engage

Engage Society for
Risk Awareness and Resilience

Deliverable 3.2 – Exploration of innovative use of communication and social media technologies

Authors: Nathan Stolero (TAU), Tomer Simon (TAU), Moran Bodas (TAU), Kobi Peleg (TAU), Bruria Adini (TAU)

Contributors: SINTEF, EENA, CA, EverBridge

Abstract: Deliverable 3.2 explores innovative uses of an AI-enabled chatbot technology with the potential for enhancing societal resilience, providing the public and practitioners with a trustworthy and resilient communication channel which can provide immediate information from every device and in every situation. The deliverable reviews existing AI solutions' strengths and weaknesses, suggests improvements, describes approaches of innovative machine learning (ML) and identifies relevant datasets for AI chatbots. Deliverable 3.2 focuses on enhancing communication channels through AI, addressing the needs and expectations of the public, reducing the workload and collapse of emergency call centres during a surge of requests, neutralising false information, and distributing unbiased messages. The result of this deliverable is a design concept and a blueprint of an AI chatbot for emergencies and disasters, addressing questions of the design and implementation of the AI chatbot.

INFORMATION TABLE

Deliverable Number	3.2
Deliverable Title	Exploration of innovative use of communication and social media technologies
Project Acronym	ENGAGE
Dissemination Level PU: Public; CO: Confidential; CI: Classified	PU
Grant	882850
Project Duration	July 2020 – June 2023
Call	SU-DRS01-2018-2019-2020
Topic	Human factors, and social, societal, and organisational aspects for disaster-resilient societies
Consortium Coordinator	SINTEF
Edition date	14 Sep. 2021
Version	01.01.03

AUTHORSHIP & APPROVAL INFORMATION

EDITOR **Bruria Adini (TAU)**	DATE 28.4.2021 – PCOS 28.7.2021 – INTERMEDIATE 22.8.2021 - EXTERNAL
CONTRIBUTORS **Matthieu Branlat, Asbjørn Lein Aalberg, Martina Ragosta (SINTEF), Alexis Gizikis (EENA), Maya Battisti (CA), Haakon Straume, Lorenzo Marchetti (EverBridge)**	DATE 28.4.2021 – PCOS 30.8.2021 – External
REVIEWED BY **Peter Burgess, ENS** **Leire Labaka & Sahar Elkady, TECNUN**	DATE 8.5.2021 – PCOS 20.8.2021 – Intermediate 30.8.2021 - External

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

APPROVED BY **Matthieu Branlat**	DATE 14 Sep. 2021

ETHICS BOARD REVIEW REQUIRED?	**SECURITY BOARD REVIEW REQUIRED?**
NO	NO

DOCUMENT HISTORY

Version	Date	Version description / Milestone description
00.00.01	28 Apr. 2021	PCOS Proposed
00.00.02	8 May. 2021	PCOS Revised
00.00.03	9 May. 2021	PCOS Approved
00.01.01	28 Jul. 2021	Intermediate Proposed
00.01.02	20 Aug. 2021	Internediate Revised
00.01.03	21 Aug. 2021	Intermediate Approved
01.01.01	22 Aug. 2021	External Proposed
01.01.02	31 Aug. 2021	External Revised
01.01.03	14 Sep. 2021	External Approved

*The project uses a multi-stage internal review process with defined milestones. Milestone names include terms (in bold) as follows:

» **PCOS**

- **proposed:** Describes planned content and structure of different sections. Document authors submit for internal review.
- **revised:** Document authors produce new version in response to internal review comments.
- **approved:** Internal project reviewers accept the document.

» **Intermediate**

- **proposed:** Document is approximately 50% complete – review checkpoint. Document authors submit for internal review.
- **revised:** Document authors produce new version in response to internal reviewer comments.
- **approved:** Internal project reviewers accept the document.

» **External**

- **proposed:** Document is approximately 100% complete – review checkpoint. Document authors submit for internal review.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

- **revised:** Document authors produce new version in response to internal reviewer comments.
- **approved:** Internal project reviewers accept the document.

» **Released:** Executive Board accepts the document. Coordinator releases the deliverable to the Commission Services.

Members of the ENGAGE Consortium

SINTEF	Stiftelsen SINTEF (SINTEF) NO-7465 Trondheim Norway www.sintef.com	**Project Coordinator:** Matthieu Branlat Matthieu.Branlat@sintef.no
deepblue CONSULTING & RESEARCH	Deep Blue Srl (DBL) IT-00198 Rome Italy www.dblue.it	**Contact:** Alberto Pasquini alberto.pasquini@dblue.it
tecnun Universidad de Navarra	University of Navarra (TECNUN) SP-31009 Pamplona Spain www.tecnun.unav.edu	**Contact:** Leire Labaka llabaka@tecnun.es
TEL AVIV UNIVERSITY אוניברסיטת תל אביב	Tel Aviv University (TAU) IL-6997801 Tel Aviv Israel www.english.tau.ac.il	**Contact:** Bruria Adini adini@netvision.net.il
Røde Kors Trondheim	Trondheim Red Cross (TRC) NO-7465 Trondheim Norway www.rodekors.no/en/	**Contact:** Marita Hoel Fossen marita.fossen@redcross.no
eena EUROPEAN EMERGENCY NUMBER ASSOCIATION	European Emergency Number Association (EENA) BE- 1060 Brussels Belgium www.eena.org	**Contact:** Alexis Gizikis ag@eena.org
	Ministry of Internal Affair, Department of Emergency Situations (DSU) RO- 010086 Bucharest Romania www.dsu.mai.gov.ro	**Contact:** Raed Arafat arafatr@smurd.ro
everbridge®	Everbridge Norway (EVBG) NO-0663 Oslo Norway www.everbridge.no	**Contact:** Håkon Straume haakon.straume@everbridge.com
ENS ÉCOLE NORMALE SUPÉRIEURE	Ecole Nationale Supérioure (ENS) FR-75005 Paris France www.ens.psl.eu	**Contact:** J. Peter Burgess james.peter.burgess@ens.psl.eu

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

	ERTZAINTZA - Departamento de Seguridad – Gobierno Vasco - (ERTZ) ES- 01010 San Sebastian Spain www.ertzaintza.eus/wps/portal/ertzaintza	**Contact:** J. Jesús Alberto Alonso Velasco 06090@ertzaintza.eus
	Cittadinanzattiva (CA) IT- 00183 Rome Italy www.cittadinanzattiva.it	**Contact:** Annalisa Mandorino a.mandorino@cittadinanzattiva.it
	Azienda Sanitaria Locale Roma 1 - – Dipartimento di Epidemiologia (ASL) IT- 00198 Rome Italy www.aslroma1.it	**Contact:** Francesca de'Donato f.dedonato@deplazio.it
	Katastrofmedicinskt Centrum (KMC) SE-58330 Linköping Sweden www.lio.se/kmc	**Contact:** Carl-Oscar Jonson carl-oscar.jonson@regionostergotland.se
	NTNU Social Research Ltd. (NTNUSR) NO- 7491 Trondheim Norway www.ntnu.edu	**Contact:** Stian Antonsen stian.antonsen@ntnu.no

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Contents

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

List of Tables

List of Figures

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Executive summary

Background: One of ENGAGE's objectives is to produce validated actionable knowledge on societal resilience by demonstrating the benefits and impact of the project solutions in different types of disasters. In addition, one of the desired results of the project is to find the best practices for communication and social media (R3).

Goal: Deliverable 3.2 aims at proposing new directions for the innovative use of Artificial Intelligence (AI) technology to improve societal resilience and citizens' engagement. The characterisation of the solution allows its potential implementation in different contexts and strengthens forms of resilience embedded culturally and socially in local contexts. We developed a design concept of a blueprint for a conversational AI-enabled chatbot to be used by authorities and first responders before and during emergencies and disasters to contribute to building societal resilience. The objective of the deliverable is to suggest directions for innovating solutions that offer: (A) The ability to test or revise the assumption of ENGAGE that AI-enabled technologies can contribute to building societal resilience; (B) an AI-enabled chatbot, allowing emergency authorities to provide a contextual, online and zero-delay response to the mass-public before, during and following emergencies; (C) Innovative solutions for neutralising false messages during disasters, based on the rapid detection and tracking of trending misinformation (e.g., false rumours) on social media. The chatbot will not distribute messages actively but will rather answer the public's questions and provide unbiased information; (D) Innovative solutions for citizens' engagement and transfer of knowledge from research to the public, leveraging the project's suggested directions to relevant population groups, according to their specific needs and expectations.

Process: The process of suggesting to design a concept of the blueprint included several steps. Initially, we reviewed already existing solutions. This was done by reviewing the scientific and grey literature and discussing the options with informants from the partners of ENGAGE, Ki-CoP members, and additional professionals. The collection was systematic but included a snowball component of collecting solutions from one informant to the next, based on the recommendations that were shared. In the second step, we analysed the solution according to criteria of algorithms and datasets in use, interface with external sources and other criteria, as described in the process section. In addition, we collected existing blueprints of AI-enabled chatbots generated by leading technology companies that present state-of-the-art solutions. Finally, we suggested a concept design for the blueprint, a roadmap and a step-by-step implementation plan.

Existing Solutions: the review identified 45 AI-enabled chatbot solutions concerning emergency and disaster management. Most of them are health-related (34), and specifically, many AI-enabled chatbots were developed in the last 18 months to provide information about COVID-19 (31). Five were related to natural disasters (water, weather, food and earthquakes), three for general disaster-related issues and two focused on women in particular. Although some are still active, several stopped working for many reasons (e.g., they were intended initially for a short-term use only). Regarding the communication platform, 25 of the AI-enabled chatbots were accessible only through the web. The rest of the 20 AI-enabled chatbots were accessible through the mobile phone (e.g., SMS, messaging apps) or mobile apps, and four of them are also available by Facebook Messenger, which is accessible by mobile or web platforms. The review of the solutions highlighted that in most cases, the approach of the AI-enabled chatbots was conservative and cautious, using closed scenarios rather than freestyle texts, highlighting several advantages and disadvantages regarding usage statistics, simplicity versus complexity, bias and more. In addition, the use of ML algorithms and datasets to train the chatbot was relatively small and not sufficient for facilitating the cutting-edge abilities of AI-enabled chatbots.

Existing Blueprints: The blueprints which were collected in this deliverable describe the different components of AI-enabled chatbot services that are provided by Microsoft, Google, IBM, Amazon Web Services (AWS) and Facebook. Components that constitute different services that are based on the same technology were clustered together. Each component is described, and whenever relevant,

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

the possible contribution to emergency and disaster AI-enabled chatbots is suggested. The most popular components of the blueprints focus on connecting the chatbot to the various communication channels of authorities and first responders, the chatbot's logic ("brain and body"), various machine learning (ML) algorithms, accessibility tools and other technological capabilities, deployment and integration, data management, processing, quality assurance and storing, analysis, monitoring and insights, security and authentication and different types of data and datasets.

The Blueprint: The suggested blueprint employs the most popular cutting-edge, state-of-the-art technologies, as they appear in current blueprints, to the possible use of authorities and first responders, in AI-enabled chatbots for emergencies and disasters, to contribute to building societal resilience. The blueprint is based on two critical working assumptions. First, during the professional meetings conducted while working on this deliverable, many of the Ki-CoP members and ENGAGE's partners expressed concerns and potential barriers of using AI-enabled chatbots, which are most probably relevant to other authorities and first responders who were not represented in those meetings. These concerns ranged from not trusting the technological capabilities of the AI-enabled chatbot, through possible biases and mistakes, to being afraid that the public will not adopt this solution and will prefer human assistance. The second working assumption is the need to validate that the chatbot is doing what it should do satisfactorily. Therefore, the primary initial recommendation is that the chatbot collaborates with a human call centre that will verify that it is functioning correctly, providing accurate answers and maintaining the treatment of false information. The chatbot will be also monitored in later stages, but in a less intense way, than at the beginning.

Recommendations: In the last part of the blueprint, which is a recommendation for implementation, we suggest additional recommendations to overcome barriers and meet needs expressed by authorities, first responders, partners of ENGAGE and Ki-CoP members. We recommend adopting either the complete blueprint or adoption of several sub-methods that make the adaptation process more stepwise. In addition, we draw a roadmap for the full implementation of AI-enabled chatbots by authorities and first responders in emergencies and disasters to contribute to building societal resilience. The roadmap draws six milestones in five categories to fully execute the blueprint – technological capability, trust, user perspectives, information management, and budget and funds. Last, Appendix D presents ten implementation stages to help authorities and first responders strategise their implementation process of AI-enabled chatbots.

Conclusions and Contributions: Deliverable 3.2 has four significant conclusions and contributions. The first conclusion relates to the objective related to the ability to test or revise the assumption of ENGAGE that AI-enabled technologies can contribute to building societal resilience. Based on the design of the blueprint and the review of solutions and technologies, we concluded that the answer to this assumption should be positive, but with a cautious adoption of the blueprint suggested in this deliverable. The second contribution is that the blueprint refers to an AI-enabled chatbot, allowing emergency authorities to provide a contextual, online, and zero-delay response to the public before, during, and after emergencies. The third conclusion, and contribution, relates to innovative solutions for neutralising false messages during disasters, based on the rapid detection and tracking of trending misinformation on social media. The suggested blueprint highlighted the technologies needed to complete this mission, but with the necessary caution of only tracking and highlighting potential false information, leaving the last decision for human fact-checkers. Last, the fourth contribution is developing innovative solutions and citizens' engagement and transfer of knowledge from research and industry to the public, leveraging the project's suggested directions to relevant population groups, according to their specific needs and expectations.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

1 INTRODUCTION

1.1 SCOPE OF THE DELIVERABLE

The deliverable reports on the result of Task 3.2: "Exploration of innovative use of communication and social media technologies". It proposes new innovative suggestions for adopting Artificial Intelligence (AI) technologies, using AI-enabled chatbots, to improve societal resilience and citizens' engagement. The suggestions also consider potential implementation methods in different contexts and ways to strengthen different forms of societal resilience embedded culturally and socially in local contexts.

The deliverable reviews current AI-enabled chatbot solutions, their strengths and limitations, and different Machine Learning (ML) approaches concerning societal resilience. As part of the deliverable, we distinguish between AI, the intelligence, and ML, the means of learning. The deliverable also highlights the relationship between the instrumental nature of AI and the organic nature of societal resilience. The outcome of the deliverable is a suggested blueprint of a potential AI-enabled chatbot for emergencies and disasters. The document creates and explains a suggested blueprint while also referring to its roadmap and technological components, highlighting the vital connection with societal resilience of each of the components in the blueprint. The blueprint is the framework that defines the limit of the solutions, while the roadmap portrays its suggested adoption process, with defined goals to be completed.

The scope of deliverable 3.2 is limited to (A) examining the assumption that enhancing communication channels through AI can contribute to building societal resilience, (B) examining whether AI can replace or temporarily augment some of the roles first responders play (e.g., providing information in call centres) in all phases of disasters and emergencies, decreasing the need for human-assisted help, and, (C) setting the guidelines, through a blueprint, for future creations by authorities and first responders of AI-enabled chatbots for emergencies and disasters. In this deliverable, we present the state-of-the-art technologies adopted by authorities and first responders to create innovative AI-enabled chatbots that can work in emergencies and disasters and build societal resilience.

The task aims at suggesting a blueprint for an AI-enabled chatbot for emergencies and disasters, augmenting some of the roles of authorities and first responders. More broadly, it aims to suggest a design concept, identifying the various technologies that authorities and first responders can use to design such chatbots, including the datasets used to train these technologies. Authorities and first responders can use the following document as guidelines and suggestions for building such AI-enabled chatbot technologies.

The deliverable's intended readers are the ENGAGE Consortium (composed of 14 partners from 7 countries), the European Commission and project reviewers, and EU emergency authorities, first responders, NGOs, and the public, as stated in one of the aspects of the ENGAGE project, regarding the "transfer of research results, solutions and knowledge to the public to contribute to societal resilience".

1.2 GOALS

The main goal of deliverable 3.2 is to leverage AI capabilities to propose advances in designing practical and valuable solutions for communication and social media. In addition, this deliverable is also based on the needs and expectations of the public, as identified in WP1, and of authorities and first responders as identified in WP2. The identified needs and expectations and the perceptions expressed towards AI technologies' opportunities, risks, and limitations (ethical or technological)

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

have served as the building blocks to producing novel design directions that also address some foreseen issues.

1.3 Objectives

One of ENGAGE's objectives is to produce validated actionable knowledge on societal resilience by demonstrating the benefits and impact of the project solutions in different types of disasters. In addition, one of the desired results of the project is to find the best practices for communication and social media (R3). Under these, the results of this deliverable are a blueprint of an AI-enabled chatbot for emergencies and disasters, addressing questions of design and implementation of the AI chatbot. These primary objectives, which were stated in the scope of the deliverable, should also offer in addition:

(1) The ability to test or revise the assumption of ENGAGE that AI-enabled technologies can contribute to building societal resilience.

(2) To allow emergency authorities to provide a contextual, online and zero-delay response to the public before, during, and after emergencies.

(3) Innovative solutions for neutralising false messages during disasters, based on the rapid detection and tracking of trending misinformation (e.g., false rumours) on social media. The chatbot will not distribute messages actively but will answer the public's questions and provide unbiased information.

(4) Innovative solutions for citizens' engagement and knowledge transfer from research to the public, leveraging the project's suggested directions to relevant population groups, according to their specific needs and expectations.

1.4 Fit within ENGAGE and relation to other deliverables

D3.2 contributes to identifying new directions for innovative communication with citizens before, during and following disasters. Therefore, it will not just follow the assumption of ENGAGE regarding the contribution of AI-enabled technologies to building societal resilience but will also enable to revise the project's assumptions based on the findings.

The deliverable is related to other deliverables in WP1, WP2 and WP3, as follows:

- **D1.2 Local perceptions, risk awareness, needs and expectations about societal resilience** and **D1.3 Communication, Social Media and Societal Resilience**: both deliverables focused on the public's needs and expectations which are at the core of developing the new directions for innovative uses of AI.
- **D2.1 identification of needs and expectations of authorities and first responders**: the use of communication channels by authorities and first responders reflects some of their needs and expectations regarding improving societal resilience and suggesting possible directions for AI uses.
- **D2.2 and D2.3 identification of formal and informal solutions**: the solutions identified in D2.2 and D2.3 are the basis for the solution, which will be suggested in deliverable 3.2.
- **D3.1 Selection of promising results**: the process of selecting promising results relates to the new directions, which will be suggested as part of deliverable 3.2.

The deliverable also contributes to future WPs, 4 and 5, as elaborated further in the report.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

1.5 Acronyms and abbreviations

Table 1. List of terms.

Term	Explanation
Artificial Intelligence (AI)	Artificial Intelligence (AI) trains machines to understand and respond to requests, identify and learn languages, recognise objects, and improve decision-making by learning from examples, sometimes based on trial and error.
Machine Learning	Machine Learning (ML) is an application of artificial intelligence. If AI represents the "brain", then ML is the process that gives the technological systems the ability to learn from experiences and data, improve themselves and solve problems without explicitly programming them.
Supervised Learning	In supervised learning, all data is being labelled in advance. For example, a list of emergency guidelines is labelled with the specific relevant emergency and steps in disaster management.
Unsupervised Learning	In unsupervised learning, the algorithm needs to identify by itself the nature of the guideline
Semi-supervised Learning	Semi-supervised learning combines the previous approaches with a small amount of labelled data and the most significant amount of unlabelled data. The ML algorithms look for patterns in the data during the learning process to create rules and make decisions.
Reinforcement Learning	Reinforcement learning differs from supervised learning by not using labelled data but rather using authentication and correction, such as unsupervised data. Reinforcement learning is based on trial and error and constant updates of the algorithms based on the results of the learning
Weak AI	Weak AI, also known as Artificial Narrow Intelligence, focuses on simulating human behaviour instead of imitating it. Weak AI technologies focus on one task and performing it very effectively. For example, customer service AI-enabled chatbots give basic information on opening hours, stores locations, prices and more
Strong AI	In strong AI, also known as Artificial General Intelligence, the technology exhibits the abilities of human intelligence. Unfortunately, while strong AI is an ideal phase of AI-enabled technologies, this concept exists primarily in

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

	theory, while almost all AI technologies are weak AI.
Super AI	Usually served only to understand the potential of AI, super AI is the best of everything, beyond what the human mind can perform.
AI-Enabled Chatbot	AI-enabled chatbots allow users to interact with an application that uses AI for communication, text and/or voice-based. AI-enabled chatbots are widely used, for consumerism, healthcare, policymaking and more, accessible via various platforms, such as messaging platforms, social media, websites, and independent platforms.
NLP (Natural Language Processing) and NLU (Natural Language Understanding)	Natural language processing (NLP) gives computers the ability to understand the text and spoken words in much the same way human beings can. Natural Language Understanding (NLU), a subfield of NLP, involves transforming human language into a machine-readable format
Decision Tree algorithm	Decision tree algorithms are statistical models that predict a possible result based on given data. In some instances, decision trees serve as a simple route, through different hubs, to the desired result.
Misinformation	Misinformation is a type of false information representing understanding wrongly or publishing false information with no intention to harm. For example, wrong interpretations of a new study. In this case, if the reader misinterpreted the information, then the information he or she will disseminate will be false – but without the clear intention to mislead.
Disinformation	Disinformation, on the other hand, is intentional. In this case, pieces of false information are spread to mislead and create chaos. Following the example of the new study, a case of disinformation would be to transfer information about false findings from the study with the intention to mislead.
Fake News	Fake news has some overlap with disinformation, sometimes being referred to in the same context. However, a critical distinction of fake news from disinformation is that fake news information is presented in the form of news articles.
Blueprint	A blueprint is a technical design plan of a model.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

2 SIGNIFICANCE

2.1 CONTRIBUTION TO THE FIELD

Built on the public's needs and expectations, as identified in WP1, and the solutions identified in WP2, deliverable 3.2 suggests a blueprint and a roadmap that authorities and first responders can use to design and build future AI-enabled chatbots augmenting human-assisted help. In addition, this deliverable aims to contribute to a rapid response to the public's queries and decrease the load on information centres during adversities, thus increasing the capacity to manage the surge of seeking data before and during disasters effectively.

2.2 SPECIFIC CONTRIBUTION TO THE ENGAGE PROJECT

Deliverable 3.2 highlights the advantages and limitations of current AI-enabled chatbot solutions, analyses existing ML approaches, technologies and emergency organisations' datasets and their role in the blueprint, and delineates a roadmap of AI-enabled chatbots in building societal resilience. The deliverable does not demonstrate an actual development of such a chatbot but rather proposes new possilbe directions for the future development of such technologies by authorities and first responders.

