires (for examples, cf. [10] and [20]). To the best of our knowledge, currently there is no overall cross-cutting assessment method in use that allows the calculation of a "citizens' ease-of-use factor" specifically for public safety applications and crisis communication services that are designed to be used by the public within unusual, stressful and difficult situations and in settings where parts of the infrastructure can become inoperational. We therefore aim to develop a new framework that facilitates a public-safety and crisis-communication related evaluation of commonly known technologies such as software-based-systems (i.e. apps, websites) and other communication technologies (i.e. TV, radio, commercial displays). The evaluation mainly aims at mandatory ease-of-use requirements of technology that is used by citizens' in disaster and crisis situations.

A primary objective of this evaluation concept is to establish a comprehensive framework that is capable of accurately categorizing and scoring communication technologies and tools for public safety. We aim to assess the citizens' ease of use by assessing preconditions (e.g. availability and impact of resources such as power or networks), interaction steps needed to use the technology, and interaction weights for an assessment of the complexity of use.

2 STATE OF RESEARCH

Since the early days of technology acceptance research, ease of use has been identified as a key determinant of people's willingness to use information systems. Ease of use exerts this influence both directly by creating a positive attitude towards using, and indirectly through its influence on a systems' perceived usefulness [7]. In the context of information and communication technology systems for disaster and crisis situations, ease of use is particularly relevant because these systems will typically be used in stressful situations where users neither have the time nor the resources to put a lot of efforts into using these systems. It is important to note that ease of use itself is greatly influenced by the systems' design features [7], which includes the user interface but goes beyond interaction design and comprises all other aspects of a system which have an

Veröffentlicht durch die Gesellschaft für Informatik e.V.
in P. Fröhlich und V. Cobus (Hrsg.):
Mensch und Computer 2023 – Workshopband, 03.-06. September 2023, Rapperswil (SG)

https://doi.org/10.18420/muc2023-mci-ws01-370

influence on if and how it can be used. It is therefore important to analyze the design features and the perceived ease of use of information and communication technology (ICT) for disaster management. One way to address this issue is to combine qualitative and quantitative research methods, e.g. semi-structured interviews with surveys and field tests (for an example, cf. [10]). These approaches have the advantage that each study can be specifically tailored towards the application(s) analyzed, and they provide valuable insights by identifying key shortcomings of the tested application(s). However, they normally have to be conducted in exercises or fictitious settings such as game-based scenarios where the stress level and possible technical deficiencies of a real disaster situation can only roughly be simulated. Another challenge comes from the fact that a lot of citizens may try new ICT tools for disaster communication ad hoc and use them for the first time in a stressful disaster setting, thus potentially creating additional „ease of use"-related issues due to the additional burden on the communication infrastructure and the possibly resulting slow speed of the applications. For the specific case of applications to support first responders, [23] published a handbook with an overview of evaluation methods for the usability of said systems, but they do not incorporate backend-related parameters nor do they provide a holistic metric for the ease of use that addresses system availability, robustness, resilience and performance. Niebla et al. [17] propose an assessment of technologies for the distribution of warning messages based on alert capability, information, coverage, availability and security, but they do neither generalize this for ICT in disaster management nor do they provide an integrated metrics together with frontend-related issues.

Other authors have proposed scales to assess interaction complexity. Constantine and colleagues, for example, developed a suite of five design metrics to support user interface designers and developers ([4],[5],[18]). The authors propose methods to measure the following aspects: layout uniformity, which „measures the graphical uniformity or orderliness of a user interface layout", task visibility, which is a „measure of accessing features", task concordance as a „measure of fit between layout and task structure", visual coherence, „a measure of semantic organization", and essential efficiency „as a measure of expected task efficiency" [18]. Alemerian and Magel elaborated further on the issue of visual cohesion and showed how a visual cohesion metric can be used to predict the usability of graphical interfaces [1]. Riegler and Holzmann [22] discuss and provide a tool to evaluate a metric for user interface complexity. These approaches provide an objective assessment of the user interface but do not allow for an overall assessment of ease of use due to the fact that other aspects of the system under analysis which go beyond user interface design are not considered, and none of these metrics-based approaches takes into consideration the specific design issues of information systems for communicating during a crisis or disaster. These design issues are also related to the backend infrastructure and system architecture, as stated by [15]. Design considerations regarding the backend and the infrastructure are particularly relevant in times of crises and have a considerable influence on the ease of use, as network congestion issues, power blackouts and the partial (or complete) destruction of the infrastructure strongly affect service performance and availability and thus the experienced ease of use. As Meissen and Fuchs-Kittowski [14] have shown, using the example of crowdsourcing integration in early warning systems, general architectural principles and foundations have a considerable impact on system effectiveness. Omilion-Hodges and Edwards [19] point towards channel specification theory and highlight that in crisis situations, one key factor for selecting a communication channel is the time which first information responders need to conduct communication. This is another argument for the relevance of ease-of-use in crisis communication, since good ease of use reduces the time needed to complete a communication task.