The deliverable is part of selecting solutions for societal resilience, focusing on innovative AI-enabled chatbots and social media approaches. Therefore, the deliverable's conclusions and recommendations contribute to setting the criteria for selecting promising solutions and the catalogue of solutions (WP3), validating the solutions (WP4), and contributing to the knowledge platform (WP5).

The deliverable adopts the approach of a "weak-AI", simulating human behaviour and not imitating it, focusing on receiving queries from the public and providing immediate and accurate information in response. This approach was adopted in order to fit to the current state of AI technology, along with the available data sets that facilitate the learning process of AI-enabled chatbots. This approach also fits the types of services that chatbot aims to facilitate (e.g., spreading informaiton, answering questions), while it is not intended to preform, for example, live triage.

This deliverable, therefore, contributes directly to the third objective of project ENGAGE, regarding validating solutions: "to produce validated actionable knowledge on societal resilience by demonstrating the benefits and impact of the project solutions in different types of disasters".

In addition, deliverable 3.2 contributes to laying the groundwork for WP 5, as follows:

- **D5.4 Website and knowledge platform & D5.5 Knowledge and innovation community:** the analysis of current solutions, innovative ML approaches and datasets, along with the AI-enabled blueprint and roadmap, is included in the ENGAGE knowledge platform.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

3 SCIENTIFIC BACKGROUND

3.1 ARTIFICIAL INTELLIGENCE (AI) AND MACHINE LEARNING (ML)

In short, the term artificial intelligence, or AI, was developed in computer sciences to describe human-like intelligence and capabilities exhibited by machines, either software ()or hardware (robots), or the simulation of human intelligence in machines. According to IBM, "artificial intelligence enables computers and machines to mimic the perception, learning, problem-solving, and decision-making capabilities of the human mind"[1]. In other words, artificial intelligence is the process of training machines to understand and respond to requests, identify, and learn languages, recognise objects, and improve their decision-making process by learning from examples, sometimes based on trial and error.

Machine Learning (ML) is an application of artificial intelligence. If AI represents the "brain", then ML is the process that gives the technological systems the ability to learn from experiences and data, improve themselves and solve problems without explicitly programming them. In other words, it allows the machine to teach itself, using already existing datasets for training or learning from new data (Jordan & Mitchell, 2015; Mohri, Rostamizadeh & Talwakar, 2018).

Three dominant approaches for ML are supervised learning, unsupervised learning, and reinforcement learning. The process of observing datasets, examples and instructions is mutual to all learning methods. In the first, supervised learning, all data is labelled in advance. For example, a list of emergency guidelines is labelled with the specific relevant emergency and steps in disaster management. By contrast, in unsupervised learning, the algorithm needs to identify the nature of the guideline. Sometimes, semi-supervised learning is used. It combines the previous approaches with a small amount of labelled data and the most significant amount of unlabelled data. The ML algorithms look for patterns in the data during the learning process to create rules and make decisions. Last, reinforcement learning differs from supervised learning by not using labelled data but rather using authentication and correction, such as unsupervised data. Reinforcement learning is grounded on trial and error and constant updates of the algorithms, based on the results of the learning (Van-Engelen & Hood, 2020).

AI and ML employ various statistical measures, from probability to statistical reasoning, algorithms that allow, for example, natural language processing (NLP), giving computers the ability to process human language and computer vision that provides the ability to analyse images. Thus, AI-enabled services are partly replacing or assisting professions that in the past were highly dependent on human-assisted help, such as customer service, sales and more (Jordan & Mitchell, 2015; Mohri, Rostamizadeh & Talwakar, 2018).

Last, the literature distinguishes between at least three types of AI. The first, and the most common which this deliverable is aligned with, is what is called "weak AI" (sometimes called "narrow AI" or "artificial narrow intelligence"). Weak AI focuses on simulating human behaviour instead of imitating it. Weak AI technologies focus on one task and performing it very effectively. For example, customer service AI-enabled chatbots, which provide basic information on opening hours, stores locations, prices and more, are considered weak-AI (Al-Rifaie &, 2015; Liu, 2021).

Despite not being at the core of this deliverable, it is essential to present the other types to understand the weak-AI concept better. The second type is "strong AI" (sometimes called "general AI" or "artificial strong intelligence"), originally coined by philosopher John Searle (1980). In strong AI, the technology exhibits the abilities of human intelligence. Unfortunately, while strong AI is an ideal phase of AI-enabled technologies, this concept exists primarily in theory, while almost all AI technologies are weak AI (Al-Rifaie &, 2015; Liu, 2021).

[1] https://www.ibm.com/cloud/learn/what-is-artificial-intelligence

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

The third type is a highly theoretical concept of AI – "super AI" (sometimes called "artificial general intelligence", "artificial superintelligence", or "superintelligence"). Usually served only to understand the potential of AI, super AI is the best of everything, beyond what the human mind can perform (Zifu, 2016).

3.2 THE TECHNOLOGY OF AI-ENABLED CHATBOTS: USES IN EMERGENCIES AND DISASTERS

AI-enabled chatbots are sometimes referred to in the scientific literature as conversational agents (Almalki & Azeez, 2020). AI-enabled chatbots allow users to interact with an application that uses AI for communication, text and/or voice-based. AI-enabled chatbots are widely used in customer service, healthcare, and policymaking, accessible via various platforms, such as messaging platforms, social media, websites, and independent platforms. One of their advantages is their ability to support omni-channel communication, allowing the users to connect to it in a variety of platforms.

AI-enabled chatbots have an unlimited range of abilities that can be elaborated and developed. From basic chatbots that only provide answers based on a predefined list of questions to fully functional chatbots that can converse with the users, understand their queries or sentences, provide answers, and decide on the necessary action following a conversation. In accordance, they are based on a range of ML approaches, from supervised learning to unsupervised learning and more (Gupta, Hathwar & Vijayakumar, 2020; Følstad, Skjuve & Brandtzaeg, 2018). These abilities are possible due to cutting-edge technologies that keep developing and large-scale datasets.

While significantly developed in other fields, such as customer service, AI-enabled chatbots have been used only infrequently in emergencies and disasters. In theory, as Hofeditz, Ehnis, Bunker, Brachten & Stieglitz (2019) portray, AI-enabled chatbots can contribute significantly to emergency and disasters management. They can provide information to help citizens prepare for emergencies, or during the emergency may be used to give moral support, collect information, and detect changes in the environment that require special attention.

For example, in 2018 The Norwegian Refugee Council, Microsoft, NetHope and University College of Dublin developed the AI-enabled chatbot called "Hakeem"[2]. The goal of the chatbot was to provide education options in places which lacks the infrastructure of schools and universities. The chatbot, designed in the form of a character of an older brother, contributed to societal resilience by enhancing education levels in poor areas for a less fortunate population.

Table 2 presents a list of possible AI-enabled chatbot uses, in emergency management, as portrayed by Hofeditz et al. (2019). The table does not present what is being done or how to do it but rather describes several necessary actions that AI-enabled chatbots' blueprints should address. Therefore, its contribution to this deliverable is in laying the ground for some of the actions and activities that the suggested chatbot's blueprint will guide to.

[2] Global refugee crisis: Refugee turned humanitarian shares reasons for optimism - NetHope

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Table 2. a List of possible AI-enabled chatbots in emergency management. Taken from Hofeditz, Ehnis, Bunker, Brachten & Stieglitz, 2019.

Name	Type	Phase	Description
Message translation bot	Chatbot	All phases	Providing a real-time translation for outgoing and incoming messages via for social media channels and app
Customer service bot	Chatbot	All phases	Providing a chatbot guide for the app of the emergency organisation
Prevention activity bot	News bot	Prevention	With systems linked bot with informs about hazard reduction fires based on geolocations
Preparedness messenger bot	Chatbot	Preparedness	Facebook messenger bot, which answers frequently asked questions automatically
Preparedness social bot	Social bot	Preparedness	Scouting for issues in communities and chatting with members on social media platforms
Smart speaker bot application	Social bot	Preparedness / Response	Emergency warning application for speech assistances using meta data such as location and app data
Emergency report bot	Chatbot / Monitoring bot	Response	Providing additional automated channel for accepting emergency reports on social media
Response messenger bot	Chatbot	Response	Answering questions regarding ongoing disasters such as information about bushfires
Intelligence bot	Monitoring bot	Response	Collection and analysis of disaster relevant material from social media platforms by using keyword and hashtag search
Emergency warning bot	News bot	Response	Automated caution & advice messages regarding the disaster on different channels

The table suggests that message translation bots, providing real-time translation for messages, and customer service bots, guiding apps of emergency organisations, are crucial for all phases of emergencies and disasters. In the prevention and preparedness stages, prevention activity and preparedness messenger and social bots, answering frequently asked questions, links to relevant information are more relevant. Emergency warning and reporting bots, intelligent bots, collecting and analysing disaster relevant material from social media platforms are most relevant in the response phase.

Another significant advantage of AI-enabled chatbots over human response is their ability to provide service to an unlimited number of users with fewer resources simultaneously. In hurricane Sandy, for example, call-centres collapsed because of a surge of calls. While the infrastructure of AI-enabled chatbots can also collapse due to an overload, the overload bar is significantly higher (Crutchfield & Harkey, 2019; Smith et al., 2016). A similar event occurred lately in Hurricane Ida[3].

While the table above suggests the possible uses of AI-enabled chatbots in emergencies and disasters, and their potential contribution to building societal resilience, another crucial aspect of this issue is their limitations. The main question that arises is why AI-enabled chatbots are highly developed in other fields, but concerning emergencies and disasters, they are still underdeveloped? While the immediate answer could be related to the state of the technology, Madianou (2020) suggests that a better explanation is lack of data and implementation, which does not allow facilitating learning processes that can improve the AI-enabled chatbots' actions in emergencies and disasters, and instead, amplify the biases of current datasets. Moreover, major emergencies, in which the public is involved in direct communication with emergency centres, are rare. Compared to the number of people reaching out for non-urgent information, for instance on Covid-19, the dataset is very low. This may pose a challenge in training a chat bot to handle such rare, yet extreme, scenarios.

Another possible reason, suggested by Palanica, Flaschner, Tommandram, Li & Fossat (2019), is first responders' apprehension. An additional reason is the distrust of some of the public, i.e. their strive to receive a human response to the questions raised during emergencies (e.g., Abd-Alrazaq,

[3] Hurricane Ida Leaves Crisis Management Lessons Behind As Recovery Begins (forbes.com)

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Alajani, Ali, Denecke, Bewick & Househ, 2021; Cheng & Jiang, 2020). The three reasons are developed in the following sections.

3.3 AI, CHATBOTS AND SOCIETAL RESILIENCE

The applicability between AI technologies and specifically AI-enabled chatbots and societal resilience is not necessarily obvious. AI has a more instrumental technological character, focused on identifying problems and suggesting solutions. Societal resilience, on the other hand, has a more organic nature. AI is a technology, and resilience is a societal phenomenon. Therefore, the applicability of one on the other needs to be established carefully.

Societal resilience is related to the natural responses and behaviours of the public, understanding their needs, which they sometimes find hard to identify, targeting many variables, contextual and target factors (Vinuesa, Azizpour, Leite, Balaam, Dignum, Domisch & Nerini, 2020). Studies in this topic use AI technologies to predict disasters (e.g., floods and fires) and cope with these adversities (e.g., Saravi, Kalawsky, Joannou, Rivas Casado, Fu & Ment, 2019).

However, here we focus on the potential of enhancing societal resilience with AI technologies through transforming the communication process between authorities, first responders and the public. For example, providing accurate information through such a technology avoids the collapse of communication means during surges of requests and neutralising false information. Deliverable 1.2 portrayed the model with the different components that contribute to societal resilience. Communication was found as a strong predictor of societal resilience, including the frequent use of different communication channels. Deliverable 1.3, examining the public's communication needs, highlighted cognitive, affective, integrative and escapist needs that the public has in the communication process with authorities and first responders. The contribution of AI-enabled chatbots is in better fulfilling these needs, which are part of the communication variable in the prediction model of societal resilience.

In general, while AI-enabled chatbots were found to have an outstanding potential contribution to emergencies and disasters, the question regarding their possible contribution to societal resilience must be treated cautiously. For example, a person in an emergency seeks help, empathy, and a human voice in many cases. Failing to meet these needs will eventually decrease societal resilience. Current AI developments, however, provides an exceptional voice experience, that in several cases, was not identified by the users as a machine voice (Natale, 2020). A development which led the EU to add in its last AI regulations a demend to fully disclose the fact that a call is with an AI[4].

In addition, AI-enabled chatbots are learning-based, and in many cases, they cannot answer the questions – a situation that creates frustration, another factor that can decrease societal resilience. Therefore, the adoption process of AI-enabled chatbots should consider both positive and negative effects on societal resilience.

In her article on the societal implications of AI, Ignatidou (2019) suggests that AI-driven personalisation in digital media has important societal resilience implications. She claims that such technologies allow, for example, personal adaptations of the information provided to the public, addressing their needs, and contributing to societal resilience. Kertysova (2018) adds a broader role for international organisations in building policies for AI in building societal resilience.

In a report by the European Commission Joint Research Centre (JRC) about AI, the EU perspective on the relationship between AI technologies and societal resilience is presented. The general claim presented in the report is that it is crucial not just to shape the development of technology to benefit society but also to prepare the institutions, policies, people, and society to become more flexible, adaptable, and ready to transform. They call for a holistic, complex and system view approach, taking into consideration the advantages of AI technologies but also their possible negative

[4] EUR-Lex - 52021PC0206 - EN - EUR-Lex (europa.eu)

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

implications, such as reducing the number of jobs in emergency call centres (Annoni, Benczur, Bertoldi, Delipertrev, De Prato, Feijoo & Junklewitz, 2018).

AI-enabled technologies, therefore, must have a clear connection to societal resilience, as concluded by the JRC report. In order to connect AI-enabled technologies with the idea of societal resilience, Alessi et al. (2018) suggest matching each parameter in an AI-enabled system with what they define as "resilience indicators", identifying the known variables that can contribute to societal resilience. Therefore, while measuring the societal or individual resilience of the public or individuals, one can explain how the system's work led to improvement in the resilience indicators. For example, matching the conversational ability with indicators of well-being, information acquisition and sense of preparedness. The contribution of AI-enabled technologies and societal resilience can also be found on the macro-level. For example, regarding preparedness, between the share of people in a society employed in science and technology (or in AI companies in particular), and the preparedness of the society for emergencies and disasters. Based on that, the JRC report concludes that "looking at AI from a resilience perspective does not only enrich our understanding of, and preparedness for AI, but also advances the analysis of resilience" (2018, p. 119).

In addition to the connection to societal resilience, in April 2021, the European Commission published the regulatory framework proposal on AI, including a reference to AI-enabled chatbots. The proposed regulations address the risks of AI and its advantages and suggest Europe's role in leading such technologies. Furthermore, the regulations suggest comparing the advantages of AI technologies within four levels of risk: unacceptable risk, high risk, limited risk and minimal risk, with a four-step approach for high-risk AI technologies, before adopting them[5].

Previous studies show that AI-enabled chatbots are primarily popular in health emergencies and disasters (e.g., Almalki & Azeez, 2020; Battineni, Chintalapudi & Amenta, 2020). In addition, they are used as online self-symptoms tests, information providers and Q&A tools. More recent uses of AI-enabled chatbots are suggested to replace simple tasks such as scheduling healthcare appointments, issuing reminders and improving medication adherence (e.g., Almalki & Azeez, 2020).

The most recent example is AI-enabled chatbots in aiding the public to receive accurate information about COVID-19. AI-enabled chatbots reached a large scope of the population in the world. For example, the World Health Organisation (WHO) has developed a messaging-based (e.g., WhatsApp, Viber, Facebook Messenger) to help the public find answers to their questions about COVID-19 and to get accurate information about the pandemic in an environment which was highly affected by mis- and disinformation. In addition, the US Centre for Disease Control (CDC) built a similar web-based chatbot, based on Microsoft's Healthcare Bot, following a closed-ended pool of answers to self-diagnose coronavirus symptoms[6]. According to the WHO, in the first month of the chatbots, 12 million users used them. By the end of 2020, it reached over 4.2 billion users worldwide, with similar numbers using the CDC chatbot[7].

However, throughout time, AI-enabled chatbots have been used in various types of applications. In 1966, as one of the first developments in this field, ELIZA, a text-based chatbot, was developed by Joseph Weizenbaum to imitate a psychotherapist (Natale, 2019). Rescue, in 2018, is a chatbot that helps users to report emergencies[8]. Other chatbots were also developed in the last years, covering various topics, primarily to provide information in the preparedness stage of emergencies and aid users in reporting about emergencies (e.g., Clark, Fox & Lappin, 2013; Tsai, Chen & Kang, 2019).

In the following sub-sections, we review three dominant contributions of AI-enabled chatbots as highlighted in the literature: providing fast and immediate information, supporting the work of

[5] EUR-Lex - 52021PC0206 - EN - EUR-Lex (europa.eu)
[6] CDC Coronavirus Chatbot Advises If You Need Medical Care | Digital Trends
[7] Chatbots provide millions with COVID-19 information every day, but they can be improved - here's how | World Economic Forum (weforum.org)
[8] Rescue.io: A Chatbot Solution for Emergency Situations | by Justin Clegg | Square One Labs | Medium

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

emergency call centres and coping with false information and fake news. Based on the literature review, these three subsections are not a scientific typology but a division we suggest.

3.3.1 Immediate information

One of the critical factors contributing to societal resilience is the ability to rapidly spread information. This was supported by the findings of Deliverable 1.3, which focused on social media and societal resilience. Participants participating in the survey conducted for the deliverable rated their cognitive needs (information) and unidirectional flow of communication needs (fast information) in the first place, compared to other needs, such as affective and escapist, which were rated last. In addition, the ability of authorities and first responders to transfer the information quickly to the public, including answering questions that arise, contributes to the public's ability to cope with emergencies and disasters and thus to societal resilience.

However, as Kertysova (2018) notes, when it comes to the ability of the public to receive immediate answers to their questions, there are several challenges. For example, the public is unaware of the various sources of information and how busy they are. Therefore, they could address busy sources while not contacting a better and less busy source of information. Furthermore, while social media is a fast and adaptable tool for disseminating information, it cannot organise information in ways that websites and offline sources can. Therefore, while users may tend to go first to social media channels to get information, authorities and first responders also favour these tools but they do not always provide the information in its best form for the public due to its unorganised nature.

Therefore, to answer the public's needs, as some of them emerged in the survey presented above, AI-enabled chatbots have a double contribution. First, they allow for better organisation of the information than social media, enabling the users to ask about a topic and then immediately get the information (Hasan, Rizvi, Jain & Huria, 2021). Second, AI-enabled chatbots can scan the different sources of information online, deliver them to the users, and answer their requests (Almalki & Azeez, 2020). For example, if an individual is unsure whether there was a change in an emergency regulation, he/she may approach the chatbot. The chatbot, connected to the various datasets of online regulations, can check, and forward the relevant updated regulation.

3.3.2 Avoiding the collapse of means of communication during request surges

During routine times, emergency call centres can facilitate a reasonable number of calls, for providing information, aiding local emergencies or any other request. However, a surge of calls to emergency call centres during emergencies and disasters may collapse the infrastructure (Carenzo, Costantini, Greco, Barra, Rendiniello, Mainetti & Cecconi, 2020; Hrabi, 2020). For example, during natural disaster events or extreme weather conditions, many people call the emergency lines, check for their family and friends, and this high volume of calls can lead to the collapse of phones lines, resulting in the emergency teams not being able to assist those in danger. In addition, call centres may collapse even due to technical failures, cyber-attacks, or other reasons. For example, in June 2021, the emergency phone lines in France collapsed for seven hours without any particular extreme load of calls. This situation resulted in three people who could not reach the emergency lines and get help[9].

Such situations are referred to in the literature as "infrastructure resilience" (MacKenzie & Barker, 2013). For example, MacKenzie (2004) analysed the breakdown of the emergency services phone lines after the collapse of the world trade centre buildings in 2001, pointing to the fact that the infrastructure could not facilitate such an amount of calls. Extensive efforts have been invested since then. However, since infrastructures are not necessarily more substantial, large-scale disasters still

[9] The emergency lines collapsed for 7 hours, civilians died - P4World.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

threaten to create an overload on call centres and emergency services. In this case, the internet, which is a more resilient infrastructure, continues working (Simon, 2013), and it is the basic technology of AI-enabled chatbots.

Some research has concluded that the work of emergency and first responders call centres should be replaced or augmented by social media solutions. For example, during a shootout in Orlando in 2016, the nightclub at the centre of the violent event used Facebook to alert people to evacuate (Richardson, 2019). During hurricanes Harvey and Irma in 2017, citizens used social media to contact the rescuers when the phone infrastructure was down (Heires, 2017). Other platforms have been developed to closely monitor social media content, use machine learning algorithms and collecting data and insights on the spread of disasters (Imran, Castillo, Lucas, Meier & Viewg, 2014). Based on such a case, Chaudhry & Yuskal (2019) suggested replicating similar abilities of the work of 911 services, allowing alternatives to the flow of information from emergency services and rescuers to the public. As they claim, this will reduce the workload on the call centres and spread the requests on more diverse channels. They developed the intelligent public safety framework, illustrated in Figures 1 and 2, allowing emergency services to handle non-emergency calls for information and SOS messages and alerts using a public safety bot.

Figure 1. : Smart Public Safety Framework. Taken from Chaudhry & Yuskal, 2019.