Our proposed assessment approach strives to incorporate all these aspects into a holistic ease of use assessment of ICT services for disaster management and crisis communication.

3 PUBLIC SAFETY TECHNOLOGY EASE OF USE ASSESSMENT MODEL

For the assessment of public safety services involving citizens and crisis communication applications (from here on referred to as Public Safety Technology) we aim to develop a baseline assessment method that enables equal evaluation of public safety related solutions of various kinds, based upon interaction complexity as well as other system properties beyond interaction design.

The underlying principle is to incorporate the user's interaction with a public safety technology as fundamental component of the a model that describes the technologies' complexity, applicability in disaster-related situations and its ease of use. Thus, our model considers the total number of interactions required for the technology to be used successfully and calculates an indicator we describe as the expected citizens' ease of use. The ease of use indicator describes the expected effort for a user while conducting all necessary tasks and steps in order to reach a specific purpose successfully. For example: The indicator may describe the citizens expected effort to receive and understand a warning in a crowded area, and evaluate the used technology whether it is an app on the smartphone or a public announcement via speaker box. The model may also consider technology underlying protocols (IP, voice, radio) or communication-pattern (i.e. bi-directional communication vs. uni-directional communication).

A **single interaction** can be defined as close as a finger-tap on a virtual button in an app or an interaction with a haptic physical button. Interactions however may further also be grouped as topic based interactions. Such as all necessary steps needed to be taken before a disaster event (disaster-preparedness). Or all preparatory configurations on a phone that are required to install an app. We may refer to these interactions as **grouped-interaction**.

We define a set of all possible single-interactions and grouped-interactions that describe a form of engagement with public safety related technologies as a set of unordered interactions denoted as I:

$$\mathbb{I} = \{i_{0,} i_1, \dots, i_n\} \tag{1}$$

All interactions that are mandatory in engaging with a specific technology (i.e. all interactions needed to use a specific early warning app) are therefore always defined as a predefined subset of all possible interactions (I) denoted as (I'). For example consider

$$\mathbb{I}' = \{i_{0,} i_2, i_5, i_8\} \tag{2}$$

It is important to note, that each interaction may be of a different difficulty or importance for a user. Interactions done in stress may be more difficult than in calmness. Preparatory grouped-interactions may be more difficult for one group of users, based on educational levels. Some technical interactions may require certain background knowledge or some form of previous experience. Each interaction or each grouped interaction needs to be considered with a weight, that takes its gravity for the user into account.

This weight may then be used to describe a disaster related influence on each single-interaction or grouped-interaction either on the level of difficulty, complexity, importance, cognitive load, or on situation-specific circumstances (i.e. stress, imminent danger). We therefore aim to provide each interaction a specific weight w defined as:

$$w: I \longrightarrow \mathbb{R}^{+} \tag{3}$$

The function w assigns a weight to every interaction in $\mathbb{I}$. The citizens' ease of use indicator can now be calculated as the sum of all mandatory interactions, with each interaction supplied with its own weight. Based on former definitions, the indicator is denoted as *Y* and determined as:

$$\Upsilon = w(i_1) + w(i_2) + \cdots + w(i_n) \tag{4}$$

or simplified:

$$\Upsilon = \sum_{i_j \in \mathbb{I}' \subseteq \mathbb{I}} w(i_j) \tag{5}$$

When considering different communication tools and channels on an aggregate level, weights may also take into consideration how widely a channel / an application is already used. One example might be alerting apps where statistics show that a certain percentage of people already uses such apps (for an example, see [20]), thus reducing the required installation effort for already existing users to zero.

A requirement for usability of public safety related technologies is their resilience against hampering circumstances in pre-, within-, or post disaster situations. These circumstances have the potential to drastically change the environment in which a user engages with a technology. Thus, the usage of each technology relies on specific mandatory preconditions that have to be set. These preconditions have an order of magnitude that influences the ease of use in a way that they either hamper the engagement with the technology or in some cases make a technology completely unusable. For example: access to the internet is a mandatory precondition for an early warning app that receives its content from a remote server. Access to electricity may hamper the usage of an app in a smartphone, however when the battery level is sufficient, a user might still be able to use the app successfully. A power outage on the other hand may make technologies such as TV or cable-based phone that are connected to a router unusable.

Regarding the assessment of public safety related technology this must be taken into account. We therefore consider the incorporation of these mandatory preconditions as crucial part of our model. We propose the incorporation of mandatory preconditions as a list of requirements for a specific technology that has to be set so that a technology can be considered as being usable. This is a precondition that must be met before the interactions can be evaluated. The model considers this list as either filled or empty, so that the calculated ease of use indicator is calculated accordingly. Therefore, the preconditions are indicated as C:

$$\mathbb{1}_P: C \rightarrow \{0,1\} \tag{6}$$

so that they can be defined as binary condition that is mapped by the function $\mathbb{1}$ to either 0 if no mandatory preconditions can be met (i.e. missing electricity) and 1 if mandatory preconditions are set (i.e. available electricity):

$$\mathbb{1}_P(c) := \begin{cases} 1 & \text{if } c \neq \emptyset, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

In another variation of C, the elements of the list that are either 0 or 1 may be multiplied with C as the result. If each of the elements is a precondition, it is 1 if met or 0 if not. Thus resulting in C = 0 as a product of multiplications of each precondition as soon as one precondition is 0. This maps the situation if an elemental precondition is not met of a specific technology in order to be used.

The precondition C can now be used as factor that influences the indicator in a way that, if the set of preconditions is not met the indicator automatically is zero, indicating non-functionality of the technology. This can be expressed as follows:

$$\Upsilon(\mathrm{c}, \mathbb{I}') = \mathbb{1}(c) \sum_{i_j \in \mathbb{I}' \subseteq \mathbb{I}} w(i_j) \tag{8}$$

The hereby presented model shows a first approach of an ease of use description model on public safety technology that inseparably combines: the users' utilization of a technology described as (grouped-) interactions, the weight of each interaction as an indicator of complexity or difficulty and the precondition that is mandatory in order to use the technology. The calculated ease of use indicator Y is a score that defines whether a public-safety technology is usable or unusable and makes multiple public-safety technologies directly comparable. It can therefore be read as follows:

A public-safety application or a public-safety service can be evaluated as unusable, if the calculated indicator Y is zero, so that $Y = 0$. This is when not all mandatory preconditions for the usage of the technology can be met. For example: the TV is not usable without electricity.

A public-safety application or a public-safety service can be evaluated as usable, if the calculated Indicator Y is bigger than zero, so that $Y > 0$. This is when all mandatory preconditions for the usage of the technology can be met. For example: A Warning-App can be used if it has enough battery and there is an internet connection that is sufficient to load content.

- Two public-safety technologies (A and B) are comparable in a way that:
  - if $Y_A == Y_B$, both technologies are **equally usable**,
  - If $Y_A > Y_B$, technology A is **less usable** or **more complex** than technology B.

4 EASE OF USE ASSESSMENT OF PUBLIC SAFETY TECHNOLOGIES IN DIFFERENT SITUATIONS

To showcase a proof of concept, we present three scenarios that demonstrate our model as an assessment method for the evaluation of the ease of use of certain public safety technologies in different disaster scenarios. The rating also helps us to understand which technologies are the easiest or most challenging to use in critical situations.