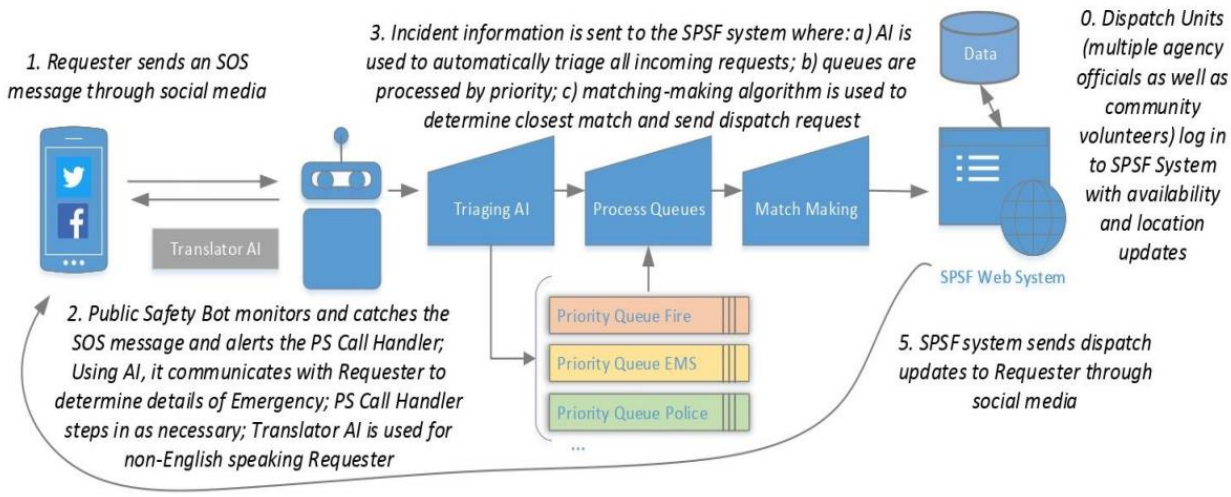
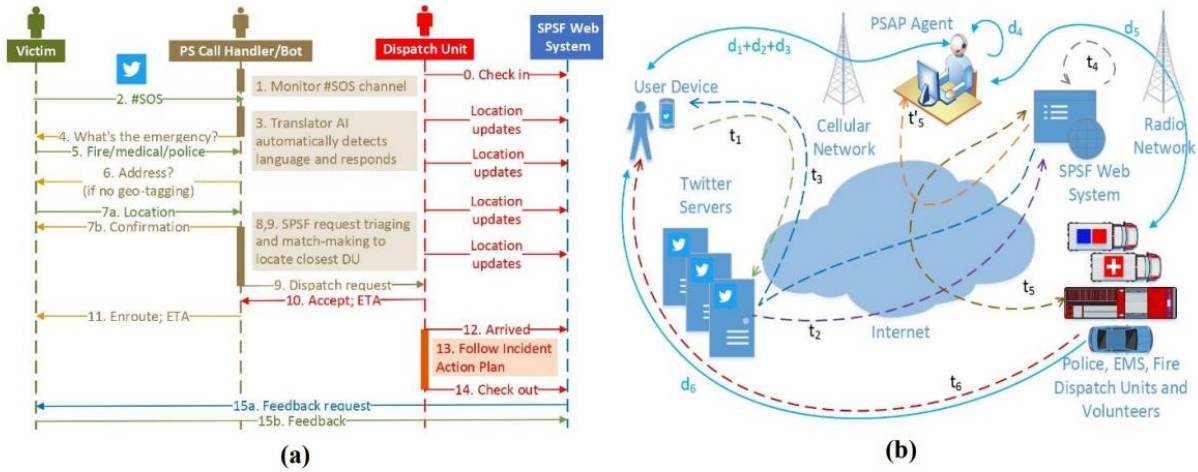


Figure 2. Smart Public Safety Framework's (SPSF's) protocol for sending distress signals. Taken from Chaudhry & Yuskal, 2019.



The bot in the framework is a chatbot integrated with APIs such as Twitter's and Facebook's to track relevant hashtags. If calls overload call centres, the chatbot is trained to ask a series of questions. It also can analyse the calls in real-time and estimate the location of the mobile subscriber and queue processing using AI capabilities. However, it is essential to note that such infrastructure is

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

also subjected to collapse and external attacks. Therefore, they should serve as alternative communication methods supporting conservative communication channels and not entirely replacing them. Hence, allowing alternatives in the possible event of infrastructure collapse.

3.3.3 Detecting false information in social media and neutralising fake news

One of the significant challenges to societal resilience is false information and rumours (Kertysova, 2018). According to Shu, Wang, Lee & Liu (2020) and Stahl (2006), there are three main types of false information: misinformation, disinformation, and fake news. Misinformation is a type of false information, which is not valid, representing understanding wrongly or publishing false information with no meaning to harm. For example, wrong interpretations of a new study. In this case, if the reader misinterpreted the information, then the information he or she will disseminate will be false – but without the clear intention to mislead. The transfer of disinformation, on the other hand, is intentional. In this case, false information is spread to mislead and create chaos. For example, confused people base their decision on this false information and may act in an ineffective and even dangerous way. The European commission's high-level expert group on fake news and online disinformation defined disinformation as "false, inaccurate, or misleading information designed, presented and promoted to cause public harm intentionally or for-profit" (European Commission, 2018).

Last, fake news has some overlap with disinformation, sometimes being referred to in the same context. However, Rogers & Niederer (2020) add that a critical distinction of fake news from disinformation is the form of news setting. According to other definitions, fake news includes fabricating information that mimics news media content forms but lacks the media's editorial norms (Lazer, Baum, Benkler, Berinsky, Greenhill, Menczer & Zittrain, 2018). In addition, it is essential to emphasise that fake news is not about unintentional reporting mistakes. It is also not about rumours that have nothing to do with the form of news articles. The last two are part of misinformation and disinformation[10].

In the last years, a new type of false information has arisen, "deepfake". Based on traditional actions of faking content, deepfake is the use of AI and ML technologies to manipulate existing or creating new audio-visual content with the primary intention to mislead (Westerlund, 2019). For example, faking photos or videos of celebrities or events. Recent studies have shown that the technology of deepfake have progressed greatly, so in many cases users could not uncover the fake (Dolhansky et al., 2020), and even computer algorithms sometimes fail to do so (Frank, Eisenhofer, Schönherr, Fischer, Kolossa & Holz, 2020 ;Maksutov, Morozov, Lavrenov & Smirnov, 2020).

False information, through all the different types, challenges societal resilience in several dimensions – it prevents certain parts of the society from receiving accurate information that they need for managing emergencies and disasters, it misleads them regarding what to do to be prepared, act and recover from emergencies and disasters, and lowers their morale. Moreover, there is much more sensitivity to false information in emergencies since people are under pressure, more stressed, and lost some of their ability to be critical – subjected to more "informational attacks". Therefore, it is crucial to neutralise cases of false information as one of the measures to contribute to societal resilience (Kim, Lyu & Gong, 2020; Puildo, Villarejo-Carballido, 2020).

Project "Builders", a sister DSR01 EU-funded project, partly examined in deliverable 1.4 the relationship between misinformation, disinformation, and societal resilience. Deliverable 1.4 of the project showed that people who do not use multiple news sources and that are less skilled are the most vulnerable to online misinformation. In addition, according to the report, users did not just fail in identifying false information, but also identified official warnings as spam or false information. Based on these findings, one of the recommendations of the project, that this deliverable follows,

[10] What is Fake News | Center for Information Technology and Society - UC Santa Barbara (ucsb.edu)

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

was to develop tools, along with skills, to evaluate the credibility of social media, but not solely, information.

Following that, AI-enabled technologies, and chatbots, can play a double role in treating false information. On the one hand, ML algorithms facilitate AI's high ability to detect false information and remove it, even with 100% success (e.g., Aphiwongsophon & Chongstitvatana, 2018; Mahabub, 2020; Patwa et al., 2021). Nevertheless, on the other hand, put in the wrong hands or even resulting from non-careful design, AI-enabled chatbots can be the source of spreading false information (Ahmen, Aliabouh, Donepudi & Choi, 2021). In 2016, Tay, a Twitter conversational chatbot developed by Microsoft, was involved in a series of racist tweets – just because its learning capabilities were highly based on what other people tweeted to him, which biased him towards racist language (Coeckelbergh, 2020). The same algorithms that help AI-enabled chatbots identify fake news can cause them to spread fake news, just because the datasets they use may include fake news. In this case, the chatbot will train itself based on the fake news bias (Sample et al., 2020).

The problem is, of course, much more complex, both from the technical and the ethical perspectives. On the technical side, recognising information as false can be wrongly classified based on a bias (Figueira & Oliveira, 2017; Hakak, Khan, Bhattacharya, Reddy & Choo, 2020). For example, when arranged online, protesters flag information as "fake" to remove it. Alternatively, when algorithms identify information as fake, only because of criticism, that later is clarified as accurate. The ethical challenge is reflected in the decision of what is true and what is wrong for society (Stroud, 2019). This has led the EU to develop ethical guidelines for AI use, addressing principles of respect, autonomy, prevention of harm fairness and explicability[11].

AI-enabled chatbots, relying on ML algorithms for false information, are based on a process of fact-checking. According to "Duke's reporters' lab" of Duke University, fact-checking projects developed to identify false information are trendy globally (Graves, 2018). However, while most of these fact-checking initiatives are conducted manually, Kertysova (2018) argues for the need to automate this process as the volume of information grows. In the last years, big technological companies, such as Facebook, Google, and other organisations that base their activity on spreading information, heavily invest in developing ML algorithms for automating this process[12]. However, while the "under the hood" access to these algorithms remain closed, they have also been criticised for being biased, for example, by ruling out information spread by specific groups as "fake news" based on political orientations (Sumpter, 2018).

3.4 How do users perceive the use of AI-enabled chatbots?

The wide use of users of AI-enabled chatbots is best exampled in the COVID-19 pandemic, as mentioned in previous sections. There are many users addressing those chatbots, querying for information and receiving answers for their questions. Therefore, it is notable that the massive use of such technologies testified about their need and the potential trust rates of such technologies by users.

In a recent analysis, Youn & Jin (2021) analysed the advantages and limitations of AI-enabled chatbots, delineating what uses AI-enabled chatbots can be reliable for and for what uses are still not acceptable. They claimed that the area is still underexplored despite the high use of AI-enabled technologies in user interactions and customer relationship management (CRM). A between-subjects experiment showed that AI-enabled chatbots tend more to build virtual assistantship relationships with the users rather than virtual friendships, which received lower scores of para-social interaction, satisfaction and trust, and less behavioural intentions instructions of the chatbot.

[11] Ethics guidelines for trustworthy AI - Publications Office of the EU (europa.eu)
[12] Facebook using machine learning to fight fake news | Internet of Business

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Moreover, the scientific findings point to several critical issues regarding the general user experience of developing AI-enabled chatbots, particularly concerning emergencies and disasters. Several studies highlight the significant limitations of AI-enabled chatbots. First, they are usually less preferred as a communication method when a human alternative is presented, regardless of the topic (Lei, Shen & Ye, 2021). Second, their trust scores are lower than human sources (e.g., Aoki, 2020). Third, they relate to fewer needs that users wish to fulfil (e.g., Ashfaq, Yun & Loureiro, 2020; Brandzaeg & Følstad, 2018), and the latest studies still find lower acceptance scores in adoption intentions (e.g., Laumer, Maier & Gubler, 2019).

On the other hand, these data were presented in the context of comparing AI-enabled chatbots to human service. Several optimistic results can be found in the scientific literature regarding the possible contributions of AI-enabled chatbots on top of what human first responders can offer. For example, especially for younger populations, AI-enabled chatbots are perceived as a good tool for basic information that saves time (De Cicco, Silva & Alparone, 2020) or offers a positive, friendly interaction (Thies, Menon, Magapu, Subramony & O'neill, 2017). They are also perceived to be more available since they do not have operating hours like call centres (Van-Wezel, Croes & Antheunis, 2020). In addition, the user rates of AI-enabled chatbots tend to be higher if a "fallback" option is available, meaning that it is possible to choose when to transfer the conversation to a human operator (Hill, Ford & Farreras, 2015).

Another aspect related to the users' perspective is the one of digital literacy. On the one hand, AI-enabled chatbots can be a sophisticated technology, that excludes the less connected and less knowledgeable. However, as several examples from the literature showed, the chatbot can be accessed through multiple platforms, including those who address digitally illiterate populations.

These previous findings suggest that AI-enabled chatbots concerning emergencies and disasters have many barriers in the public's eyes. Barriers that may prevent the adoption process of AI-enabled chatbots. However, as shown by Nadarzynski, Miles, Cowie & Ridge (2019), most users would be receptive to using AI-enabled chatbots, although hesitancy regarding this technology is likely to compromise engagement. Therefore, it is recommended to adapt a framework that suggests where AI-enabled chatbots can be integrated to gain the public's trust moderately. This can be relevant since while they have a more resilient infrastructure, AI-enabled chatbots can also collapse, leading to a lack of communication or be subjected to cyber-attacks (Verma, Chandra & Joshi, 2021; Yamin, Ullah, Ullah & Katt, 2021).

3.5 Other applications of AI-enabled technologies for communication and social media

Although the focus of deliverable 3.2 is on AI-enabled chatbots, other AI applications were identified in the literature for communication and social media in emergencies and disasters. One of these applications, which is also at the core of the previous discussion of fake news, is social media listening, using AI capabilities to monitor social media channels for tags, mentions, topics and specific keywords (Rao, 2016). In this case, concerning emergencies, social media listening is conducted, identifying tags, keywords and topics, and employing them with a range of algorithms, such as sentiment analysis, used to track the sentiment of content, to produce insights that can highlight necessary actions that should be done (Herrera, Majchrzak & Thapa, 2021).

The same implies false news detection (e.g., Barojan, 2021; Luccioni, Pham, Lam, Aylett-Bullock & Luengo-Oroz, 2021), mentioned in the previous section. In addition, when AI is applied to the large flow of incoming reports during an emergency, this might help to better understand the gravity of a situation. For example, a single report of "someone looking suspicious and maybe had a gun somewhere" may not be enough to raise the necessary awareness by a single emergency operator, but if there were similar reports from multiple citizens in the same area this could be identified by AI and raise the appropriate attention.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Another application provides a continuous feedback loop for authorities and first responders about the public's reaction to emergency-related content (Hayes, Britt, Evans, Rush, Towery & Adamson, 2021). For example, how the public understands or reacts to new guidelines, how they experience the emergency, how they respond to a speech, and other types of content. In this case, similar analyses using ML algorithms are conducted that provide feedback to improve how information is presented to the public concerning the different phases of emergencies.

3.6 Societal resilience, needs and expectations

Deliverables 1.3 and 2.4 analysed the communication process between authorities and first responders, and the public. The first focused on the communication needs and expectations of the public, and the second, on the communication strategies of authorities and first responders. In addition, Deliverables 1.2 and 2.1 analysed the needs and expectations of the public, on the one hand, and authorities and first responders, on the other hand, about societal resilience and how to improve it. Several findings found in these deliverables are essential to our work on innovative directions using AI-enabled chatbots to build societal resilience, as elaborated in the following two sub-sections.

3.6.1 The public needs and expectations

In the survey conducted in Deliverable 1.2, several variables were found to predict individual and societal resilience: coping skills, digital literacy, sense of responsibility, individual preparedness, risk awareness, age, communication needs, level of education, trust and social norms and communality. In addition, individual resilience was a predictor of societal resilience and vice versa.

Deliverable 1.3 identified and elaborated on the complex matrix of the public's communication needs from cognitive to the affective, integrative, escapist, unidirectional and multidirectional flow of communication needs. However, in the literature, most of the focus AI-enabled chatbots, especially emergency and disaster management, is given to the fulfilment of cognitive and unidirectional flow of communication needs, meaning to the ability of the chatbots to provide accurate and trustworthy fast information.

However, as deliverable 1.3 highlights, before, during and after emergencies and disasters, the public has also other important, sometimes complex, communication needs. They want to feel better, to be able to connect to other members of society, to submit information and queries and not just to receive information passively. In addition, deliverable 1.3 showed that these needs exist in parallel, obligating authorities and first responders to address them simultaneously in the same communication process.

These needs, found in deliverables 1.2 and 2.4, may contradict the use of AI-enabled chatbots and must be implemented into the process. For example, studies showed that interacting with AI-enabled chatbots decreased the sense of a better service. Moreover, even though several studies highlighted the ability of AI-enabled chatbots to create para-social interactions with their users, the overall affective experience is still lower than in human-to-human interactions. On the other hand, other examples of interactions between AI-enabled chatbots to human-beings showed the opposite. For example, a 2018 conversation between Google Assistant to a hair dresser, being able to make a complex call and schedule an appointment[13] or 2016 McDonalds & Microsoft AI drive-through capabilities[14].

[13] Google Assistant calling the hairdresser for an appointment - YouTube
[14] McDonald's brings cognitive to the drive-through - Software - iTnews

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Yet, Failing to answer this matrix of needs can harm the potential of AI-enabled chatbots to be effective in emergencies and disasters and build societal resilience.

3.6.2 The authorities and first responders point of view

Deliverable 2.1 identified what different members of emergency organizations and government officials expect from society to handle a crisis better. The findings of this deliverable indicate that authorities and emergency responders value the involvement of the members of society and volunteers in handling disastrous situations and look for various communication channels to disseminate information and engage the public.

Deliverable 2.4 completed this view by focusing on how authorities and first responders manage the communication process from their side. Here, also, we identified a complex picture regarding possible barriers. From the size of organisations, preventing medium and small organisations from adopting such innovative measures, through lack of communication guidelines that allow the implementations of such technologies, to individual objections, thinking that such technologies will not work.

Therefore, the implementation process of AI-enabled chatbots should address these possible barriers. First, it should address the barriers and needs identified in deliverables 2.1 and 2.4. Second, and not less critical, the roadmap to adopting AI-enabled chatbot technologies should be moderated to fit the preparedness level of every organisation. For example, organisations that never adopted social media strategy cannot start by implementing and fully developing such chatbots, compared to organisations that tried AI-enabled chatbots in the past, but the solution was not applicable to their needs. Third and last, the suggested future directions for designing such chatbots should consider the current phase of chatbots use by organisations. New directions cannot offer a far more sophisticated design than what is currently available but should fit into what can be considered as the next steps.

3.7 Summary of the literature and critical points for the current project

The scientific literature review introduced the areas of AI and ML, highlighting their potential contribution, and not less important, their connection to building societal resilience. While this deliverable relies heavily on AI and ML, its centre is still societal resilience. The aim at the core of this deliverable is not how to create state-of-the-art technologies, but in what ways it is possible to build societal resilience. Therefore, after introducing the technological fields, the following discussion focused on their relationship with societal resilience.

In the first stage, we reviewed the field of AI and ML. We defined AI as a process of trying to imitate human reasoning and learning abilities. Afterwards, we introduced ML as an application of AI, elaborating on the different approaches for learning: supervised, unsupervised, and semi-supervised. We also introduced the different types of AI: weak, the most popular, solid and the only theoretical "super-AI".

Later, in the second stage, based on the literature, we defined the limitations of what AI can do and cannot do and the borders of this deliverable – what we intend to do and what are the boundaries, even if the technology allows more sophisticated solutions. So, for example, despite initial abilities to provide emergency treatments and triage of AI-enabled chatbots, we do not recommend, at this stage, to fully employ such capabilities in the work of the chatbot, given the status and adoption level of authorities and first responders.

In addition, it should be emphasized that the focus is on current development in AI, which its progress is swift and exponential – two generations of progress in every year. Therefore, limitations that are relevant while writing these lines might be less relevant a few months later, whilst

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

introducing new and advanced capabilities. Finally, based on Hofeditz et al. (2019), we reviewed the current use of AI-enabled chatbots in emergency management. The review gives a clear understanding of what is being done in an emergency and societal resilience and what could be considered too ambitious for the preparedness level of the field. Alongside the limitations, we also highlighted the potential that AI-enabled chatbots have in building societal resilience.

In the third stage, we tried to define the theoretical connection of an instrumental technological field, AI and ML, to societal resilience, which is more organic. Previous studies emphasised that societal resilience set the goals for what problems AI should solve and not vice-versa. We introduced a system-view approach that clarifies those connections between societal resilience indicators and components of AI-enabled systems and chatbots. We showed that AI-enabled chatbots are used primarily on health systems but gave a few examples of other types of emergencies.

We then focused on three possible dominant contributions of AI-enabled chatbots to building societal resilience, first, giving immediate information. We elaborated on the contribution of providing immediate and non-delayed information for the public and from the public to authorities and first responders and discussed the possible ways to achieve such bidirectional communication through AI-enabled chatbots. Second, in avoiding infrastructure collapse, especially call-centres, by being able to mimic and augment the work of first responders. Third, based on tracking fake-news technologies, we showed ML algorithms' contribution to providing accurate information and fighting false information that can harm societal resilience.

Last, based on deliverables 1.3 and 2.4, we highlighted the communication needs of the public, as identified in a cross-national survey, and the communication strategies of authorities and first responders. We called for AI-enabled chatbot solutions that consider the needs of the public and work structure and perceptions of professionals in authorities and first responders.

In the following sections, we will draw the process based on the literature for suggesting new directions for designing AI-enabled chatbots for building societal resilience.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

4 PROCESS

In order to create the blueprint of the AI-enabled chatbot, we followed a process that included reviewing existing solutions, systematic analysis and review of the scientific and grey literature, consultation, reviewing solutions, algorithms and datasets, and suggesting the blueprint based on existing blueprints.

4.1 REVIEWING EXISTING SOLUTIONS

One of the essential steps in suggesting innovative directions and design for AI-enabled chatbots for authorities and first responders is to map the existing solutions on the field. A solution was defined in deliverable 2.2 and deliverable 2.3, "to refer to this set of means that emergency responders and authorities can use and implement to reach out to the public and improve the interaction with them" (Deliverable 2.2, p.11). In this case, the solutions were the set of AI-enabled means focused on chatbots. It is essential to understand the current state-of-the-art interface between existing technologies and the organisations that use them: what technologies do organisations already use? What is their adoption level in these technologies? How ready are they to adopt innovative technologies? In addition, it was essential to map what chatbots and capabilities already exist and what problems concerning societal resilience they solve.

We, therefore, collected and reviewed current and past solutions of AI-enabled chatbots of authorities, first responders and other types of organisations, which focus on different types of emergencies and disasters (e.g., health, natural disasters, security events). Our initial sources of information for collecting the solutions, as elaborated below, were the scientific literature, grey literature and the snowball sampling method, starting with ENGAGE partners and the Ki-CoP organisations. To complete the list, we also searched public sources (e.g., Bing, Google). Several solutions appeared in more than one collecting method.

The selection criteria for considering an AI-enabled chatbot as part of the corpus were comprehensive, and the list of solutions included 45 examples. Therefore, we included AI-enabled chatbots used by official authorities and first responders to deal with emergencies and disasters in all available platforms (e.g., mobile, web, call-centres). The list of solutions appears in Appendix A.

4.1.1 SCIENTIFIC & GREY SYSTEMATIC LITERATURE ANALYSIS AND REVIEW

We used the following academic databases to locate articles and other scientific material about the use of AI-enabled chatbots concerning societal resilience in general, and emergencies and disasters and particular: Ebsco, Google Scholar, JSTOR, Medline, Proquest Central, Pubmed, ScienceDirect, Scopus, Sociological Abstracts and Web of Science. We also used Bing/Google as general search engines for grey literature.