The first scenario describes a risk situation at a mass gathering in which the citizens must be informed about the immediate cancellation of the event. The second scenario describes a post-blackout situation in a disaster-struck area with still functional mobile network infrastructure. The third scenario describes a void post-blackout situation with finally failing mobile network infrastructure. For the sake of understandability and to overcome some of the shortcomings of the proposed model in its early stage, we simplify some steps in the evaluation process:

1. We define interactions as singular interaction (or task) with a technology that can be measured and evaluated based on cognitive load. For example, one tap on a screen or reading a message is one interaction that has the potential to be cognitively more difficult in stressful situations.
2. We summarize generic interactions as grouped interaction. For example, installing an app is a combination of multiple sub-interactions. We define this as one interaction that represents a group of sub-interactions.
3. Since there is currently no data provided in order to apply universally valid weights, we choose the factor 1.5 as a general weight to show a higher complexity of a specific interaction compared to others. We also use the same factor as weight for grouped interaction to mark a higher complexity. For example, interactions that increase disaster preparedness are considered as grouped interaction with a higher complexity.
4. We consider preconditions as mandatory conditions for the technology to remain technically functional and enable users to perform basic tasks. For example, telephones connected to a fixed telephone line via router are technically non-functional if there is no electricity.

4.1 Mass gathering communication

There are different definitions of mass gatherings ([2],[25]). All in common they describe that a certain amount of people meeting spontaneously or planned for different reasons from religion to a sports event and there is an increased risk potential. Usually, the event organizers and other parties involved are in a qualified cooperation with the authorities and organizations with security tasks. In addition to hazards such as severe weather or fire, crowd accidents - "situations where mass gatherings of people lead to deaths or injuries" [8] are one of the most common disasters at large events. Some very well-known examples in the media are the accident at the Love Parade in Germany (2010), a fire at a nightclub in Basil (2013), an accident at the Hajj in Saudi Arabia (2015), or

Table 1: Calculated scores for selected public safety applications and crisis communications services at a mass gathering

Technology	Preconditions	Preconditions Met	Interaction steps	Interaction weight	Score
Giant screens	Electricity	C = 1	i_1 Look a screen	w_1 =1,5 (difficult visibility)	1,5
Loudspeaker systems	Electricity	C = 1	i_1 Listen to loudspeaker	w_1 =1,5 (difficult audibility)	1,5
Social media (app)	Battery power, mobile data	C = 1	i_1 Install application i_2 Make account i_3 Open the app i_4 Pro actively search and find the warning i_5 Read the warning	w_1 =1 w_2 = 1,5 (complex interaction) w_3 =1 w_4 =1,5 (grouped interaction) w_5 =1	6
Event mobile app	Battery power, mobile data	C = 1	i_1 Install application i_2 Make account i_3 Receive message i_4 Read the warning	w_1 =1 w_2 = 1,5 (difficult action) w_3 =1 w_4 =1	4,5
Alert system (warning app)	Battery power, mobile data	C = 1	i_1 Install application i_2 Receive message i_3 Read the warning	w_1 =1 w_2 =1 w_3 =1	3

more recently the accident during the Halloween celebrations in South Korea (2022).

Choosing the right public safety technology is therefore an important step during the planning and preparation of an event in order to be able to warn and inform the participants immediately at the event. In the research project LETSCROWD a communication toolkit was developed to support event organizers, law enforcement agency officers and first responders [12]. Their choice of crisis communication during „warning the crowd in the case of something critical occurs" are the following communication channels: giant screens, public announcement and loudspeaker systems, face-to-face communication, social media, event mobile app, alert system, emergency signs and traditional media like TV or newspaper.

In Table 1 the proposed technical public safety applications are presented together *with* their precondition, the interaction steps required for the user and a proposed weight about the complexity of the step. For each application a total score is calculated that defines the ease of use. The assessment shows clearly that for a large number of people within a restricted area on-site applications like screens and loudspeakers have a much lower score and therefore a better ease of use compared to mobile applications that a user must set up and actively operate. The complexity of using screens and speakers arises primarily from limited visibility (taller people, buildings, trees) or limited audibility (loud music, loud conversations). There is also the issue of understanding the transmitted message in terms of language or cultural context, as well as local knowledge. The various mobile applications are primarily used for individualization and offer the advantage of being able to provide comprehensive information. On the other hand, they are more complex to operate. Social media applications such as Twitter are helpful in providing an overview, but must be used actively and do not automatically alert the user. Television and newspapers were not considered, as it is assumed that participants on site do not have access to these technologies and that this is mainly used to inform arriving participants or local residents.