We used combinations of the following keywords to locate for relevant sources: "AI"/"Artificial Intelligence", "Chatbots", "ML"/"Machine Learning", "Societal Resilience", "Emergency"/"Emergencies", "Disaster"/"Disasters", "Crisis"/"Crises". We included all types of content: empirical, theoretical, case studies and more if it referred to an actual chatbot example. We also included theoretical suggestions for models, even if they were not implemented. Thirty-eight relevant solutions were located by this method.

4.1.2 SNOWBALL, ENGAGE PARTNERS AND KI-COP ORGANISATIONS

After the literature search, we also collected a list of solutions from ENGAGE partners, Ki-CoP organisations, and other relevant sources. Those were collected through personal emails, meetings,

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

and personal references to other informants. By this method, we added five chatbots that were not on the initial list. Thus, the final list includes 45 AI-enabled chatbots, which is included in appendix A.

4.1.3 Reviewing the solutions

After we compiled the list of solutions, we reviewed them using predefined criteria. The list of criteria and reviewing template is attached as appendix B. The reviewing criteria included the name of the chatbot. This organisation developed/used it, types of crises it handles, the list of datasets the chatbot uses (e.g., call scenarios, diseases symptoms), the technology it uses (i.e., the platform, the algorithms, types of chat scenarios and other services it uses), the approach of the chatbot to societal resilience (i.e., providing information actively or only responding to queries/questions, integration with other services, referring the user to other sources of information), its advantages, disadvantages and gaps/weaknesses of operation.

4.2 Algorithms

Algorithms were also classified in each chatbot which used them whenever they were transparent, and they are also included in Appendix A.

4.3 Identifying relevant datasets

One of the main findings of deliverables 2.4, focusing on the communication strategies of authorities and first responders, was that many organisations hold several datasets with information about the activity of authorities and first responders during emergencies. These datasets can be, for example, treatment or apply event protocols relevant for public emergency scenarios (e.g., how to diagnose a disease and conduct a triage, what steps to follow when handling an emergency event), call scenarios for call centres (i.e., mass alerting scenarios based on incoming event data), transcription of calls, guidelines and more. Such datasets can significantly contribute to the work of AI-enabled chatbots in emergencies and disasters for building societal resilience. Therefore, for each solution that we analysed, we identified, if it was available, the relevant datasets which were used to operate or train the chatbot, and they are included in Appendix A.

4.4 Identifying existing blueprints of AI-enabled chatbots

Compared to other implementation processes of AI-enabled technologies, the implementation process of an AI-enabled chatbot is considered more straightforward. While most AI technologies require deep knowledge and experience professionals, AI-enabled chatbots can be implemented using published services and technological capabilities. This can be done by following a blueprint.

A blueprint is a technical design plan of a model. In this deliverable, we are not suggesting a blueprint for a specific model of a chatbot, but a general suggestion of the possibilities that authorities and first responders have in using such technologies and the various components they should include. Therefore, as part of the identifying and reviewing stage, we mapped the existing blueprints of AI-enabled chatbots and their services, which authorities and first responders can use. We reviewed the blueprints of AI-enabled chatbots of big providers, such as Microsoft, IBM, Amazon Web Services (AWS), and others, identifying the components and services they use to help authorities and first responders (e.g., classification algorithms, data processing services). The complete list appears in Appendix C.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

4.5 Creating the blueprint

In the last step, after identifying and analysing the existing solutions and their components, we generated a suggested blueprint for designing an AI-enabled chatbot for emergencies and disasters. The blueprint suggests directions for the design of such chatbots by authorities and first responders. While the blueprint offers the most cutting-edge technologies that authorities and first responders can use, it also allows simpler chatbots to fit smaller organisations.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

5 THE AI CHATBOT CONCEPT DESIGN AND BLUEPRINT

5.1 EXISTING SOLUTIONS: STRENGTHS AND WEAKNESSES

As mentioned before, the review identified 45 AI-enabled chatbot solutions related to emergency and disaster management. Most of them are health-related (34), and specifically, many AI-enabled chatbots were developed in the past 18 months to provide information about COVID-19 (31). Five were related to natural disasters (water, weather, food and earthquakes), three for general disaster-related issues and two focused on women in particular. Although some are still active, several stopped working for many reasons (e.g., they were intended initially for short-term use only). The complete list of solutions appears in Appendix A.

Regarding the communication platform, 25 of the AI-enabled chatbots were accessible through the web only, despite the fact that it is able to support omni-channel communcation. The rest of the 20 AI-enabled chatbots were accessible through the mobile phone (e.g., SMS, messeging apps) or mobile apps, four of them, also by Facebook Messenger, which is accessible by both mobile or web. Neither of the AI-enabled chatbots were accessible through the web and mobile apps, although most web-based AI-enabled chatbots could be accessed through mobile browsers. Thus, only four AI-enabled chatbots were available on more than one platform. Besides Facebook, the AI-enabled chatbots were accessible also through other apps, such as WhatsApp, Telegram, Viber and in three cases (one each) on Line, Slack and an emergency self-developed app.

The review of the solutions highlighted that in most cases, the approach of the AI-enabled chatbots was conservative and cautious. Most AI-enabled chatbots adopted a narrow, usually closed-scenario-based, interacting approach. Twenty-six solutions only allowed the user to choose from a closed or almost closed list of questions and topics and provided predefined answers. Out of the rest, the chatbots allowed free writing, but in most cases either provided basic information (e.g., providing emergency phone numbers and evacuation centres' addresses) or limited the range of queries by asking specific questions that instructed regarding what topics it can handle. Only in five examples, an open approach of asking questions freely could be found: Desi – COVID-19 AI-enabled chatbot developed by Kinoa, GetJenny COVID-19's AI-enabled chatbot developed in cooperation with a Finish health institute, SPeCECA, a general-emergencies AI-enabled chatbot, Everbrige's emergency public alert system and IBM's Watson AI-enabled chatbot which was adapted in several countries for the COVID-19 pandemic. In the first two cases, the chatbots are not currently available. In the case of SPeCECA, it was a theoretical development that was not yet implemented in practice.

Only in the case of IBM's Watson AI-enabled chatbot, an entire complimentary conversation was still available. Another IBM's Watson AI-enabled chatbot (Smartex), focusing on food disasters, also allowed free conversation, but it is still under development.

The review of these solutions highlighted several advantages and disadvantages. One advantage relates to usage statistics. According to WHO's news release from April 2021, more than a year after the chatbot's launch, more than two billion users could receive information and answers to their questions around the world[15]. A similar number of users used the coronavirus self-checker by the CDC.

Another advantage is the simplicity of most chatbots. Unlike more complex chatbots that sometimes confuse the users, fail or partly fail to process their requests (e.g., "I did not understand your request. Please try again"), the closed-scenarios visual AI-enabled bots provide a clearer picture of options for the users. In addition, the closed list of options defines the borders of the chatbot clearly and does not allow the conversation to shift to places and dialogues that the system is not trained to follow.

[15] WHO Health Alert brings COVID-19 facts to billions via WhatsApp

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

The advantage is that simplicity also leads to another benefit of responsibility and caution. In free-style chatbots, the margin of error in the answers or information is more extensive than in closed-scenarios dialogues. In addition, if the chatbots are trained to identify and alert, they may miss real emergencies expressed by the users. Therefore, the more popular version of AI-enabled chatbots include a closed list of scenarios which prevents in advance such situations. In addition, they provide a fallback option – to receive an emergency phone number that they can call to immediately.

Other advantages of the chatbots under review were the diversity of communication platforms, availability in several languages, and unique services (e.g., the Japanese weather chatbot that addressed tourists and foreigners in particular).

In addition, while most AI-enabled chatbots' closed-form could be considered an advantage, several chatbots that used a more open approach (e.g., free-style writing) could also highlight an advantage. Allowing users to express themselves positively might inquire about information that the closed system will not access. Therefore, the openness versus closeness of the chatbots is, at the same time – an advantage and a disadvantage. From the technological point of view, and based on the innovative capabilities that AI and ML can offer authorities, first responders and building societal resilience, the closed approach is considered a disadvantage because of the limits it puts. From the responsibility point of view, some could argue that it is an advantage, but as reviewed in the above sections, this concern can be answered by giving a seamless trasnfer to a human operator and stating the emergency numbers and working hours in the beginning and other phases in the dialogue process. Another way to settle the advantages and disadvantages of closed and open approaches, is to build new AI-enabled chatbots in steps. In the first stage, to build a closed system, and as more data is collected, to develop more topics and to open the system.

For the same reason, the focus of most current chatbots on a closed list of issues also limits the users and might not lower the barriers that prevent them from using an AI-enabled chatbot instead of calling the busy call centre. A limited number of chatbots allowed a wide range of topics, but the list was very narrow in most cases.

5.2 Critical lessons from existing solutions and relevant ML approaches

Most of the analysed solutions were not open-sourced, with tiny details regarding the ML algorithms in use. However, from the interface of the chatbots, three dominant algorithms and approaches could be identified. The first is decision trees, allowing the AI-enabled chatbot to sort the data, follow paths, identify criteria, and decide which nodes to follow to provide the answers. For example, regarding the COVID-19 chatbots, following the system's questions helped the chatbot navigate through the different categories until it received all the information to extract the data relevant to the last step in the dialogue.

AI-enabled chatbots that allowed free-style writing or answering open questions in texts added another layer of Natual Language Processing (NLP), the AI technology concerned with giving computers the ability to understand the text and spoken words in much the same way human beings can. In several cases, they used the algorithm of Natural Language Understanding (NLU), a subfield of natural language processing (NLP), which involves transforming human language into a machine-readable format. It allowed the chatbot to identify the meaning of the text, the users' input, their intentions, what they are inquiring about, and what types of information they are hoping to get as an answer. On top of the decision trees and NLP (or NLU) algorithms, many chatbots used a Q&A (questions and answers) framework, generating a list of questions and the matched answers from already existing manuals, guidelines and other documents.

In addition, the dominant learning approach, as emerged by reviewing the chatbots, was supervised learning, with tagged and labelled datasets. The mostly closed, with almost no or little request for feedback from the users, emphasising that no additional learning was required from the AI-enabled

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

chatbot. Although it was not stated officially in the more open or fully open chatbots, it is possible that either reinforced or maybe even unsupervised learning was employed.

These findings from the analysis strengthen the conservative nature of AI-enabled chatbots concerning emergencies, disasters and societal resilience, as used by authorities and first responders. However, it also reflects a significant gap between what authorities and first responders are currently doing, from the ML algorithms point of view to the wide range of possibilities that could be adopted. This calls for many improvements, on the one hand, and to consider the current technological state of emergencies and disasters.

5.3 Relevant datasets

In general, the ability of the AI-enabled chatbot to create meaningful and helpful conversations with the users is based on two factors. The first is the technology, which was broadly discussed before and will be discussed further in the following sections. The second is the authorities and first responders' datasets, the AI-enabled chatbots' core information. The datasets include the information that the organisation stores during its activity. Datasets include guidelines, documentation of previous events, logged and transcribed conversations with callers, fact sheets, statistics and many other data types.

Data help the organisation to improve itself. The managers, workers and volunteers can learn from those datam wgucg can also facilitate the training process of the AI-enabled chatbot – which can be considered another vital worker in the organisation that needs to learn.

As part of the work of deliverable 2.4, regarding the communication strategies of authorities and first responders, and by reviewing the AI-enabled chatbots in this deliverable, many datasets were identified. The AI-enabled chatbots which were reviewed in this deliverable used several types of datasets. For example, they used health decision protocols (e.g., emergency triage, coronavirus symptoms) to guide the user step by step in deciding what to do. They also used Q&A generated documents, a list of phone numbers and other contact details of emergency organisations and other essential statistics, facts and recommendations, in the form of lists. In the closed, decision-trees-based AI-enabled chatbots, these were the most popular datasets.

Other AI-enabled chatbots, especially the more open-approach ones, also used alternative datasets. For example, they used information guidelines, databases of previous disasters, their results and recommendations, dictionaries of synonyms and antonyms to facilitate better dialogues, preprogrammed answers from experts, and live data updated a few times a day more. However, despite the detailed list of examples, very few AI-enabled chatbots used them.

One of the main findings in deliverable 2.4 was that authorities and first responders manage many datasets. However, some interviewees even mentioned that, except for managing those datasets, they have not as yet found the proper use for them, in order to improve. Instead, they mentioned that one of the pieces of advice they need is how to use those datasets more appropriately and what can be done with them.

Of course, not every type of dataset could be relevant for training the AI-enabled chatbot. For example, transcribed conversation in the call centres and guidelines regarding handling an emergency call are more relevant for training the chatbot to conduct meaningful dialogues than information sheets, statistical data and lessons learnt from previous disasters. However, the latter is more relevant for the chatbot's ability to extract information for the users' inquiries.

This is a significant part of the communication goal of ENGAGE and the objective of helping authorities and first responders to make more information accessible for the public in order to contribute to building societal resilience. Therefore, the current map of datasets used by authorities and first responders' AI-enabled chatbots, compared to the types of datasets they hold, calls for an

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

approach that will extend the use of such datasets to facilitate better and broader AI-enabled chatbots.

5.4 What is working? What is not working? What can be improved?

The review of advantages and disadvantages highlighted several issues regarding what is working in the use of authorities and first responders of AI-enabled chatbots, what is not working, or less working, and what could be improved.

One issue is the conservative use of AI-enabled chatbots by authorities and first responders. The field of AI-enabled chatbots offers innovative technologies for improving. Is is very common, with tens of thousands examples. However, the field of emeregncies, disasters and societal resilience is still far behind, with many barriers prevent this change. For example, barriers include fears of widening the use of AI-enabled chatbots and extending the areas of topics they deal with; or disapprovals of specific situations that authorities and first responders do not want to be responsible for (e.g., triage, identification of emergencies).

Another issue was the closeness versus the openness of the AI-enabled chatbots, with advantages and disadvantages for each approach. On the one hand, even closed-scenario chatbots (e.g., CDC's coronavirus symptoms checker) attracted billion of users and provided important information for the public. On the other hand, this approach also prevents many users from receiving other types of information they need. In contrast to the closed approach, the open approach allows more information, but at the same time, may lead to more mistakes.

The third reccurring issue relates to the minimal use of specific algorithms and datasets as the "core" of the AI-enabled chatbots. Authorities and first responders based their chatbots on a very selective number of sources for providing answers based on them, using, in most cases, basic ML algorithms and ML approaches. While it worked for the leading cause, to provide focused information, as mentioned before, the limitations are also the challenges.

Therefore, in order to improve the current situation, specific considerations should be taken into account. As will be elaborated in the next section, various AI technologies can facilitate innovative AI-enabled chatbots. However, the component that allows for slow but continuous growth, development and constant human inspection is crucial. Therefore, future AI-enabled chatbots should take into consideration the adoption level of authorities, and first responders, and how comfortable they feel with using such technologies. However, they should also be well familiar with the technological options and how they can facilitate the work of such chatbots and answer some of the stakeholders' fears and organisational barriers.

In addition, future developments should keep both options regarding the closeness and openness of AI-enabled chatbots. On the one hand, the closed nature of the chatbot provides more order and, at least in the beginning, a clear definition of what could and could not be achieved by using the chatbot. On the other hand, in the context of emergencies and disasters, the organic nature of societal resilience is more complex than the problem-solving instrumental approach of AI. Therefore, it is also essential to combine free questions and queries that allow the AI-enabled chatbot to provide more complex situations.

The same implies the combination of various ML algorithms, learning approaches and datasets. On the one hand, more algorithms, more complex learning approaches, and various data sets provide the AI-enabled chatbot with more data and learning capabilities. As a result, it can provide better training and abilities, and as a result, more important information for the users. On the other hand, as shown in the literature review, more is not necessarily better and can bias the chatbot, resulting in the users' information. Therefore, the improvement in combining more algorithms, learning approaches and datasets should be done in steps, with manual human inspection and monitoring for each stage, allowing another expansion only when reaching a satisfactory level of information provided by the chatbot.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

In the next section, we review the components of current blueprints of innovative AI-enabled chatbots produced by leading technological companies, discussing their relevancy to emergencies, disasters and societal resilience and the possibilities to use them in AI-enabled emergency chatbots, taking into consideration the suggestions of this section.

5.5 Existing blueprints and their components

Apart from mapping, documenting and reviewing the current AI-enabled chatbot solutions that are already in use concerning emergencies and disasters, we collected and analysed existing blueprints/frameworks of innovative leading AI-enabled chatbots that are available in the market. The results are presented in Appendix C. As this field involves global commercial companies, some of the data is not publicly available; however, extensive data could be found in several cases.

This analysis aims to understand the technological capabilities available to authorities and first responders, which may be incorporated in AI-enabled chatbots that deal with emergencies and disasters and contribute to societal resilience. Therefore, the character of this section will focus on the technological-instrumental nature of AI-enabled chatbots through the existing blueprints and their components, while the following sections will connect the framework to emergencies, disasters and societal resilience.

The table in Appendix C describes the different components of Microsoft's AI-enabled chatbot services, Google, IBM, Amazon Web Services (AWS) and Facebook. Components that constitute different services that are based on the same technology are clustered together in the same line. Each component is described, and whenever relevant, the possible contribution to emergency and disaster AI-enabled chatbots is suggested.

5.5.1 Connecting the chatbot to the communication channels

One of the components of AI-enabled chatbots is the services that allow the interaction between the chatbot itself and the user, through the chosen communication channels. Several chatbots were developed under the same system. However, most chatbots interact with external communication platforms (e.g., WhatsApp, Facebook Messenger, Telegram). Since such platforms are not part of the development of the chatbot, there is a need to facilitate a connection and transference of data between the two systems. This, of course, is not limited to one platform, but one AI-enabled chatbot can facilitate and support conversations in all platforms.

To achieve this, several commercial services facilitate the communication between the user and the chatbot through relevant communication channels. One example is Microsoft's Bot Framework Services (BFS), which can connect the chatbot to popular communication services that users use worldwide, regardless of the system it is based on and how it was developed.

5.5.2 The chatbot's logic

Another critical component in the existing blueprints of AI-enabled chatbots is the chatbot's logic. It can be defined in simpler terms as the "glue" of the system. These services create the most critical "hub" in the framework, through which all the data flows. It connects all the components, applications, services and technological capabilities. The pre-analysed data are sent there before the analysis and are received via this hub after completing the data analysis. An example of the chatbot's logic components is Cloud Pub/Sub by Google.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

5.5.3 ML ALGORITHMS: LANGUAGE, IMAGE AND VOICE PROCESSING

The communication process between the users and the AI-enabled chatbot requires the substitution of a communication process between two human beings. Therefore, the framework requires the capacity to process the user's message, understand it, and respond to it, but also reciprocally – to understand the analysed data, respond to it or make it accessible to the user (e.g., by answering).

Many services use ML algorithms to provide these capabilities. From Natural Language Processing, or partial Natural Language Understanding, through image analysis and voice processing. All allow AI-enabled computerised systems to understand and analyse different types of data.

IBM's NLP capability, for example, is called DeepQa. It allows the IBM Watson chatbot to conduct conversations and provide answers to questions. Google's Vision API adds the layer of image analysis, identifying objects and text included in it, and other services to provide similar solutions, even among other types of data (e.g., video, voice). This allows the development of more complex chatbots that can interact with users not just in a close-ended way (choosing from limited options) but in a more accessible way that can include more than just text. This expanded option may be necessary in emergencies and disasters when users sometimes need to send a live picture and not just describe it.

5.5.4 ACCESS TOOLS: INDEXING, SEARCHING AND GENERATING EXTERNAL LAYERS FOR THE DATA

AI-enabled chatbots are based on data and datasets. When a user queries the chatbot, the system scans for the relevant data to answer even by just a greeting. In informational chatbots that aim to make information accessible to the user, several services and technological capabilities exist to create indexes, search documents, and additional external layers for the data. Those external layers are additional features on the raw data that make it more organised and accessible.

For example, there are services that index documents and make them more accessible or user-searchable, thus facilitating the extraction of relevant information from a document. Microsoft's Azure Search, Google's Dialogflow, and Amazon's Kendra are good examples of a technological capability. Other services, such as - Microsoft's QnA Maker and Google's actions, provide additional external layers, such as generating lists of questions and the matched relevant answers, based on semi-structured content of FAQs loaded into the system.

Such services can provide authorities and first responders with an opportunity to present important information that they hold accessible. For example, guidelines, transcripts of emergency calls, and other documented data can be loaded into the chatbot system, while it is currently documented electronically and sometimes not in a public database accessible to the public.

5.5.5 DEPLOYMENT AND INTEGRATION

Another component in the blueprints is related to deployment and integration. Since AI-enabled chatbots are a technological feature, they need to be accessible to the users. While one way to achieve this is to use a service that facilitates the interaction between the chatbot and the various communication channels, another option is to integrate the chatbot, for example, into a web platform. This way, users can access the chatbot through a browser after entering the organisation's website, and thus they are not dependent on social media accounts, like in the case of Facebook, or need to expose their phone number to the chatbot in the case of WhatsApp. Services such as Microsoft's Web App or Google's Web Integrations are examples of deployment and integration services.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

5.5.6 Data management, processing, quality assurance (QA) and storing

The activity of the AI-enabled chatbot relies heavily on data and information. Therefore, prior to explaining the role of analysis, which will be done in the next section, one mission that the system has, that should be included in chatbots' blueprints, is data management. Data management includes extracting data from various types of datasets, whether manually or automated on defined time slots, and managing the processing of the data in the system. For example, defining which type of data needs to be transferred, to which components in the system, for what purposes, at what time, when not to do so and more.

Quality assurance (QA) is another part of the data management and processing components, which sometimes appear as an independent component in the blueprint. A technological system, like any other system, is not error, mistakes, and wrong-decisions free. In the QA process, the results of actions and decisions of the AI-enabled chatbot are analysed, examined, and, if needed, are forwarded to other parts of the system for improvement. QA can also strengthen the good parts of the chatbot and not just make decisions regarding bad results. Examples for data management, processing and QA services can be Apache UIMA, a Linux open-source system used in the IBM reference diagram.

The last part of this component, again, sometimes independent in several blueprints, is storing. As part of the data management of AI-enabled chatbots, the data is stored to make it easier to extract and analyse. In most cases, the current servers of authorities and first responders might not be enough for this purpose. Therefore, external services, such as Amazon's S3 or Microsoft's Azure blob storage, are optimised to store big datasets in various formats and types, analysis-ready. The storing process also includes logging the data the chatbots create (e.g., conversation logs, extracted information and other data that can be used for QA, statistics or other purposes).