4.2 Post-Blackout Communication

4.2.1 Receiving a public safety warning in a Post-Blackout scenario (infrastructure still functional). We assess a list of technologies that are used by authorities to warn citizens or respectively can be used by citizens to receive warnings from authorities. The scenario defines a disaster struck area without electricity but still functional cellular network infrastructure. This example shows the grouping of interactions. For instance, if the warning is received by an app, the app must be installed before, and the phone must be set with the correct settings (push-notifications, location-services) before. In this example the usage of a battery-powered radio is assessed. We increase the weight of having batteries due to disaster preparedness.

Within this scenario, there are smartphone users who are prepared for emergencies with a fallback battery-powered radio. This can be attributed to the fact that it is more difficult to prepare against a disaster by buying additional equipment compared to having an app and to download it. In current literature it is stated that it is a difficult task for citizens to motivate themselves to prepare for disasters, in particular if the task is difficult or requires a lot of effort and people have doubts to be able to cope with it (as discussed in [11] based on protection motivation theory as described in [16]).Therefore, we increase weight on disaster preparedness as preparational step.

Table 2: Calculated score for early warning technologies in a Post-Blackout with working cellular network

Technology	Preconditions	Preconditions Met	Interaction steps	Interaction weight	Score
SMS	Mobile network, battery power	C = 1	i_1 Open the SMS in app i_2 Read the SMS	$w_1 = 1$ $w_2 = 1$	2
Radio (wired)	Electricity	C = 0	- *Nonfunctional*		0
Radio (battery)	Electricity	C = 1	i_1 Disaster preparedness and having batteries i_2 Turn on the radio i_3 Search & find a channel i_4 Listen to the warning	$w_1 = 1,5$ (grouped interaction) $w_2 = 1$ $w_3 = 1$ $w_4 = 1$	4,5
TV	Electricity	C = 0	- *Nonfunctional*		0
Warning via. app (I.e., KATWARN, NINA)	Mobile network, internet	C = 1	i_1 Install application i_2 Make correct phone settings i_3 Tap on push notification i_4 Read the warning	$w_1 = 1$ $w_2 = 1$ $w_3 = 1$ $w_4 = 1$	4
Cell Broadcast	Mobile network	C = 1	i_1 Make correct phone settings i_2 Read the Alert on Screen	$w_1 = 1,5$ (grouped interaction) $w_2 = 1$	2,5
Social Media (I.e., Twitter via app)	Internet, mobile network, electricity	C = 1	i_1 Install application i_2 Make account i_3 Open the App i_4 Pro actively search and find the warning i_5 Read the warning	$w_1 = 1$ $w_2 = 1,5$ (complex interaction) $w_3 = 1$ $w_4 = 1,5$ (grouped interaction) $w_5 = 1$	6
Social Media (I.e., Twitter via Web)	Internet, mobile network, electricity	C = 1	i_1 Make an account i_2 Visit & login via web-browser i_3 Pro actively search and find the warning i_4 Read the warning	$w_1 = 1,5$ (complex interaction) $w_2 = 1,5$ (grouped interaction) $w_3 = 1,5$ (grouped interaction) $w_4 = 1$	5,5

Our assessment in Table 2 shows that SMS and Cell Broadcast have the best ease of use indicator showing more usefulness compared to other public safety technologies on the condition of remaining battery power and provided that the mobile network is not overloaded. However, the usage of a warning app could have the same score as the SMS if the preparational steps that are installing the app and configuring the phone are already done. The usage of social media is more complex. Social media requires an account. Since the creation of an account requires several complex sup-interactions and processes (I.e. having an email account, creating passwords, verification via phone number, etc.) we see this as a more complex interaction. Furthermore, it excludes every person without a social media account on the specific platform where the warning information is shared. The use of social media using an app is slightly more complex due to the installation process, compared to the usage of a pre-installed web browser on a smartphone.

4.2.2 Providing Emergency Communication between citizens and authorities in a Post-Blackout Scenario (infra-structure non-functional). In this scenario we assess the usefulness of public safety technologies in a void-scenario as defined in [13] in which there is complete radio-silence in licensed spectrum due to finally failed mobile network infrastructure. Further there are void- communication services that provide emergency mobile network services in restricted areas using nomadic emergency networks that provide basic cellular communication services (Phone, SMS, Cell-Broadcast) but no access to World Wide Web [13].