Current use of AI will almost always require access to some cloud service. However, the data provided by the public during an emergency often includes personal information (PII) as well as even more sensitive information such as personal health information. The use of such cloud services to handle sensitive data may be in direct conflict with security and privacy policies, and therefore, demends meetings requirements of security and authentication, as will be elaborated below.

5.5.7 Analysis, monitoring and insights

Apart from data management capabilities and based on ML algorithms, other crucial components of the blueprints include analyses, monitoring and providing insights. For example, in the AI-enabled chatbot's framework, the analyses services process the data and provide relevant information as outputs. For example, analysing medical data from big datasets, using NLP, tree sorting, and other ML algorithms to answer users' queries. Another example could be in analysing data from the datasets to train the AI-enabled chatbot. The difference between the component of ML algorithms and the act of analysis is that the first refers to the technological capabilities (e.g., the algorithms), and the second adds the layer of processing data based on those capabilities. Examples for analyses services could be Google's BigQuery, combining data from different types.

In addition to analysing data that serves to train the chatbot or answer users' queries, the analysis capabilities also contribute to monitoring. While the primary goal of the chatbot is to provide information and conduct meaningful conversations with the users, monitoring is also essential. Monitoring includes going over the logged data, identifying patterns, discovering anomalies, pinpointing messages that might demand special attention or human observance (e.g., emergencies, situations that raise danger to the user or cases in which the chatbot cannot provide the necessary assistance and needs to transfer the user to a human-assisted conversation).

Based on monitoring and logged data, another aspect of these components is to provide insights. Those insights could be in the form of warnings, as abovementioned, or in a more standard form of

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

statistics, dashboards or any other type of meaningful insights. Several examples for monitoring and insights services could be Microsoft's Application Insights, monitoring the logged chatbot's metrics, for diagnostic and analytical purposes, or Power BI, for creating automated dashboards and other business intelligence (BI) data, or Google's Chatbase, tracking predefined KPIs, insights based on the content of users' messages and insights regarding user journies.

5.5.8 Security and authentication

Apart from possible analysis and meaningful insights, another vital component of AI-enabled chatbots is security and authentication. Since AI-enabled chatbots are technological systems that heavily rely on big datasets, securing the data and authenticating the legitimacy of users is crucial. This is a significant consideration because of protection and preventing information from leaking and privacy regulations, such as GDPR. Authorities and first responders cannot use such technological systems if they do not meet those regulations.

The security and authentication components help design AI-enabled chatbots that can handle different types of data, personal data, and other sensitive data, such as restricting access to the system or preventing unauthorised access to data that regular users should not access. In order to do so, the system of the chatbot should be able to authenticate the user and the access platform and provide a relevant key, or token, that opens the desired access to the system. Examples to security and authentication systems, from AI-enabled chatbots' blueprints, are Microsoft's Azure key vault and Azure active directory, Amazon's Identify and Access Management (IAM) that store credentials and authenticate the users, and from a similar perspective, Google's Cloud Data Loss Prevention (CDLP), that redacts sensitive information (e.g., addresses and phones numbers) remaining in transcripts before storage.

5.5.9 Datasets types

The last repeating component in the AI-enabled chatbots' available blueprints is the datasets types. A dataset is a collection of data. That could be a list of customers, a set of guidelines, FAQs or different types of files, from the text (e.g., word, pdf), photos, videos or any other type of file. The most common division of datasets is between structured and unstructured data. Structured data is less sophisticated to analyse. It is tabular data, divided into rows, columns and cells, that are defined. Analysing structured data demands defining the rows, columns and cells, allowing the AI-enabled chatbot to understand what type of information to expect.

On the other hand, most of the information in current datasets is unstructured. It follows no structured order. There are no two single datasets, or pieces of data, that look the same and follow the same logic, and therefore, it is more complex to analyse. In the significant data era, the exponential growth of data also resulted in the ongoing increase in different data types.

Therefore, the choice of datasets that the AI-enabled chatbot will rely on affect the system's necessary components and technological capabilities. As a result, it can change the authorities and first responders' decisions regarding what services to use in any of the components. However, as illustrated above, the available services nowadays offer the relevant technological capabilities for all types of datasets in various proficiency levels.

5.6 The blueprint

This section will present the blueprint addressing the design and implementation of the AI chatbot blueprint. It will review the following parts of the blueprint: design concept, case studies, adaptation options, roadmap and implementation stages.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

5.6.1 INITIAL WORK ASSUMPTIONS

The blueprint is based on two critical working assumptions. First, during the professional meetings conducted while working on this deliverable, many of the Ki-CoP members and ENGAGE's partners expressed concerns and barriers of using AI-enabled chatbots, which may be relevant to many other authorities and first responders who were not represented in those meetings. The barriers and concerns that were expressed ranged from not trusting the technological capabilities of the AI-enabled chatbot, through possible biases and mistakes, to being afraid that the public will not adopt this solution and will still look for human assistance. Yet, one of the work assumptions is still that AI-enabled chatbots use a popular technology in the private sector, which answers some of the perceived barriers of the public.

We address these barriers in the blueprint in several ways. We elaborate on how the technological processes operate, explaining how they can address those concerns, suggest a functioning AI-enabled chatbot, and under which conditions. As mentioned before, we also suggest a roadmap and a step-by-step adoption model that allow authorities and first responders to adopt and adapt to "thinner" solutions. This allows facing one barrier at a time, moving to the second step only after overcoming it. We also refer to the organisations that provide these capabilities that can also offer, beyond the capabilities themselves, support and elaboration to the adoption process.

Another way we address these barriers relates to the second working assumption – the need to validate that the chatbot is doing what it should do at a satisfactory level. Therefore, the primary initial recommendation is that the chatbot collaborates with a human call centre that will inspect that it is functioning correctly, providing accurate answers, maintaining the treatment of false information and more. The human professional accompaniment will vary according to the different stages. In the initial stage, the monitoring will be of most conversations, ensuring that there is good accuracy in the work of the chatbot, according to predefined accuracy rates. After it reaches a satisfactory level of accuracy, the monitoring will continue, but in the crosspoints elaborated in the blueprint (e.g., false information validation, checking alerts, quality assurance, regular maintenance, statistical analysis, creating dashboards and more).

5.6.2 DESIGN CONCEPT: COMPONENTS, TECHNOLOGIES AND ALGORITHMS

Based on models of societal resilience, the existing technological AI capabilities, published blueprints and frameworks of AI-enabled chatbots and previous solutions, we suggest the following blueprint, as appears below. The blueprint is a general design concept with suggestions for the different components of an AI-enabled chatbot for emergencies. In the following sections, we introduce the different components, but without recommendations for specific products of specific commercial companies. Instead, the description includes the services that could be integrated into the chatbot, their advantages and suggested use, and above all, the contribution to societal resilience. At first, we describe the most comprehensive and innovative form of a chatbot, while later, we suggest "thinner" versions that can be adapted.

There are two "entry" or "base" points to the framework – the users and the organisations. The users interacting with the chatbot can submit a request and receive a response to their query. In such a situation, starting a conversation with the chatbot, even only by accessing the page or the app without typing anything, is considered a request that will be responded to. Regarding the users, several input types can be received, and AI-enabled chatbots can facilitate the response. For example, users can type text, record a message, send any other audio file, or upload photos or videos. However, as the audio analysis may be less accurate concerning a voice of a person the system did not train on, audio inputs should be limited to a predefined set of requests (e.g., recording the sound of water flow after being asked to). The same also applies to video files.

Affected citizens receiving a public warning message from an emergency center (e.g., "reverse 112") will often be able to not only share details by voice (phone) or text (i.e., SMS), but also upload

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

images and video directly to the emergency operators. Coupled with the senders estimated location, as provided by the mobile network, this information might be extremely useful for emergency operators and first responders on the ground, but making sense of large amounts of incoming data may be a very tedious process for the human emergency operator. A potential use case for AI would be to "parse", sort, organise and categorise incoming data from the public in an early stage, making it easier for emergency operators to see patterns and piece things together. For instance, images taken by affected citizens during an emergency can be geo-coded on a map and organised by location and timestamp, thus providing operators with a visual view of the situation on the ground.

The connection between the input type and societal resilience lies in discussing the public's communication needs. One of the key findings in deliverable 1.3, which examined the public's communication needs regarding societal resilience, elaborated on the importance of allowing multiple public methods to report or query about emergencies and disasters. In their interaction with authorities and first responders, users can find themselves with different needs, and some can be communicated through text, voice, or the need to send an example of something they find difficult to explain. For example, when a user cannot explain where he or she is located, the ability to send a location, using the phone's GPS, and receive a more accurate answer – is essential.

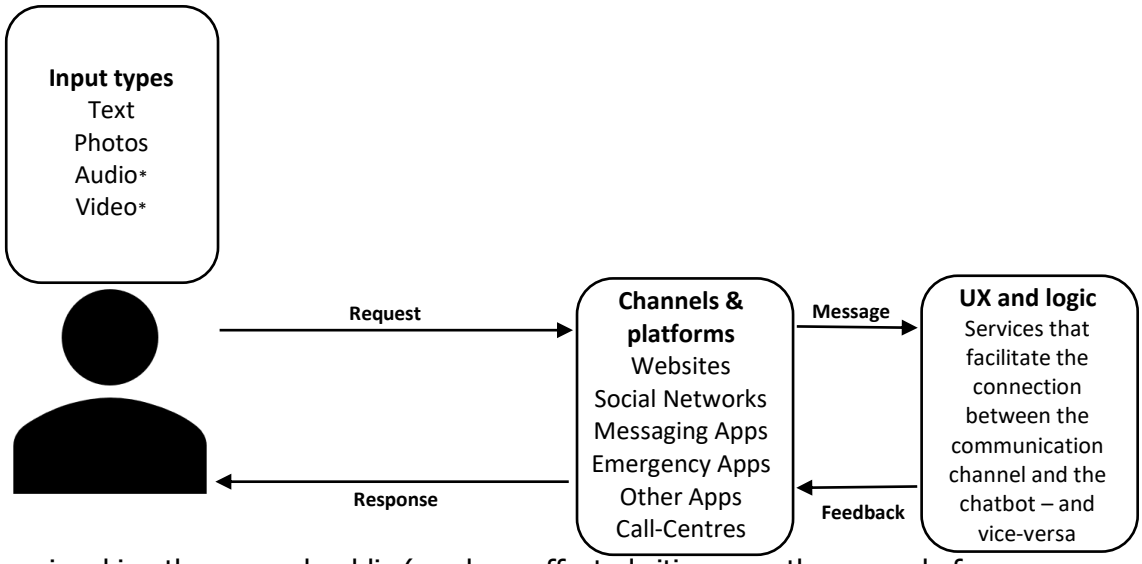
It must also be considered that in many situation the communication or engagement with the public might be triggered by the authorities or first responders. So it cannot always be assumed that individuals will initiate contact. In the case communication is initiated by authorities and first responders, the AI must be able to assist in handling of potentially large volumes of responses, and guide users of the AI to various outcomes including call for action.

Another aspect to consider is the User Experience (UX) and the platform of the chatbot. As previous sections show, AI-enabled chatbots can be implemented in various platforms, from websites to popular messaging apps (e.g., Facebook Messenger, WhatsApp, Viber, Telegram, SMS) or other apps. They can also be integrated as part of call centres (voice). Organisations can implement the chatbot within one channel, or preferably, implement it in all available platforms that they use (omni-channel). Along with the choice of digital platforms, the user should also have the possibility to interact with a human, and not solely to the chatbot, whenever it is possible. The interaction between the channels and platforms to the bot is conducted through a service that facilitates the connection. This is the "body" of the AI-enabled chatbot that connects all the internal and external components of the framework.

The connection between the UX and platform and societal resilience is similar to the previous aspect, which discussed the input types. The idea of using multiple platforms to communicate with authorities and first responders was another important finding of deliverable 1.3. The ability to address authorities and first responders through multiple channels, and not just through one of the channels, was perceived to improve the communication process between the public and authorities and first responders – and vice-versa. As shown in the survey reported in deliverable 1.2, communication was one of the contributing variables to building societal resilience.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Figure 3. The user's entry point.



When invoking the general public (random, affected citizens on the ground of an emergency), users will prefer communicating though a "known" channel which they already trust and are comfortable using. Some may prefer to use plain SMS, some Facebook or Messenger, while other users will only trust a web page where the domain of the URL is familiar and trusted. This has resulted in always to assume the multi channel approach - not just when broadcasting alerts the public, but when receiving feedback engaging in two-way communication with citizens. Another important note here is that there are various scenarios (e.g., an active shooter) where users on the ground cannot raise their voice and so are forced to use a text based channel.

Three other components of the AI-enabled chatbot's framework are connected to the component of UX and logic. The first is the component of security and privacy. This refers to most technical components securing the information stored in the chatbot system, including users' details and other data that require specific security and privacy measures. However, despite its technical orientation, it also has a contribution to the idea of societal resilience. The possibility to interact with a security system can encourage users to use it. In addition, data leaks will reduce the trust levels in the AI-enabled chatbot, decreasing the sense of societal resilience.

The second component, and one of the most central in the framework, is related to cognition and intelligence. This can be considered as the "brain" of the chatbot. As discussed before, and in other deliverables, one of the crucial challenges in communicating emergencies and disasters concerning societal resilience is identifying the needs of the public. Another significant challenge that completes the first is the ability to tailor the information to individuals and groups. For example, when a citizen calls an emergency service or even inquires for information in a call centre of authority, one of the first tasks of the human responder is to understand what the citizen needs. The responder then needs to provide the information and examine whether it answered the citizen's needs, and in some instances, whether it was understood or there is a need for more clarifying information. AI-enabled chatbots need to augment this process, which is at the core of the cognition and intelligence component.

The component of cognition and intelligence involves the question of which ML algorithms to use. As shown in figure 4, several available algorithms, from NLP to speech recognition and image analysis, could be integrated into the chatbot. Those algorithms are part of services provided by several organisations. Such services and algorithms can give the chatbot the ability to understand the users' requests, and as will be explained later – also to use its database to respond.

Another aspect of the cognition and intelligence component, as shown in figure 4, is the ability of the AI-enabled chatbot to identify intents, entities and create dialogues. Intents refer to what the user wants or is looking for. For example, does he or she want to ask a question? To report on an

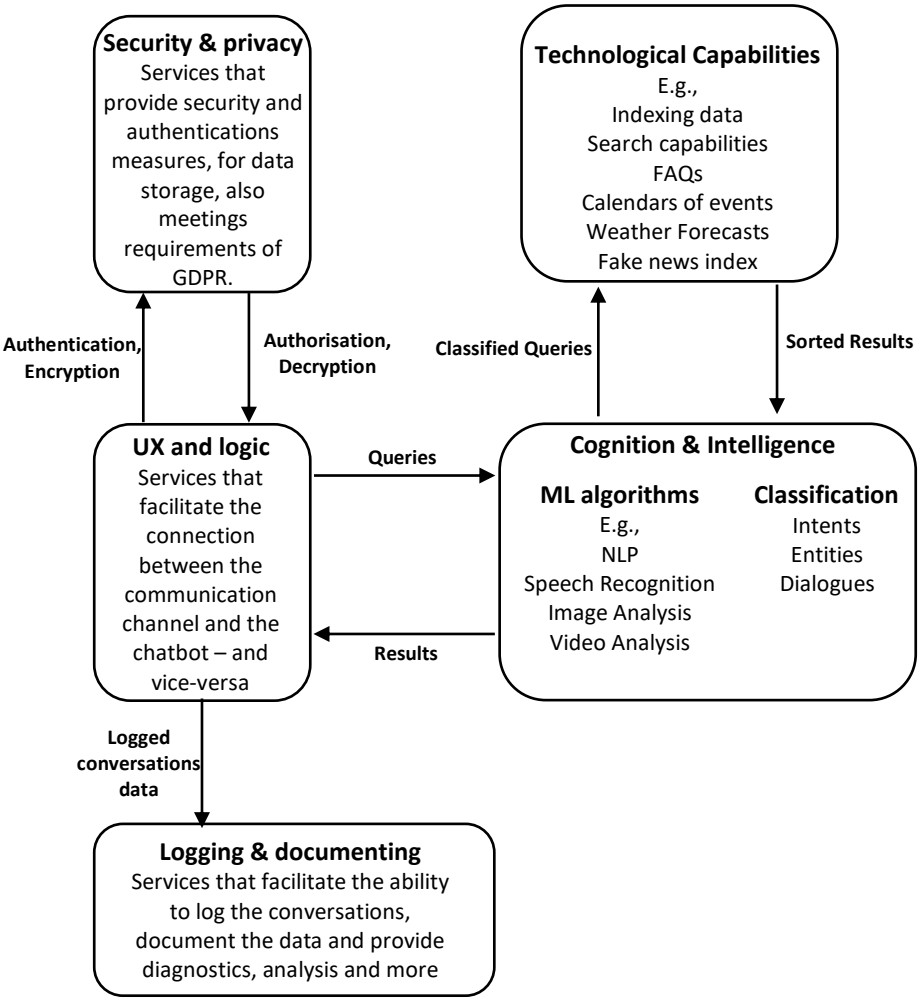
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

event? Or maybe just to politely greet the chatbot? Entities allow capturing specific values within the user utterance. It unfolds the ability of the chatbot to understand what the subject of the conversation is or what the user wants to do or understand. For example, locations, types of events and organisations are examples of entities. News updates, regulations, opening hours and forecasted weather conditions are additional examples of entities. Lastly, dialogues are how the AI-enabled chatbot conducts the conversation with the user. There are rules, for example, how to progress in the conversation when an intent is identified, when to go back and look again for intents, how to respond to specific entities, and more.

As mentioned before, intents, entities, and dialogues aim to partly augment the work of call centres and other human multidirectional communication processes of authorities and first responders concerning emergencies and disasters. Both AI-enabled chatbots and human operators' mission is to identify what callers and users want and expect. Both aim to understand the subject and follow call scripts to manage the communication process with the caller/user.

The component of cognition and intelligence can connect with various technological capabilities and services that use the ML algorithms to provide necessary actions to an AI-enabled chatbot of emergencies and disasters: the ability to search within documents or datasets (in order to match the need of the user with an adequate response), generating and indexing frequent questions and answers that can be a reasonable basis for designing and improving the dialogue process, create a calendar of events, a list of recommendations (e.g., during COVID-19, several chatbots generated a list of recommendations and matched relevant recommendations with requests of users), provide a list of external sources for the user, match information with a database of known fake news stories and more.

Figure 4. Main components connected directly to the UX and logic.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

The third component in figure 4 relates to logging and documenting the information. When operated in an advanced capacity, the AI-enabled chatbot can have millions, and even many more, conversations with users. Several processes are needed to facilitate these conversations. The first is data that is received and serve as the base for conducting these conversations. Second, in addition to using the data, the system needs to store and process the information. For these purposes, the AI-enabled chatbot requires logging and documenting capabilities. It needs a place to store all the information, the conversations, the details, and the data. However, storing the information is just one requirement. Such information can also be analysed, for statistical analysis, improving the chatbot's activity, creating dashboards, either for the activation of the chatbot or for other purposes and more.

These actions replicate the processes conducted by emergency authorities and first responders and, in some cases, improve them. For example, authorities and first responders usually keep documentation of calls and inquiries for future use. Such processes contribute to improving their communication with the public and the services they provide. As a result, improving authorities and first responders' work also contributes to building societal resilience. Automated logging and documenting capabilities streamline the process and improve the feedback authorities and first responders receive about the public's societal resilience and how to enhance it.

The component of logging and monitoring is connected in addition to other components. One of them is monitoring and reporting. This relates to the process of analysing the data and its diagnosis, as mentioned before, by authorities and first responders. However, it also relates to another crucial aspect of the work of authorities and first responders in building societal resilience – prompt monitoring and identifying live occurrences.

When there is a surge of calls to the call centre, or any other communication channel, from a specific geographic location, with similar reports – it immediately raises an alert that an ongoing emergency might be occurring. Fast identification of occurrences can lead to a quick response that maintains societal resilience. Conversely, failing to identify the emergency may lead to a late response that harms the societal resilience. In this case, while human forces are able in many cases to identify an occurrence, when there is a surge of calls, in many other cases, it is unable to do it on time. For example, in cases of dissemination of fake news, when people keep calling and asking questions based on a fake source, the human operators are unable to track it on time and alert on its existence.

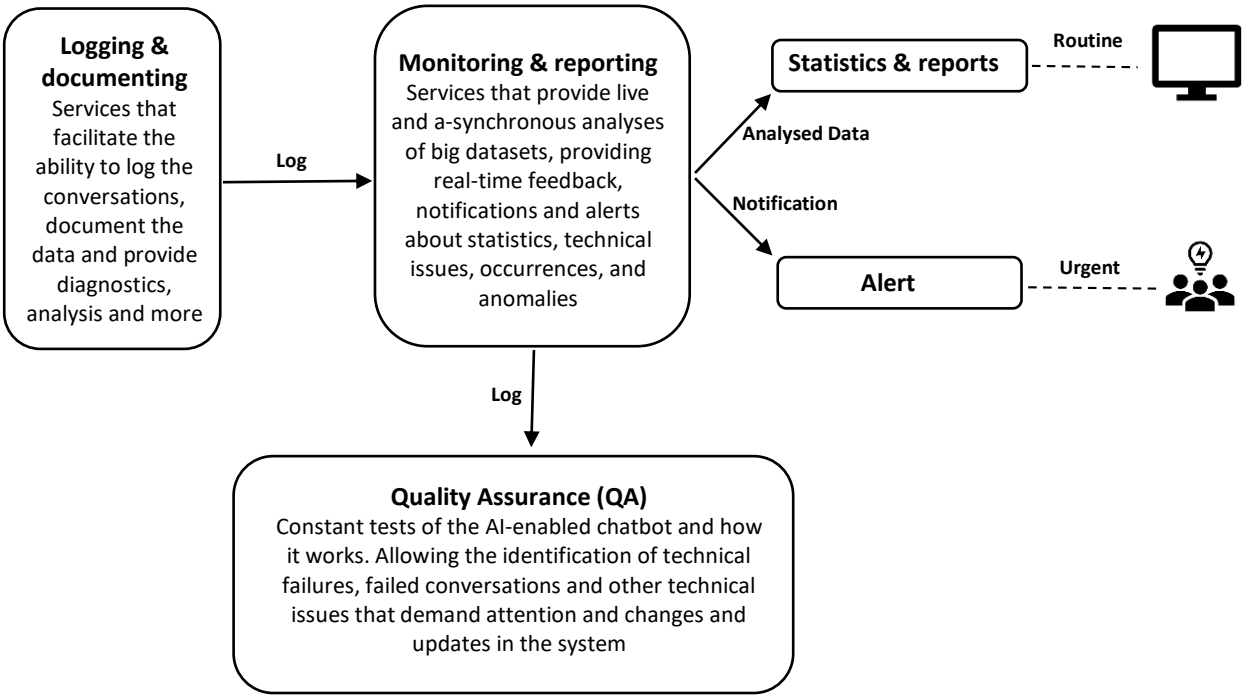
The services and technological capabilities that can analyse and diagnose the data can monitor and alert several occurrences. They can alert about technical failures or on a concentration of cases where the chatbot could not conduct the dialogue properly. They can identify and immediately highlight anomalies or repeated patterns in the conversations that might demand special attention or even about individual conversations that demand the intervention of a human operator. The alerts are not unidirectional, from the public to authorities and first responders, but can be processed by the system and let authorities and first responders alert the publc, in a circular process. The sender of an urgent alert to the public will often prepare for receiving incoming responses from citizens on all applicable channels. Processing such incoming replies (text, images, video) and handling two-way communication directly with affected citizens on the ground is a strong potential use case for AI in scope of public warning.

The monitoring component of the AI-enabled chatbot can also connect to specific sources, such as social networks, discussion boards and news platforms, and provide similar capabilities. For example, identifying repeated patterns of information that do not match the official information that the authorities and first responders disseminate (i.e., suspicious mis/this information or fake news).

Therefore, the output of the monitoring and reporting component can go in three directions. One is for statistics, analysis, and information reports. The second is for informing and alerting about occurrences. The third is for technical issues. This relates to the quality assurance (QA) of the system and its constant updates and improvement. The QA component is connected to many other components in the framework since the QA results can change and affect those.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Figure 5. From logging & documenting to statistics, reports, and alerts.



The last component of the framework/blueprint refers to the data and information that authorities and first responders hold. Deliverable 2.4, focusing on the communication strategies of authorities and first responders and the collection of solutions described the different types of datasets that authorities and first responders hold. Datasets that include guidelines, conversation logs, treatment protocols, history of events and more. In addition, as mentioned above, some of the information that the AI-enabled chatbot deals with may be found on social media and other external websites that are monitored. Some of the data is structured, fixed in rows, columns and other types of fixed fields and records (e.g., CRM, SQL, conversation logs). However, most data is unstructured and needs to be organised or structured (e.g., FAQs, guidelines, information sheets). This data serves both for the training of the chatbot and the retrieval of information for conducting the dialogues with the users and providing answers.

Concerning emergencies and disasters, these datasets are one of the most vital contributors to building societal resilience. Authorities and first responders created those datasets to facilitate better instructions and knowledge that the public needs to be better prepared and cope effectively with emergencies and disasters. However, as noted in the interviews of deliverable 2.4, one of the challenges that authorities and first responders sometimes face is in communicating and making this data accessible to the public.

In this case, the role of the AI-enabled chatbot is essential. Using technological capabilities and external services, such datasets are being prepared for training, analysis and communication, integrated into the knowledge system of the chatbot and used in the communication process with the users. This is the component of preparing, processing, and training. For example, when a user asks about recommendations for flood preparedness, the relevant instructions from guidelines and information sheets can be extracted, matched with actual data about current risks (e.g., recent weather forecast and how to adapt the instructions to the current predicted situation). The component of preparing, processing and training is related not just to the datasets themselves.It is affected by the QA processing (e.g., when interactions with the chatbot highlight the need to update the information) and is co-effecting the cognition and intelligence component (e.g., on the one hand, it uses the algorithms to analyse the data for intents, entities and dialogues, and on the other hand, it gets the input from the users to extract information from datasets).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Figure 6. Processing the datasets and training the chatbot.

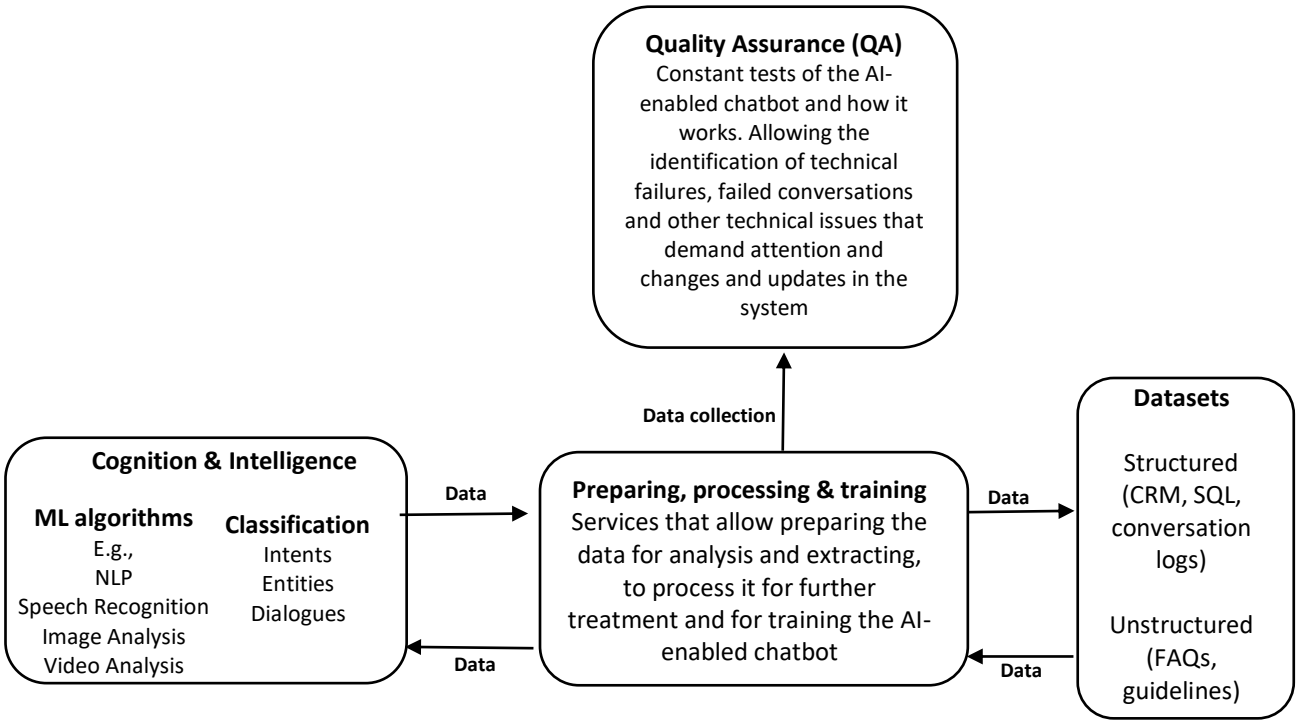
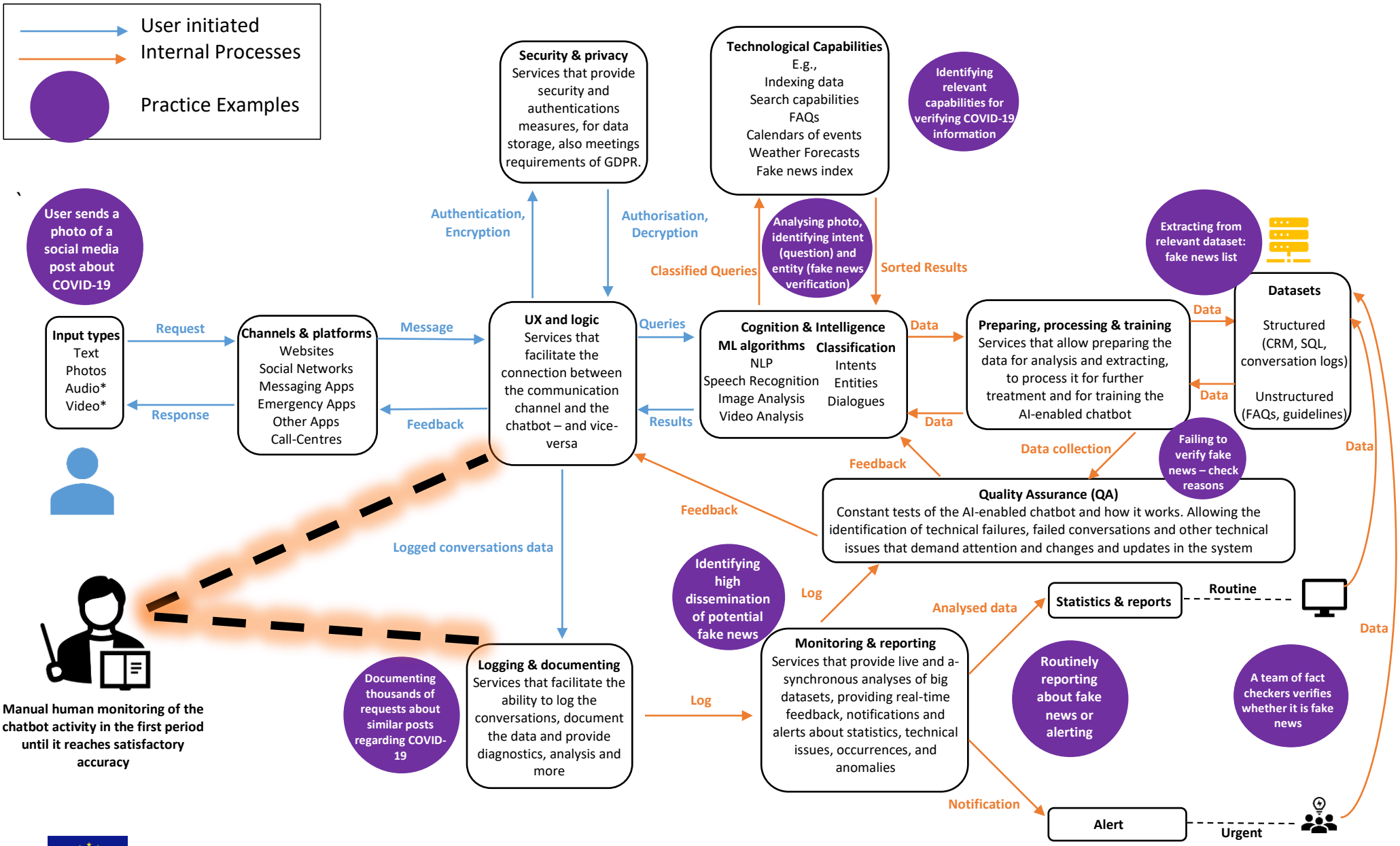


Figure 7 below presents the complete framework/blueprint. It illustrates all the components of the blueprints and how they connect. In addition, the figure highlights the actions performed by the user (e.g., asking a question) and other actions which include internal processing of the system (e.g., analysing data from social media, regardless of user inquiries). It also provides several examples for uses and actions in the practice of the system. In addition, the framework suggests human monitoring of two core processes until the chatbot achieves a satisfactory level of accuracy.

In the following sections, we will present several examples/case studies that will elaborate on the different functionalities of such a system and will clarify their contribution to societal resilience. In addition, we will draw the different options for step-by-step, partial, and different methods to adopt this framework among emergency authorities and first responders.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Figure 7. The complete blueprint/framework.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

5.6.3 Case studies/Examples of the blueprint

Example #1: User inquires about covid-19 fake news

The first example illustrates a situation in which the user initiates the process. In this example, a user encountered a post about COVID-19 on social media and wants to verify its authenticity. The user either "copy and paste" the post into the conversation with the AI-enabled chatbot or takes a screenshot (input types). The user then opens the Facebook messenger and starts a conversation with the chatbot (channels & platforms and UX and logic). After clarifying the request, the user submits the text or photo. Next, the query is analysed (cognition & intelligence) using various algorithms (e.g., image analysis to extract the text in the photo, NLP to understand the user and the post), identifying its intent (asking a question), entity (fake news verification) and conducting the dialogue step by step. At the same time, the user's query is logged into the database of the organisation (logging & documenting). Then, the query of the user is processed (preparing, processing & training). By using the relevant technological capability, for example, fake news list (technological capabilities), the relevant dataset is extracted, the information in the post is found in the dataset (preparing, processing & training) with 89% similarity of the text and information, employing the relevant ML algorithms (cognitive and intelligence) and then the answer is returned to the user through the chatbot's interface (UX and logic) in the platform of the conversation (channels & platforms).

Example #2: Fake news alert

In the first example, a simple process was described – inquiring about fake news and finding it in the dataset, confirming it as fake news. In this second example, it is assumed that this data is missing and could not be verified. In this case, after processing (preparing, processing & training), the result, and failing to verify it, it is sent as feedback to the chatbot interface (UX & logic) and logged (logging & documenting). It is then analysed and checked with external sources (monitoring & reporting), and after no significant match or anomaly is found, it is documented as part of the usage statistics (statistics & reports – routine).

However, later on, thousand of inquiries regarding similar posts and pieces of information are logged in the chatbot's database (logging & documenting) and identified in the monitoring process (monitoring & reporting). Since it is a significant exception, the system alerts the fact-checking staff (alert) who manually examine the post and define it as fake news. In response, they manually update the system's data (datasets) and tag the information as fake news for future interactions.

Example #3: Training the chatbot to provide emergency information on natural disasters

As mentioned above, there are two entry points to the framework. While the first two examples focus on the user starting point, the third example starts from the authorities and first responders. An organisation has many datasets it has stored through its years of operation, from guidelines on conducting a rescue mission during floods through safety recommendations for the public to hundreds of thousands of transcribed calls to its emergency call centre. These documents (datasets) are the basis for training the AI-enabled chatbot (preparing, processing & training), either by letting it analyse them (unsupervised learning) or partly tag them (e.g., marking Q&As, marking examples for a list of recommendations) and then letting the chatbot learn (semi-supervised learning).

The data is then classified for intents (e.g., recommend, advise, question), entities (e.g., guidelines, recommendations, facts), and matched with the dialogues, using ML algorithms as illustrated in figure 7 (cognition & intelligence). In addition, the processed data used for training is also going through checks and authentications (QA) with the relevant feedback given back (cognition and analysis) for future use in the conversations with the users.

Such examples would typically be found in the use of Public Warning systems where autorities and first responders reach out to potentially large amount of people, for example via Location-based

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

SMS. The purpose could to seek feedback from the target population, request for information related to a certain incident (e.g., missing child situation), or to ask the target population to report back if they need help, information or other type of assistance. The challenge with asking for feedback from large amounts of people is obviously how to process, interpret and follow-up on all responses. In particular during critical events this must be done without any delay and with precision. AI-assisted handling of such feedback is an obvious alternative to allow authorities and first responders to handle the volume of information. The AI capability in this case would be more oriented towards structuring incoming data by identifying pre-defined answers and intents (e.g. bing able to interpret many variations of answers to predefined questions, such as "do you need medical assistance at your location?"). The AI-enabled solutuon would be able to group, categorize and enrich responses such that autorities and first responders more efficiently could decide follow-up action.

5.6.4 Various adaptation models

The blueprint in figure 7 illustrates an innovative approach for employing an AI-enabled chatbot by authorities and first responders. While this blueprint represents a complex AI-enabled chatbot with many features and technological capabilities, the solutions reviewed as part of this deliverable showed simpler uses. Therefore, on the one hand, one of the goals of this deliverable is to suggest state-of-the-art technology as an inspiration for authorities and first responders. However, on the other hand, based on findings from deliverable 2.4 and consultations made with several professionals and members of the Ki-CoP, we wish to suggest simpler ways to adopt this blueprint.

One way to do so is by limiting the types of datasets and the scope of the chatbot. For example, to make it a FAQ chatbot or a fact-checking chatbot. In this way, the complexity of the chatbot decreases, but it can still offer substantial value to the public. By employing this strategy, we also limit the necessary technological capabilities even in other components of the chatbot (e.g., fewer ML algorithms that we need to use, no need for monitoring & reporting, etc).

A second way to do so is by limiting the input types to text only, or, as many other chatbots which were reviewed are doing, or to a closed list of options, whether visual or not. This also reduces the complexity of several components, requires fewer ML algorithms and follows a more structured process, using tree sorting.

A third way is to omit the logging & documenting, monitoring & reporting and QA components. On the one hand, it, again, simplifies the chatbot's design, reducing its complexity. However, on the other hand, by doing this, we lose significant components that contribute to the work of the chatbot.

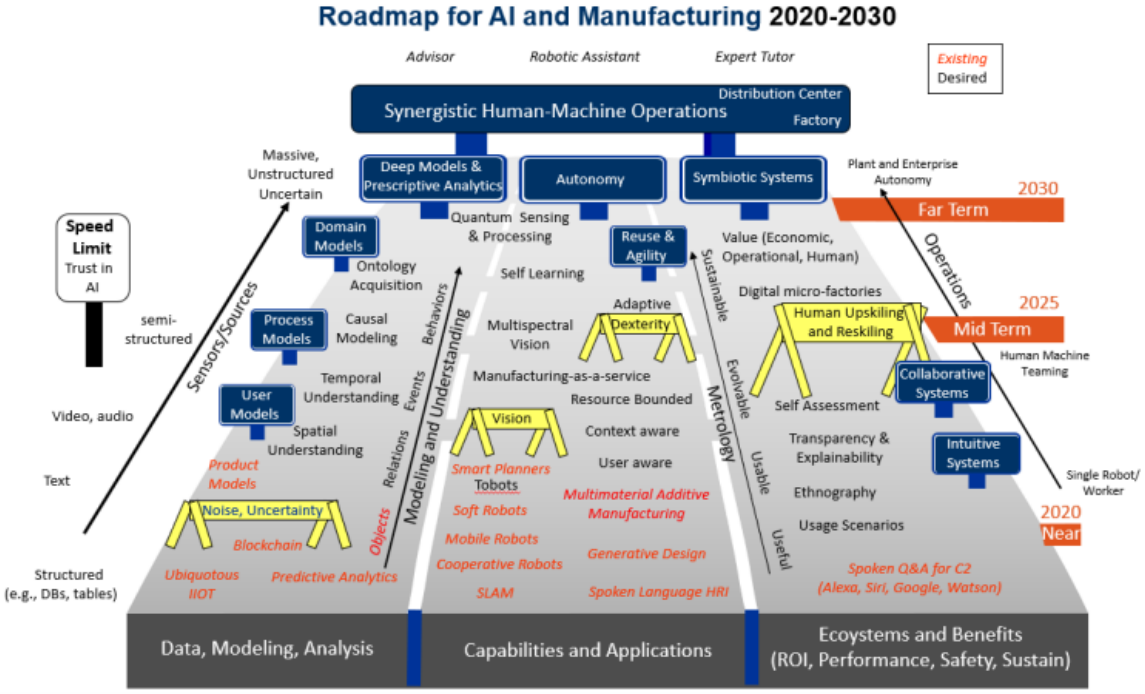
The fourth and last option is related to the learning component the chatbot uses. This blueprint does not necessarily aim for a specific type of ML learning type. While unsupervised and semi-supervised learning can broaden the scope of the chatbot, they also complicate the process, compared, for example, to supervised learning, relying solely on tagged data. In addition, supervised learning, compared to unsupervised and semi-supervised learning, reduces the chances for bias based on unknown data from various external untagged sources.

5.6.5 Roadmap

The scientific review summarised the development of chatbots, conversational chatbots, and AI-enabled chatbots along the history. As AI-enabled chatbots progressed in line with the development of new technologies, they keep developing. For example, Figure 8, taken from Mayburi (2020), illustrates the roadmap in propitiating the technology in three categories: data, modelling and analysis; capabilities and applications; and ecosystems and benefits. The figure illustrates the roadmap, with the desired steps in the way.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Figure 8. Roadmap for AI manufacturing 2020-2030. Taken from Mayburi, 2020.



The existing solutions analysed in this deliverable, the organisational's preparedness levels and technological adoption of authorities and first responders in emergencies and disasters, and the review of the available state-of-the-art technology reflect many gaps. For example, between what technologies have to offer to what authorities and first responders are willing to adopt, between what AI-enabled chatbots require in order to function efficiently to the level of data organisation that authorities and first responders are in, from the technological affordances to the acceptance and engagement levels of the public and the level of responsibility, authorities and first responders are willing or should be willing, to accept.

Therefore, in addition to the technological roadmap, we offer a roadmap of adoption, suggesting the steps towards full adoption of the blueprint. The roadmap is divided into five categories that reflect the gaps between the state-of-the-art technology on the preparedness levels of authorities and first responders. The roadmap considers only available and developed technologies and does not draw the progress in years, presenting the steps and necessary time-frames needed between every step. As varied authorities and first responders might find themselves in different steps, it is recommended to use the roadmap individually. Figure 9 illustrates the roadmap.

Technological capability. As reflected in the solution review, the discussion of technological capabilities reflects the current technological level that authorities and first responders have. Most AI-enabled chatbots adopted a very narrow approach, mainly of a closed visual set of choice, that users could choose from and receive answers to their questions. Even when some AI-enabled chatbots extended the technological framework, it was still far from the presented state-of-the-art technology.

According to figure 9, the current (or initial) step authorities and first responders may be in is the "zero technological capability", meaning that they have never adopted or adapted any technology or process that can empower them forward into full adoption of a functioning AI-enabled chatbot, on the other end of the roadmap.

In between, there are several steps. After the initial phase, the second step is to adopt and adapt to various technological communication channels, from different social media, messaging apps, and more. This step allows authorities and first responders to extend their communication points with the public. The third step is adopting logging and documenting capabilities based on the online and

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

phone interactions with the public. The fourth step is adopting more analytical technologies and using ML algorithms, as presented in the blueprint. Finally, the fifth step is developing a basic model of the chatbot, based on the different implementation suggestions in the blueprint section, until reaching the sixth and final stage in the roadmap of full adoption of the chatbot's operation.

Trust. In order to adopt progressed technological capabilities, one of the essential factors is trust in the technology, in its ability to fulfil its goals and prevent errors, mistakes and failures. However, one of the significant barriers to adopting AI-enabled chatbots by authorities and first responders in their work in emergencies and disasters is a lack of trust in the technology/system. Therefore, another category in the roadmap is building trust by authorities and first responders.

Here, the initial level is of no trust, ranging up to complete trust, resulting in the adoption of the full capabilities of the blueprint. In between, there are different steps to measure the improvement in trust levels. The second stage, for example, is examined by the first stakeholders in authorities and first responders that agree to consider augmenting some of the roles done by humans to chatbots, even without taking any significant actions. The third stage can be agreeing to try a small chatbot pilot, with feedback that lets the stakeholders review the criteria that can help them increase their trust levels. In the fourth stage, the pilot is already a part of the organisations' working process, while in the fifth stage, the stakeholders have enough trust to use more datasets and rely on more issues to be dealt by the chatbot. The last and sixth stage represents complete, or almost complete, trust levels in the AI-enabled chatbot.