Table 3: Calculated score for emergency communication technologies in a Post-Blackout with non-functional cellular network

Technology	Preconditions	Preconditions Met	Interaction steps	Interaction weight	Score
SMS	Mobile network, battery power	C = 1	i_1 open the SMS in app i_2 Read the SMS	$w_1 = 1$ $w_2 = 1$	2
Phone	Terrestrial-cable network, electricity (Router)	C = 0	- *Non-functional*		0
Emergency call app (Nora, HessenWarn)	Mobile Network, Internet Battery Power	C = 0	- *Non-functional*		0
Cell Broadcast	Mobile network, battery power	C = 1	i_1 Make correct phone settings i_2 Read the alert on screen	$w_1 = 1{,}5$ (grouped interaction) $w_2 = 1$	2,5
Social Media (App)	Internet, mobile Network, electricity	C = 0	- *Non-functional*		0
Social Media (Web)	Internet, mobile network, electricity	C = 0	- *Non-functional*		0

Our assessment in Table 3 shows the obvious fact that due to the circumstances of this scenario nearly no mandatory preconditions are met for apps, phones or web-based public-safety technologies. Since most phones are connected in some form to a router, they cease work without electricity. Without access to the internet, no apps can be used to either make emergency calls nor receive warnings or information from authorities. However, when authorities need to contact citizens, they might use either mass SMS as communication channel to all connected phones in range of the nomadic emergency network or either use a cell broadcast as warning or information channel to contact a huge number of persons. The assessment also shows that the usage of SMS in these scenarios is due to its low ease of use indicator (Y = 2) preferable. SMS as public safety technology in void-scenarios proves itself to be trivially accessible for most people due to less interactions needed to function, if there is at least some battery power and a simple mobile network.

5 CURRENT LIMITATIONS AND CONSIDERATIONS FOR THE FURTHER DEVELOPMENT OF THE ASSESSMENT MODEL

As presented in this work, the proposed assessment model is clearly in an early stage, as some considerations have to be considered for its next development iteration. One of the considerations regards the usage of weights for each interaction. Although it is imminent that each interaction may be considered differently compared to other interactions, there is at this point no universally applicable rule-set in place that describes an unambiguous process of how to apply weighting to distinct interactions with a selected technology in pre-, within-, or post-disaster scenario. There is also a weighting-specific order of magnitude or sizing-guideline missing that enables an accurate weighting of each interaction. As a solution for this we propose multiple steps to be taken in the future to overcome these current limitations of the model.

To be able to define an adequate weight that describes the difficulty connected to an interaction, the user experience must be integrated. This includes with regards to software-based technologies with graphical user interfaces the actual user interface design. User interfaces utilized by a user with wrong prerequisites or missing education make the completion of a task using the technology more difficult by overloading the user's short term memory [3]. On the other hand, interfaces that are less demanding on short term memory and thus less demanding on the cognitive capacity of a user increase the usability [6]. We therefore propose the conduction of surveys and experiments with user groups. These investigations should include measuring of cognitive load while utilizing a technology [21]. The goal would be to understand the level of difficulty and complexity of grouped or single interactions with public safety related technologies and services in mind. We expect the results to enable us to define specific orders of magnitudes or sizes of weightings based on cognitive load related results. This would code the users' expected experience into the assessment model. The measurements obtained and integrated as weight may describe the perceived effort of an interaction with a public safety technology on a scale from "easy" to "difficult" or "uncomplex" to "complex". Furthermore, cognitive-load based measurements would enable to account for 'stress-resilience' of interactions and consider users' capabilities, education as well as expected usage environment, since external factors also influence the amount of cognitive resources [24]. We further aim towards an adequate weighting procedure that includes socio-technical aspects. By incorporating user demographics and categorizing weight-sizes based on user demographics, the assessment model is enabled to include socio-technical perspectives into the final score. This approach considers not only the relationships between social and technical factors, but also takes into account organizational, economic, cultural, and human factors that can not only influence users' behavior but may potentially exclude user groups from the access to public safety technology. The new assessment model would penalize for instance disaster preparedness related interactions that cost money and exclude disadvantaged groups of users with a higher weight. Nevertheless, open access to the public might also be considered as mandatory precondition that would be rendered as unusable by the model if not met. However, simply not having access to a technology does not make it technically unusable but rather makes its utilization more difficult by adding a layer of complexity. The calculated ease of use score calculates the sum of all weighted and unweighted interactions multiplied. These interactions may be categorized as prerequisite interactions and in-situ interactions. Prerequisite interactions are in case of public-safety technologies all interactions that are done in order to increase disaster preparedness. In-situ interactions are all interactions that are done shortly before or within a disaster. Thus, a formerly assessed technology might have a worse score as ease of use indicator in general but also provide a better score due to less interactions within the disaster. We therefore propose to extend the score with an in-situ score as additional ease of use indicator that shows the usefulness of a technology within a disaster that benefits from the user's disaster preparedness. This can be seen in the assessment of warning apps in chapter 4.2.1, since in-situ interactions of receiving an app (Y=2) is equal to receiving and reading an SMS (Y=2). These apps might also be referred to as trivially accessible during a disaster due to their simplicity. We see the importance that emergency communication should always be simple to use or 'out of the box'.