User perspectives. The discussion in the literature about users perspectives towards AI-enabled chatbots, in general, showed that the interaction with robots is well recognised but sometimes perceived to be unpleasant. In deliverable 2.4, although not focused on AI-enabled chatbots, interviewees from authorities and first responders mentioned the concern of user responses to the use of technology by authorities and first responders during emergencies. In addition, in several professional meetings conducted as part of deliverable 3.2, significant concerns were raised regarding how users will perceive the use of AI-enabled chatbots during emergencies and disasters. The concerns led to the claim that instead of contributing to societal resilience, the use of chatbots in such situations may actually decrease it.

Therefore, the roadmap regarding user perspectives is not solely concerned with how users perceive AI-enabled chatbots operated by authorities and first responders for emergencies and disasters but with how authorities and first responders perceive the public's potential concerns. Here, the first stage reflects the most significant concerns, believing that the public will have a very negative perspective. On the other hand, the "final" stage in the roadmap is not zero or "zero concern", but a minimal concern that allows the full adoption of the framework, with enough engagement and interaction with the public, while still allowing amendments and changes.

In between these approaches, the second stage focuses on the belief that a small group of users will agree to try the AI-enabled chatbot, while the central part of the public will not only ignore it but will also severely reject its idea. The third stage is more balanced whereas the group of people willing to use the chatbot somewhat grows, while the rest of the public reduce their rejection. In the fourth stage, the positive change is expressed not just in attitudes and willingness to try the chatbot but also in a significant increase in the number of people using the chatbot. Finally, in the fifth stage, the increase in numbers is also expanded to additional features of the chatbot that are in use by the public. Those actions have a direct effect on how stakeholders in authorities and first responders perceive the user perspectives.

Information management and datasets preparation. Deliverable 2.4 showed that authorities and first responders may have big datasets, from different types: from call transcriptions, through books of guidelines, documents of requests received by the public, research papers and more. In a significant number of authorities and first responders, those datasets are documented and logged, with manual learning and almost no computerised data processing. Here, one of the goals of the steps in the roadmap is to prepare those datasets and make them ready for integration in the AI-enabled chatbot infrastructure.

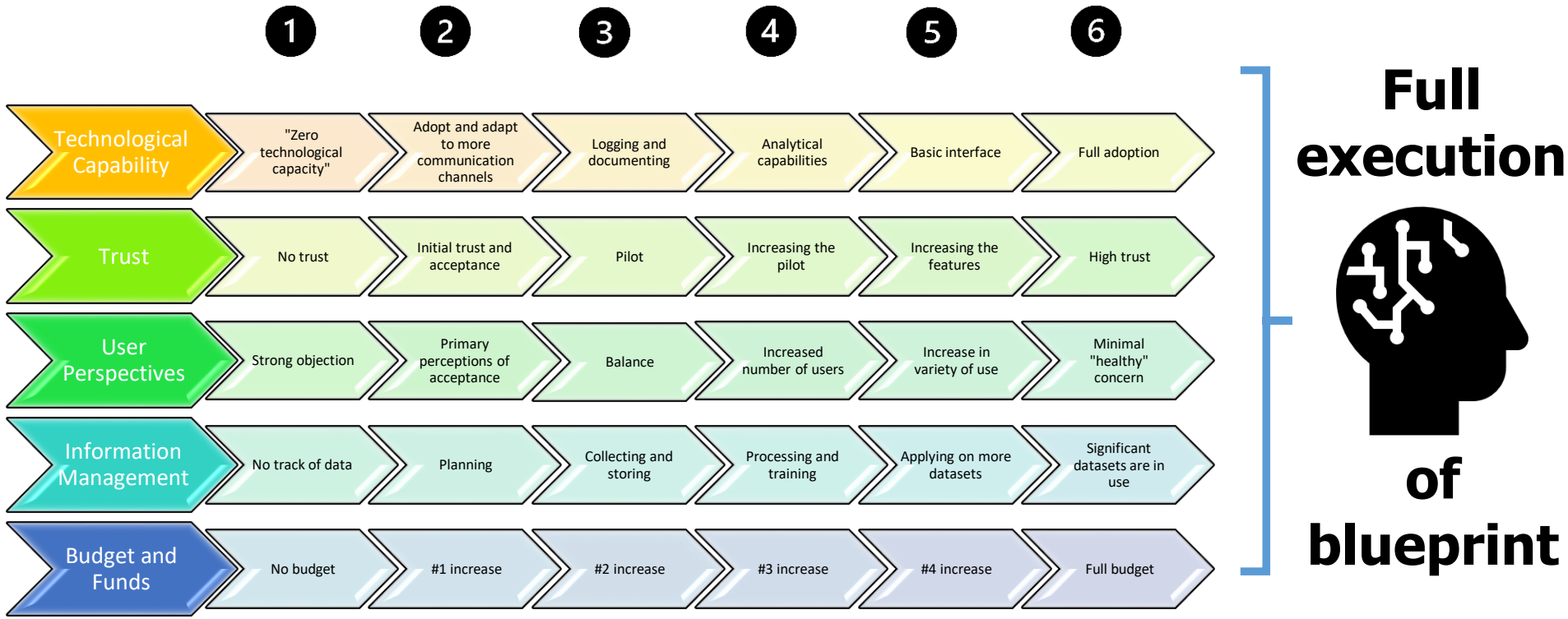
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

This category includes several steps. In the first and initial steps, the organisation has no track of logged data from almost any kind. The second stage is planning, where the organisation can start creating relevant datasets, and how and where it can store the data. In the third stage, the organisation starts collecting the data, storing it in the different datasets, and expanding it to other categories. In the fourth stage, the organisation start processing the datasets for training the chatbot, for example, by letting it use transcribed calls to learn and create relevant dialogues and chat scenarios. Finally, in the fifth stage, more datasets are prepared for training and extracting by adding more relevant processing services, up to a significant amount which characterises the final step.

Budget and allocation of funds. The last category is related to budget and allocation of funds. Although not as costly as before, with more affordable options, authorities and first responders' designing, programming, and implementing AI-enabled chatbots necessitates a budget. As authorities and first responders have different organisational sizes and budgets, the roadmap reflects increasing the budget from no funding whatsoever to the sixth step of budget increase for developing the chatbot.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Figure 9. A roadmap for the complete execution of the blueprint.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

5.6.6 Implementation stages

Last, Appendix D illustrates suggested implementation stages of AI-enabled chatbots by authorities and first responders. The illustration summarises the proposed steps, based on this deliverable, with the relevant questions and key points to consider. The figure is illustrated in steps but could also be used as a list of recommendations, conducted differently.

The first step refers to the framework of the AI-enabled chatbot. It includes answering general questions such as what we want the chatbot to do, and not less important – what we do not want it to do. In this step, it is essential to define the scope of the chatbot's activities and actions, determine the goals and objectives, define the necessary budget for investing in the chatbot, and whether it is possible to develop it inside the organisation or with the service of external sources.

The second step deals with the communication channels of the organisation. It includes first mapping of the current channels in order to understand how to integrate them in the chatbot's system. However, it also includes considering additional channels that the organisation might not have used before, due to lack of workforce, but will now utilize with the AI-enabled chatbot.

The third step relates to the "brain" of the AI-enabled chatbot. This is one of the most essential and fundamental components. This is when the organisation needs to decide about the chatbot's technological requirements, how it wants to communicate with its users, and what available services it needs to use to facilitate the chatbot.

The fourth step concerns security and privacy. In this step, important decisions need to be made regarding security measures. The data used as part of the chatbot's system needs to be identified, matched with the appropriate security measures, examined according to the relevant regulations and following it, creating groups of users – with the appropriate level of access to each group.

In the fifth step, the datasets of the organisation need to be prepared. First, current and ongoing datasets should be mapped and processed to fit the chatbot's architecture. Second, the organisation should also think about creating new datasets that will be generated more easily by the chatbot, and in addition, will improve its functionality.

The sixth step is about the technological capabilities. It defines what the organisation wishes to ahcieve. For example, is it interested to authenticate news for the users, track and warn them about mis- and false information? Does it want to allow the chatbot QA abilities? To identify when the user is in stressful situations and alert about it? For each capability, organisations should choose the relevant services from those which were reviewed.

The seventh step relates to monitoring. It includes questions regarding what information is about to be logged and for what purposes. Here, the organisation needs to decide about the monitoring processes and their outcomes, matching them with the facilitating technologies.

The eighth step is about defining the relationships between the different components of the AI-enabled chatbot and the relevant processes. For example, what processes affect one another and what processes can only affect but not be affected by others.

The ninth and tenth steps are almost connected. The ninth, building, relates to the building process of the chatbot – programming it according to the framework. The tenth step is the implementation process, which includes human monitoring of the chatbot's activity until satisfactory results and accuracy are achieved.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

6 DISCUSSION

Deliverable 3.2's main goal was to propose new innovative directions for using social media in building societal resilience. In particular, the goal was to suggest a design concept for an AI-enabled chatbot's blueprint/framework that partly augments the communication roles of authorities and first responders. An AI-enabled chatbot that can converse with users and provide essential information in all phases of emergencies and disasters contributes to building societal resilience.

In the first stage, we introduced AI and ML in general and their relationship to AI-enabled chatbots. Then, we started with setting the borders regarding what could and could not be done by this technology, as well as what should and should not be done. After that, we focused on two significant challenges that are at the core of this deliverable.

The first challenge concerns the relationship between the instrumental problem-solving field of AI and the more organic nature of societal resilience. On the one hand, AI technologies are faced with a problem that they need to solve or provide an answer for. For example, based on big datasets, calculating the probability of an occurrence of a disaster in a given period or finding a solution for water disasters. However, on the other hand, societal resilience is more complex, involving a wide range of variables, and takes effect in many areas of life, environmental, mental, health, economical and more.

In order to face this problem, we suggested adopting Alessi et al.'s (2018) framework, also used by the European Commission Joint Research Centre. According to this framework, by adopting AI solutions, each system component should be paired with a factor of societal resilience. In the case of the suggested blueprint, we focused on the role of improving the communication process, providing accurate and immediate information, solving cases of misinformation and false news, allowing the public to transfer information and data to authorities and first responders, and to define how this data can be processed and improve the communication system – all contributors of societal resilience. All these factors of societal resilience can be achieved by different components of the system, from dynamic datasets that can provide immediate and accurate information, NLP algorithms that allow the users to express themselves, Q&A and search functions that organise the information for the users and more.

The second question concerns whether AI-enabled chatbots can augment part of the roles of human workers of authorities and first responders in all phases of emergencies and disasters. This is crucial due to one of the basic needs that this deliverable faces – to reduce the surge of calls on emergency call centres during adversities. Here, we showed that current AI-enabled chatbots are very conservative in their approach, providing a minor replacement for routine information from the call centres and other online communication channels of authorities and first responders. However, extending the abilities of AI-enabled chatbots should be done with caution, preventing possible biases and mistakes.

The main section of this deliverable reviewed the current AI-enabled chatbots solutions, their strengths and weaknesses, pointed out what is already functioning and what is not and what can be learnt from the current ML approaches. Finally, we suggested an approach that merges the varied aspects, starting from closed-scenario parts in the chatbot and allowing freestyle conversations that enable the users to elaborate on the information they wish to explore. In this way, we balance the current conservative nature of the AI-enabled chatbots, as used by authorities and first responders, setting the boundaries of what can be done and what as yet we are unable to do, concerning the cutting-edge technology of AI-enabled chatbots.

We suggested a roadmap for adopting AI-enabled chatbot solutions, step by step, with milestones to follow. This roadmap was based on the solutions review that suggested that most authorities and first responders are still in the very early stages of considering such innovative solutions. In the following sections of the discussion, we highlight several key points raised concerning the blueprint.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

6.1 Alternative or supplementary channels for emergency call centres

As mentioned above, an important goal of the deliverable, and the AI-enabled chatbot in particular, is to examine whether and how AI-enabled chatbot can augment, as an alternative or supplementary channel, the role of human operators in authorities and first responders call-centres, or other communication channels (e.g., new media accounts, websites). We pointed out that one of the major problems during emergencies and disasters, and in several cases also before and after, is a surge of requests that can lead to the collapse of the infrastructure (Carenzo, Costantini, Greco, Barra, Rendiniello, Mainetti & Cecconi, 2020; Hrabi, 2020). AI-enabled chatbots can contribute to the idea of infrastructure resilience – to provide information for non-emergency inquiries, leaving the human operators to handle emergency calls, complex inquiries, special populations in a lower volume of calls.

In our review of AI-enabled chatbot solutions, we assumed that one of the reasons for the conservative approach is a concern of not identifying these emergency calls or other users that need human attention. Therefore, to avoid this barrier but still augment some of the roles of the human operators, we suggested a system that allows a wide range of possibilities and answers but includes options of fallback and a component of live monitoring. These allow the user to ask to be transferred to a human operator, if possible, or to be given an emergency phone number to call and monitors distress chats and other emergencies that should alert the first responders.

This idea is partly based on suggestions in the literature. As shown, AI capabilities are already somewhat used in emergencies and disasters (Heires, 2017; Imran, Castillo, Lucas, Meier & Viewg, 2014; Richardson, 2019). In addition, the suggested AI-enabled chatbot uses several suggestions by Chaudhry & Yuskal (2019), enhancing the flow of information between authorities, first responders and the public.

The review of current solutions, existing blueprints and ML approaches showed a wide range of technologies that can be implemented in AI-enabled chatbots that may partly augment the role of human operators, allowing to reduce the volume of calls. They range from NLP algorithms that allow the interpretation of inquiries by the public, through processing systems that help organisations to work better and arrange their datasets, up to technological capabilities that create Q&A platforms, search in databases and datasets, visualise information and other actions that not just augment the role of call-centres, but can also provide new advantages.

However, as mentioned several times during the presentation of the blueprint, the adoption process should be done cautiously and slowly. Using such technological capabilities can risk bias and mistakes that may prevent the augmentation process and lead to worse results and inaccurate information.

6.2 Addressing society's needs and expectations

Another critical aspect of the blueprint is related to the needs and expectations of the public, on the one hand, and the perceptions of authorities and first responders on the other hand. The suggested blueprint considers the needs and expectations reviewed in deliverables 1.2 and 1.3 and the approach of authorities and first responders, as presented in deliverables 2.1 and 2.4. On the one hand, as reviewed before, there is a need for rapid availability of information, a two-way flow of communication, accurate data and more, along with other influential, integrative and even escapist needs. On the other hand, the size, lack of communication guidelines and fear of adopting innovative solutions, as expressed by the interviews conducted in deliverable 2.4, along with the different ways organisations seek to engage the public in emergencies and disasters, as discussed in deliverable 2.1, must be considered.

In order to address this complex approach, we suggested referring to it in several places in the blueprint. From the society's point of view, the various components suggested in the blueprint can

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

introduce many opportunities to fulfil these needs, and provide a narrower channel for those who wish to avoid the sense of overload.From the authorities and first responders' perspective, we suggested a roadmap and a step-by-step approach, followed by the possibility of adopting only part of the blueprint, which may help them adapt and thus adopt the suggested model.

The step-by-step adoption suggestion also allow adressing the question of digital literacy. Allowing the adoption of more simple interfaces of the chatbot, aiming to address more digitally iliterate populations, up to more sophisticated populations.

6.3 The contribution to all phases of emergencies and disasters

Another essential question regarding the blueprint is the possible contribution of the AI-enabled chatbot in every phase of emergencies and disasters. On the one hand, it can be suggested that the chatbot will be more popular not during the active phases of emergencies and disasters, when the public's inquiries might be more urgent and necessitate human assistance.On the other hand, despite this need, surges of calls and requests usually occur during ongoing emergencies and disasters, and thus the assistance of an AI-enabled chatbot is more noticeable.

The suggested blueprint provides directions for AI-enabled chatbots that can contribute to the public's needs and expectations and build societal resilience in all phases of emergencies and disasters. As in other disaster-related activities, the crucial part of building the chatbot is in the initial phases, in routine times before crises occur. This gives time to the public to get used to the chatbot, examine its advantages and disadvantages and provide feedback. The authorities and first responders are thus offered a more calm environment to design, build, inspect, and revise the system.

The contribution of the chatbot in the early phases, before emergencies occur, is also essential for the informational and other needs of the public. It can provide crucial information that helps the public prepare for emergencies and be more accessible (e.g., in the middle of the night, when a phone call is impossible, or even save time). The same implies for the later stages of crises, after disasters.

During disasters, the role of the AI-enabled chatbot changes, but not significantly. At these phases, it can be expected that more users will address the chatbot with repeating, and maybe more complex, questions. Based on its activity during the early stages, the chatbot should be trained and more functioning during disasters, even at this stage, more effectively reduce the load on call centres and identify urgent requests. These factors were addressed in the monitoring component of the chatbot.

6.4 Fact-checking, mis- and false information and fake news in social media

Another crucial aspect of the blueprint, as previously addressed, is the work of validating information. Concerning emergencies, in all phases of disasters, fact-checking, mis- and false information, and fake news play an important role (Shu, Wang, Lee & Liu, 2020; Stahl, 2006). They threaten societal resilience, especially during emergencies and disasters, when people are under pressure and may lose some of their ability to be critical, and therefore are exposed to false information "attacks" (Kertysova, 2018). Thus, an essential aspect of the blueprint is to suggest how to neutralise false information (Kim, Lyu & Gong, 2020; Puildo, Villarejo-Carballido, 2020).

The blueprint suggested what technologies could be adopted to achieve this purpose, including a case study, based on the suggestions in the scientific literature (Aphiwongsophon & Chongstitvatana, 2018; Graves, 2018; Kertysova, 2018). However, while AI-enabled chatbots and other technological capabilities allow fact-checking to a particular stage, we suggested doing it in a more limited

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

capacity, due to several reasons which were described, such as the opacity of these technologies and possible biases (Ahmen, Aliabouh, Donepudi & Choi, 2021; Sumpter, 2018).

Fact-checking for false news should be semi-automatic and focus on identifying and alerting, leaving the part of ruling out false information to human fact-checkers. The relevant technologies which were reviewed can contribute to several areas. First, they can help identify repeating questions relating to similar pieces and sources of information. Second, they can track social media for specific channels, or in general, and gather data about trending topics, hashtags or any other symptoms that demand inspection. It should though be noted that, in the end, the result should be a list of "suspicious" posts or other types of content to be classified by human fact-checkers. Following this, the content can serve as input to the system and extracted to respond to questions and inquiries.

6.5 AI-ENABLED CHATBOTS, DIVERSITY AND SOCIETAL RESILIENCE

The last issue is the relationship between the activity of AI-enabled chatbots, diversity and societal resilience. The reviewed solutions showed addressing diversity mainly concerning languages and nationality. However, as noted in previous deliverables, there was a minimal reference to diversity, such as gender. Two chatbots, among them only one that was still active, addressed women-only issues.

This lack of diversity can also lead to biases in AI-enabled chatbots that cannot handle the diverse population. However, this is because learning is enabled through what is currently done – actions, datasets, and other sources. Furthermore, if those sources are not diversity-originated, then new material produced will only replicate the situation.

This is not a pivotal point to be addressed in the narrow context of the blueprint, but more generally, in the context of authorities, first responders, communication and diversity.Changes in this subject will also affect the activity of the AI-enabled chatbots.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

7 STRENGTHS & LIMITATIONS

7.1 BLUEPRINT'S LIMITATIONS

The suggested new innovative directions and the blueprint have several limitations to be considered. One of the limitations is the lack of a current operational, full functioning model of an AI-enabled chatbot, similar to the blueprint. In addition, the current blueprint resembles more current cutting-edge technology chatbots, which are not yet in use concerning emergencies, disasters and societal resilience. Therefore, as in other theoretical suggestions in the literature, it was not yet fully examined.

Another limitation is related to the process in which this project was conducted. While in this project, we reviewed current technologies and solutions and consulted with members of the partners of ENGAGE, including the Ki-CoP members, it did not include two crucial components: formative assessment research or consultation of the blueprint and receiving feedback from the public, for example, in the form of a low-fi prototype. This feedback is crucial for suggesting new directions and could be done in the re-evaluation process of this deliverable as part of deliverable 3.4.

Last, another limitation this project faced is the significant gap between the cutting edge, state-of-the-art AI technologies and AI-enabled chatbots that were developed in the past years and the current readiness of authorities and first responders to adopt them. This significant gap can result in barriers in adopting the blueprint, which the suggested roadmap and step-by-step adoption process should contribute to.

7.2 BLUEPRINT'S STRENGTHS

On the other hand, as mentioned, one strength of the deliverable is the adoption of cutting-edge, state-of-the-art technologies to by used by authorities and first responders in building societal resilience. The deliverable maps the technological capabilities, services and products of leading companies in the market that develops professional AI-enabled chatbots, along with suggestions on how to implement them in the field of societal resilience.

In addition, the roadmap, partial adoption suggestion and the step-by-step process is another strength of the blueprint. It allows savvy technological organisations to adopt this blueprint with different preparedness levels, from more considerable sizes to smaller organisations.

The last strength is in the inclusion of fact-checking and false news treatment. While prompt work has been done in this area, as shown in the literature, this is one of the first suggestions to implement such technologies in authorities and first responders concerning emergencies, disasters and societal resilience. Since false information dramatically affects the public during disasters, we believe that this implementation of technologies and models can significantly contribute to societal resilience.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

8 CONCLUSIONS

Deliverable 3.2 followed one of ENGAGE's objectives, to produce validated actionable knowledge on societal resilience by demonstrating the benefits and impact of the project solutions in different types of disasters. In addition, it aimed at finding the best practices for communication and social media. The goal of a design concept and a blueprint of an AI-enabled chatbot for emergencies and disasters, addressing questions of design and implementation of the AI-enabled chatbot, aimed to suggest new directions for innovative solutions that offer four main contributions.

The first is the ability to test or revise the assumption of ENGAGE that AI-enabled technologies can contribute to building societal resilience. This deliverable reviewed many of the constraints and barriers this assumption has, from low cooperation of both entities (authorities and first responders) and the public to technological limitations through possible biases and mistakes. However, while these barriers are significant, we offered several solutions and presented various technologies that can help overcome these barriers. Therefore, this assumption should be optimistic, but with a cautious adoption of the blueprint suggested in this deliverable.

The second contribution is that the blueprint refers to an AI-enabled chatbot, allowing emergency authorities to provide a contextual, online, and zero-delay response to the public before, during, and after emergencies. The suggested blueprint is based on current technologies that can facilitate this objective. As suggested by the blueprint, the online processing of information and the enrichment of the authorities and first responders datasets facilitate a contextual, online and zero-delay response.