With a clear guideline of weighing specific interactions of public safety technologies, the model would provide a rather comprehensive ease of use indicator as a comparable score of a technology.

6 CONCLUSION AND FUTURE WORK

This paper proposed the concept of a metric to assess the ease of use of information and communication system for disaster management and crisis communication from a holistic perspective. It was shown how the metric can be applied in two different scenarios and for a number of selected crisis communication / public safety technologies. Using the metric can support practitioners in the selection process of systems while planning and establishing communication infrastructures for a specific disaster case, depending on the characteristics of the case (e.g., expected damage to the infrastructure or expected level of stress etc.). Future work can also extend to pre-assessing different types of available or emerging communication infrastructures for disaster situations, such as ad-hoc peer to peer networks involving technologies such as LoRa modems [9].

However, additional steps are needed in the future to make it easier to apply the proposed metric and to apply it in a standardized and consistent way. This future work can address the following aspects:

- How to define the weights in the formulae in a standardized manner.
- More precise definition of minimum requirements (indispensable criteria which lead to an assessment as "unusable" in the metric) for a catalogue of disaster scenarios and tools / applications / channels.

After this detailing has been completed, it will be possible to develop a decision support tool for disaster management planners and actual disaster managers based upon the proposed approach.

Acknowledgements

The authors would like to thank Leonhard Löffler-Dauth for checking the correctness of the mathematical formulae presented in this paper. Eridy Lukau's and Janine Hellriegel's work was conducted within the Fraunhofer Center for the Security of Socio-Technical Systems (Fraunhofer SIRIOS) funded by the German federal government and the state of Berlin. Michael Klafft's contribution was funded by the German Federal Ministry of Economic Affairs and Climate Action through the SPELL project.

REFERENCES

[1] Alemerien, K., & Magel, K. (2015). SLC: a visual cohesion metric to predict the usability of graphical user interfaces. In Proceedings of the 30th Annual ACM Symposium on Applied Computing (pp. 1526-1533).

[2] BaSiGo. (2015). BaSiGo – Bausteine für die Sicherheit von Großveranstaltungen. The BaSiGo-Guide. URL: https://basigo.vfsg.org/index.php accessed June 16th, 2023

[3] Cohen, M. H., Giangola, J. P., & Balogh, J. (2004). Voice user interface design. Boston: Addison-Wesley Professional.

[4] Constantine, L. L. (1996a). Usage-Centered Software Engineering: New Models, Methods, and Metrics, in Purvis, M. (ed.), Software Engineering: Education & Practice. Los Alamitos, CA: IEEE Computer Society Press.

[5] Constantine, L. L. (1996b). Measuring the Quality of User Interface Designs, Software Development Conference Proceedings. San Francisco: Miller Freeman. 1996.

[6] Cooper, G. (1998). Research into cognitive load theory and instructional design at UNSW. Sydney: University of New Sout Wales.

[7] Davis, F. D. (1987). User acceptance of information systems: the technology acceptance model (TAM). Ann Arbor: University of Michigan.

[8] Feliciani, Claudio & Corbetta, Alessandro & Haghani, Milad & Nishinari, Katsuhiro. (2023). Trends in Crowd Accidents Based on an Analysis of Press Reports. SSRN Electronic Journal. 10.2139/ssrn.4367162.

[9] Höchst, J., Baumgärtner, L., Kuntke, F., Penning, A., Sterz, A., Sommer, M., & Freisleben, B. (2023). Mobile Device-to-Device Communication for Crisis Scenarios Using Low-Cost LoRa Modems. In Disaster Management and Information Technology (pp. 235–268).

[10] Kaufhold, M. A., Rupp, N., Reuter, C., Amelunxen, C., & Cristaldi, M. (2018). 112. SOCIAL: design and evaluation of a mobile crisis app for bidirectional communication between emergency services and citizen.

[11] Klafft, M., Harari, I., Bustos, R. G., Dudzińska-Jarmolińska, A., & Duarte, S. B. (2022). CITADINE: A platform to provide the experiences of survivors of natural disasters as open educational resources for risk communication. In *Advances and New Trends in Environmental Informatics: A Bogeyman or Saviour for the UN Sustainability Goals?* (pp. 187-196). Springer International Publishing.

[12] Letscrowd (ed., 2017). Law Enforcement agencies human factor methods and Toolkit for the Security and protection of CROWDs in mass gatherings. URL: https://letscrowd.eu/ accessed June 16th, 2023.

[13] Lukau, E., Schiller, J., & Meissen, U. (2023). Towards efficient Post-Blackout Emergency Communication based on Citizens' Smartphone State of Charge. In Jaziar Radianti, Ioannis Dokas, Nicolas Lalone, & Deepak Khazanchi (Eds.), Proceedings of the 20th International ISCRAM Conference (pp. 76–87). Omaha, USA: University of Nebraska at Omaha.

[14] Meissen, U.; Fuchs-Kittowski, F. (2014). Towards a reference architecture of crowdsourcing integration in early warning systems. In Hiltz, S. R.; Pfaff, M. S.; Plotnick, L.; Shih, P.C. (eds.): Proceedings of the 11th International ISCRAM Conference – University Park, Pennsylvania, USA, 339-343.

[15] Meissen, U., Hardt, M., & Voisard, A. (2014). Towards a general system design for community-centered crisis and emergency warning systems. In: Hiltz, S.R. (Hrsg.), 11th inter- national conference on information systems for crisis response and management (ISCRAM 2014)

[16] Neuwirth, K.; Dunwoody, S., Griffin, R.J. (2000). Protection motivation and risk communication. Risk Analysis 20, 721–734.

[17] Niebla, C. P., Chaves, J. M., & De Cola, T. (2016). Design aspects in multi-channel public warning systems. In *Wireless Public Safety Networks 2* (pp. 227-261). Elsevier.

[18] Noble, J., & Constantine, L. L. (1996). Interactive design metric visualization: Visual metric support for user interface design. In Proceedings Sixth Australian Conference on Computer-Human Interaction (pp. 213-220). IEEE

[19] Omilion-Hodges, L.M., & Edwards, A. L. (2021) Students as Information Responders and Creators during a University Shooting, Communication Studies, 72:4, 701-719

[20] Reuter, C., Kaufhold, M. A., Leopold, I., & Knipp, H. (2017). Katwarn, NINA or FEMA? Multi-Method study on distribution, use and public views on crisis apps. Darmstadt: TU Prints.

[21] Reuter, C., Kaufhold, M. A., Schmid, S., Spielhofer, T., & Hahne, A. S. (2019). The impact of risk cultures: Citizens' perception of social media use in emergencies across Europe. Technological Forecasting and Social Change (TFSC), 148(119724). https://doi.org/10.1016/j.techfore.2019.119724

[22] Riegler, A., & Holzmann, C. (2015). UI-CAT: Calculating user interface complexity metrics for mobile applications. In Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia (pp. 390-394).

[23] Theofanos, M.; Choong, Y. Y.; Dawkins, S.; Greene, K. Stanton, B.; Winpigler, R. (2017). Usability Handbook for Public Safety Communications - Ensuring Successful Systems for First Responders. NIST Handbook 161. Gaithersburg: National Institute of Standards and Technology.

[24] Tracy, J. P., & Albers, M. J. (2006). Measuring cognitive load to test the usability of web sites. In *Annual Conference-society for technical communication*, Vol. 53, p. 256.

[25] World Health Organization. (2015). Public health for mass gatherings: key considerations.