The third contribution is innovative solutions for neutralising false messages during disasters, based on the rapid detection and tracking of tending misinformation on social media. However, as highlighted in the objectives section, the work of the AI-enabled chatbot is not in actively distributing messages and notifications about false information but in detecting a surge in social media regarding specific issues, internally analysing it and operating human fact-checkers to verify or refute them. Therefore, the conclusion here is that the AI-enabled chatbots should act, at least at this stage, as identifiers and "under the hood" analyser of the information, rather than to decide automatically about it.

Last, the fourth contribution is developing innovative solutions and citizens' engagement and transfer of knowledge from research to the public, leveraging the project's suggested directions to relevant population groups according to their specific needs and expectations. This deliverable is part of ENGAGE's knowledge platform and should serve as the basis for authorities and first responders who wish to develop other AI-enabled chatbots. It will be integrated as part of D5.4, the website and knowledge platform, and 5.5 knowledge and innovation community.

8.1 NEXT STEPS

Deliverable 3.2 will be followed and further developed in a later stage of the project in deliverable 3.4, which will revise the results based on WP4. As part of of the next step, we will introduce the blueprint to groups of AI experts, using the modified Delphi method, in order to receive feedback and improve the blueprint. We will also conduct a formative evaluation of the blueprints, by additional emergency professionals from authorities and first responders, along with potential users, to address the differtent points for improvement. For example, we will explore the possibility to advance from more closed AI-enabled chatbots to more open chatbots, using significant datasets.

In addition, as mentioned before, the progress of AI technologies is exponential. Two generations every six months. Therefore, deliverable 3.4 will incorporate the new affordances which will be offered by AI technologies at that time.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Last, as part of deliverable 3.4 we intend to expand deliverable 3.2 to additional possibilities beyond AI-enabled chatbots, as mentioned in previous sections, and discuss sentiment analysis, detecting emerging topics in social media, gathering information from user discussions and more.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

9 REFERENCES

Abd-Alrazaq, A. A., Alajlani, M., Ali, N., Denecke, K., Bewick, B. M., & Househ, M. (2021). Perceptions and opinions of patients about mental health chatbots: Scoping review. Journal of medical Internet research, 23(1), e17828.

Alessi, L., Benczur, P., Campolongo, F., Cariboni, J., Manca, A. R., Menyhert, B., & Pagano, A. (2018). *The resilience of EU Member States to the financial and economic crisis. What are the characteristics of resilient behaviour?* Joint Research Centre (Seville site).

Almalki, M., & Azeez, F. (2020). Health Chatbots for Fighting COVID-19: A scoping review. *Acta Informatica Medica*, *28*(4), 241.

Al-Rifaie, M. M., & Bishop, M. (2015). Weak and strong computational creativity. In *Computational creativity research: Towards creative machines* (pp. 37–49). Springer.

Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, B., De Prato, G., Feijoo, C., Macias, E. F., Gutierrez, E. G., Portela, M. I., & Junklewitz, H. (2018). *Artificial Intelligence: A European Perspective*.

Aoki, N. (2020). An experimental study of public trust in AI chatbots in the public sector. Government Information Quarterly, 37(4), 101490.

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. Telematics and Informatics, 54, 101473.

Barojan, D. (2021). Building Digital Resilience Ahead of Elections and Beyond. In Disinformation and Fake News (pp. 61-73). Palgrave Macmillan, Singapore.

Battineni, G., Chintalapudi, N., & Amenta, F. (2020). *AI chatbot design during an epidemic like the novel coronavirus*. *8*(2), 154.

Brandtzaeg, P. B., & Følstad, A. (2018). Chatbots: changing user needs and motivations. Interactions, 25(5), 38-43.

Carenzo, L., Costantini, E., Greco, M., Barra, F., Rendiniello, V., Mainetti, M., Bui, R., Zanella, A., Grasselli, G., & Lagioia, M. (2020). Hospital surge capacity in a tertiary emergency referral centre during the COVID-19 outbreak in Italy. *Anaesthesia*, *75*(7), 928–934.

Chaudhry, S., & Yuksel, M. (n.d.). Using Social Media for Crowd-Sourced Public Safety. *MMTC Communications-Frontiers*.

Cheng, Y., & Jiang, H. (2020). AI-Powered mental health chatbots: Examining users' motivations, active communicative action and engagement after mass-shooting disasters. Journal of Contingencies and Crisis Management, 28(3), 339-354.

Clark, A., Fox, C., & Lappin, S. (2013). *The handbook of computational linguistics and natural language processing*. John Wiley & Sons.

Coeckelbergh, M. (2020). AI ethics. MIT Press.

Crutchfield, A. S., & Harkey, K. A. (2019). A comparison of call volumes before, during, and after Hurricane Harvey. The American journal of emergency medicine, 37(10), 1904-1906.

De Cicco, R., e Silva, S. C., & Alparone, F. R. (2020). Millennials' attitude toward chatbots: an experimental study in a social relationship perspective. International Journal of Retail & Distribution Management.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Dolhansky, B., Bitton, J., Pflaum, B., Lu, J., Howes, R., Wang, M., & Canton Ferrer, C. (2020). The deepfake detection challenge dataset. arXiv e-prints, arXiv-2006.

Figueira, Á., & Oliveira, L. (2017). The current state of fake news: challenges and opportunities. Procedia Computer Science, 121, 817-825.

Følstad, A., Skjuve, M., & Brandtzaeg, P. B. (2018). *Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design*. 145–156.

Frank, J., Eisenhofer, T., Schönherr, L., Fischer, A., Kolossa, D., & Holz, T. (2020, November). Leveraging frequency analysis for deep fake image recognition. In International Conference on Machine Learning (pp. 3247-3258). PMLR.

Gupta, A., Hathwar, D., & Vijayakumar, A. (2020). Introduction to AI chatbots. *Int J Eng Res Technol (IJERT)*, *9*(7), 255–258.

Hakak, S., Khan, W. Z., Bhattacharya, S., Reddy, G. T., & Choo, K. K. R. (2020, December). Propagation of fake news on social media: challenges and opportunities. In International Conference on Computational Data and Social Networks (pp. 345-353). Springer, Cham.

Hasan, I., Rizvi, S., Jain, S., & Huria, S. (2021). *The AI-enabled Chatbot Framework for Intelligent Citizen-Government Interaction for Delivery of Services*. 601–606.

Hayes, J. L., Britt, B. C., Evans, W., Rush, S. W., Towery, N. A., & Adamson, A. C. (2021). Can Social Media Listening Platforms' Artificial Intelligence Be Trusted? Examining the Accuracy of Crimson Hexagon's (Now Brandwatch Consumer Research's) AI-Driven Analyses. Journal of Advertising, 50(1), 81-91.

Heires, K. (2017). Technology to the Rescue. *Risk Management*, *64*(11), 32–35.

Herrera, L. C., Majchrzak, T. A., & Thapa, D. (2021, September). Ecosystem of Social Media Listening Practices for Crisis Management. In Conference on e-Business, e-Services and e-Society (pp. 710-722). Springer, Cham.

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. Computers in human behavior, 49, 245-250.

Hrabí, M. (2020). Call centres: Going voice-first in the post-Covid world. *Biometric Technology Today*, *2020*(8), 10–12.

Ignatidou, S. (2019). *AI-driven personalisation in digital media: Political and societal implications*.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260.

Kertysova, K. (2018). Artificial Intelligence and Disinformation: How AI Changes the Way Disinformation is Produced, Disseminated, and Can Be Countered. *Security and Human Rights*, *29*(1–4), 55–81.

Laumer, S., Maier, C., & Gubler, F. T. (2019). Chatbot acceptance in healthcare: Explaining user adoption of conversational agents for disease diagnosis.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. Science, 359(6380), 1094-1096.

Lei, S. I., Shen, H., & Ye, S. (2021). A comparison between chatbot and human service: customer perception and reuse intention. International Journal of Contemporary Hospitality Management.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Liu, B. (2021). 'Weak AI' is Likely to Never Become" Strong AI", So What is its Greatest Value for us? *ArXiv Preprint ArXiv:2103.15294*.

Luccioni, A. S., Pham, K. H., Lam, C. S. N., Aylett-Bullock, J., & Luengo-Oroz, M. (2021). Ensuring the Inclusive Use of Natural Language Processing in the Global Response to COVID-19. arXiv preprint arXiv:2108.10791.

MacKenzie, A. (2004). Infrastructural Resilience: Lessons from the WTC Collapse. *E. Nz Magazine: The Magazine of Technical Enterprise*, *5*(1), 4–13.

MacKenzie, C. A., & Barker, K. (2013). Empirical data and regression analysis for estimation of infrastructure resilience with application to electric power outages. *Journal of Infrastructure Systems*, *19*(1), 25–35.

Madianou, M. (2020). Non-human humanism: When AI for good turns out to be bad. AoIR Selected Papers of Internet Research.

Mahabub, A. (2020). A robust technique of fake news detection using Ensemble Voting Classifier and comparison with other classifiers. SN Applied Sciences, 2(4), 1-9.

Maksutov, A. A., Morozov, V. O., Lavrenov, A. A., & Smirnov, A. S. (2020, January). Methods of deepfake detection based on machine learning. In 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus) (pp. 408-411). IEEE.

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of machine learning*. MIT press.

Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. Digital health, 5, 2055207619871808.

Natale, S. (2020). To believe in Siri: A critical analysis of AI voice assistants.

Natale, S. (2019). If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. *New Media & Society*, *21*(3), 712–728.

Palanica, A., Flaschner, P., Thommandram, A., Li, M., & Fossat, Y. (2019). Physicians' perceptions of chatbots in health care: cross-sectional web-based survey. Journal of medical Internet research, 21(4), e12887.

Patwa, P., Sharma, S., Pykl, S., Guptha, V., Kumari, G., Akhtar, M. S., ... & Chakraborty, T. (2021, February). Fighting an infodemic: Covid-19 fake news dataset. In International Workshop on Combating On line Ho st ile Posts in Regional Languages dur ing Emergency Situation (pp. 21-29). Springer, Cham.

Rao, N. R. (Ed.). (2016). Social media listening and monitoring for business applications. IGI Global.

Richardson, M. (2016). The role social media played in wake of Orlando nightclub massacre. *Bizjournals. Com*.

Rodríguez Cardona, D., Werth, O., Schönborn, S., & Breitner, M. H. (2019). A mixed methods analysis of the adoption and diffusion of Chatbot Technology in the German insurance sector.

Sample, C., Jensen, M. J., Scott, K., McAlaney, J., Fitchpatrick, S., Brockinton, A., ... & Ormrod, A. (2020). Interdisciplinary lessons learned while researching fake news. Frontiers in Psychology, 11, 2947.

Saravi, S., Kalawsky, R., Joannou, D., Rivas Casado, M., Fu, G., & Meng, F. (2019). Use of artificial intelligence to improve resilience and preparedness against adverse flood events. *Water*, *11*(5), 973.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–424.

Simon, T. (2013). Developing an Emergency Response Framework (procedure), Leveraging Social Media Collaboration Tools, to Promote Regional Cooperation Following an Earthquake. Ben-Gurion University of the Negev.

Smith, S. W., Braun, J., Portelli, I., Malik, S., Asaeda, G., Lancet, E., ... & Goldfrank, L. R. (2016). Prehospital indicators for disaster preparedness and response: New York City emergency medical services in hurricane sandy. Disaster medicine and public health preparedness, 10(3), 333-343.

Stroud, S. R. (2019). Pragmatist media ethics and the challenges of fake news. Journal of Media Ethics, 34(4), 178-192.

Thies, I. M., Menon, N., Magapu, S., Subramony, M., & O'neill, J. (2017, September). How do you want your chatbot? An exploratory Wizard-of-Oz study with young, urban Indians. In IFIP Conference on Human-Computer Interaction (pp. 441-459). Springer, Cham.

Tsai, M.-H., Chen, J. Y., & Kang, S.-C. (2019). Ask Diana: A keyword-based chatbot system for water-related disaster management. *Water*, *11*(2), 234.

Van Engelen, J. E., & Hoos, H. H. (2020). A survey on semi-supervised learning. *Machine Learning*, *109*(2), 373–440.

Van Wezel, M. M., Croes, E. A., & Antheunis, M. L. (2020, November). "I'm Here for You": Can Social Chatbots Truly Support Their Users? A Literature Review. In International Workshop on Chatbot Research and Design (pp. 96-113). Springer, Cham.

Verma, A., Chandra, P., & Joshi, A. (2021). AI Chatbots, its Feasibility and Reliability in Modern World.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Nerini, F. F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nature Communications*, *11*(1), 1–10.

Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology Innovation Management Review, 9(11).

Yamin, M. M., Ullah, M., Ullah, H., & Katt, B. (2021). Weaponized AI for cyber attacks. Journal of Information Security and Applications, 57, 102722.

Zifu, C. (2016). Strong AI and super AI: technology rationality and its criticism. *Science and Management*, 05.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

10 APPENDICES

10.1 APPENDIX A: LIST OF SOLUTIONS

General		Technology					Approach		
Name (Organisation) (Country) (Period)	**Description**	**Types of Crises**	**Platforms**	**Datasets**	**Algorithms**	**Scenarios**	**Provides Information/Only Responds**	**Connection to other services/Reference to other sources**	**Strengths, Weaknesses and Gaps**
Coronavirus self-checker - CLARA (CDC) (USA) (2020-)	Allows the user to self-diagnose themselves with coronavirus symptoms	Health (Covid-19)	Web	Covid-19 decision protocol	Decision trees, based on a Microsoft technology	Close	Only respond to choices	For us residents, it provides relevant phone numbers of clinics	The chatbot allows for diagnosis but does not provide help for more complex situations
WHO health Alert (WHO) (International) (2020-)	Provides information related to covid-19 pandemic (vaccines, protection, statistics, etc.) in Arabic, English, French, Hindi, Italian, Spanish and Portuguese	Health (Covid-19)	Facebook Messenger, Viber, WhatsApp, Line	Covid-19 closed information guidelines	Decision Trees	Close	Only respond to choices	WHO website	Allows access to official health information, with constant updates, but still have limited options
DESI (KINOA) (Italy) (2020)	A chatbot aimed to reduce the workload on call centres. Provided answers to questions of the public regarding the coronavirus	Health (Covid-19)	Web (implemented on authorities and first responders' websites)	Q&A documents, dictionary of synonyms and antonyms,	NLP	Freewriting	Respond to open questions	No	Live agents monitored the chatbot for a month to test its ability to answer questions. Only after it was validated for more than 95% accuracy, the monitoring stopped. DESI was able to understand and answer the standard questions but had trouble with more complex questions that needed human response

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Coronavirus chatbot (MOH) (Israel) (2020-)	Provides basic answers to common topics and predefined questions, with the possibility to talk to a human operator on the chat	Health (Covid-19)	Web (implemented on authorities and first responders' websites)	Covid-19 treatment protocols + QnA	Decision trees, based on a Microsoft technology	Close	Respond to questions but allows asking to take to a human operator	Israeli MOH website	Integration between a closed chatbot with the possibility to talk to a human operator
IBM-Watson Covid-19 chatbot (International) (2020-)	Provides information and recommendations regarding the COVID-19 pandemic	Health (Covid-19)	Web (implemented on authorities and first responders' websites)	Covid-19 information sheets, QnA, Information provided by experts	NLP	Open	Allows free conversation about Covid-19 questions	Depends on the local implementation of the chatbot. It was implemented in several countries.	Based on IBM's technology and NLP capabilities. Allows free questions and broad topics, but still needs to be adapted to each country that uses it.
ADA (ADA) (USA) (2020)	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	WHO + CDC websites	Limited options
Apple	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices		Rich information with many possibilities but limited user control
Babylon	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	WHO + CDC websites	Limited options
Bobbi	Answering frequent questions about coronavirus	Health (Covid-19)	Web	QnA	NLP	Open	Responds to open questions		Limited options
Cleveland Clinic	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	WHO + CDC websites	Limited options
Corona Bot	Symptoms check for coronavirus	Health (Covid-19)	Facebook Messenger	Covid-19 treatment protocols + QnA	Decision trees + NLP	Close + open (depends on the question and situation)	Only respond to choices	WHO + CDC websites	Rich information with many possibilities but limited user control
HSE	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	No	Limited options
Covid-19 chatbot	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	WHO + CDC websites	Limited options

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Dubai Department of Health	Answering frequent questions about coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees + NLP	Close + open (depends on the question and situation)	Only respond to choices	No	Rich information with many possibilities but limited user control
e-Bot	Answering frequent questions about coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees + NLP	Close + open (depends on the question and situation – in most cases, typing is blocked)	Only respond to choices	No	Limited options
German Red-Cross	Answering frequent questions about coronavirus	Health (Covid-19)	WhatsApp	Covid-19 treatment protocols + QnA	Decision trees	Close (needs typing but from a limited list)	Only respond to choices	WHO + CDC websites	Limited options
Health Buddy	Answering frequent questions about coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	WHO + CDC websites	Limited options
Infermedics	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols	Decision trees	Close	Only respond to choices	No	Rich information with many possibilities but limited user control
Ivan Mask	Answering frequent questions about coronavirus	Health (Covid-19)	Telegram	Covid-19 treatment protocols + QnA	Decision trees + NLP	Close + open (depends on the question and situation)	Only respond to choices	No	Limited options
Martha	Symptoms check for coronavirus and answering frequent questions about coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees + NLP	Close + open (depends on the question and situation)	Only respond to choices	No	Limited options
Missouri Health Department chatbot	Symptoms check for coronavirus and answering	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	NLP	Allows open questions	Responds to open questions	No	Limited options

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

	frequent questions about the coronavirus					but asks specific questions to respond to			
MTI Singapore	Answering frequent questions about coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees + NLP	Close + open (depends on the question and situation)	Only respond to choices	No	Limited options
Providence	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	No	Limited options
Russian MOH	Answering frequent questions about coronavirus	Health (Covid-19)	WhatsApp	Covid-19 treatment protocols + QnA	Decision trees	Close (needs typing but from a limited list)	Only respond to choices	No	Limited options
Suve	Symptoms check for coronavirus and answering frequent questions about coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees + NLP	Close + open (depends on the question and situation)	Only respond to choices	WHO + CDC websites	Rich information with many possibilities but limited user control
Smartex	Aiding food supplies in cases of natural disasters – in development	Natural Disasters (food)	Facebook Messenger, SMS, WhatsApp, and Slack	IBM's Watson, disaster protocols, previous disasters	NLP	Open	Asks questions to identify food shortage	Emergency services	Based on IBM's Watson AI-enabled chatbot. It can identify food shortages according to the information provided by the users and prioritise the work of first responders and emergency authorities
Symptoma	Symptoms check for coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	No	Limited options
Your.MD	Symptoms check for coronavirus and answering frequent questions about coronavirus	Health (Covid-19)	Web	Covid-19 treatment protocols + QnA	Decision trees	Close	Only respond to choices	WHO + CDC websites	Rich information with many possibilities but limited user control

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

U-Report chatbot (UN) (International)	Through communication channels like SMS, Viber, Facebook Messenger, and WhatsApp, users can ask U-Report questions about the Corona Virus and received pre-programmed answers from experts on the matter.	Health (Covid-19)	SMS, Viber, WhatsApp, Facebook Messenger	Pre-programmed answers from experts	NLP + Decision trees	Open	Allows asking questions but provides only the pre-programmed answers	WHO + UN + UNICEF websites	The COVID-19 bot strengthens UNICEF's ability to assess needs, tackle misinformation, and in partnership with governments, share reliable information on where communities can seek assistance
GetJenny – Coronavirus chatbot	Provided answers to questions about the coronavirus	Health (Covid-19)	Web – could be implemented on any website	Information sheets about the corona manually entered	NLP + NLU	Open	Provides information as a response to questions	No	It is limited to manual updates and can miss information. It was subject to manual updates and, therefore, could not respond fast to changes.
Japan Earthquakes chatbot	Allows local Japanese to get basic information in a case of an earthquake	Natural Disasters (Earthquakes)	Mobile App + SMS	Pre-programmed basic recommendations	NLP + Decision trees	Mostly close, but allows free writings	Provides basic information per request	Just reference to emergency phone numbers	Subject to updates by authorities and to basic information guidelines, such as emergency and evacuation centres
POS – Covid-19	FAQ chatbot about coronavirus	Health (Covid-19)	Web	Pre-programmed basic recommendations	Decision trees	Close	Provides information per request	No	Basic visual chatbot with essential information that the user can choose from
Woebot	Mental health during disasters	Mental Health in Disasters	App	Pre-programmed advice, big datasets of therapeutic sessions and CBT programmes	NLP		Unlike many other chatbots, this chatbot can contact the user, ask questions, and provide basic information actively and not just per request	Mental health services, hotlines	While basically, this is a regular mental health chatbot, it is mainly used during disasters to provide mental support to those who need it
SPeCECA	A theoretical suggestion for an emergency chatbot	General first aid disasters	Multiple platforms	Undefined (theoretical)	NLP	Open	Providing information and answering questions	Subject to decision	This is a theoretical suggested model only

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Ask Diana	A chatbot for information about water-related disasters	Water-related disasters	Mobile	Water-related disaster database	NLP + Decision trees	Mostly close, but users could also type in some instances	Providing information and answering questions	No	The chatbot could give basic information about water disasters, such as where to get water, flood alerts and more
SuperWomen	Provides first-aid women-related information in emergencies	Gender (women) related	Facebook messenger	Pre-programmed information sheets with essential information and phones numbers	NLP + decision trees	Open, but allows choosing from limited options	Providing information and answering questions	Emergency services and health centres	A basic messenger chatbot that enables essential basic information for gender-emergency cases
#MeTooMastricht	Collects information about sexual and other attacks	Gender (women) related	Telegram	Pre-programmed messages and commands – only collect information	Decision trees	Open but allows choosing from limited options	Mainly asks for information	Emergency services and support centres	This is a student-developed chatbot that aims mainly at collecting information
Bebot (Japan)	Weather disaster chatbot for foreigners.	Natural disasters – weather	SMS	Basic information sheets	NLP + Decision trees	Open	Respond with essential information to basic requests	No	The chatbot does not provide much information but provides basic answers to basic situations, such as emergency numbers, evacuation centres during earthquakes, etc.
Antidotos_bot	A basic chatbot that provides first aid for poisoning cases	Health (Poisoning)	Telegram	Basic information sheets and a list of chemical material and their antidots	Decision trees	Close	Respond with basic information about the effect of chemical substances, recommendations, and suggested antidotes	Emergency and health services	A basic chatbot with limited options
Water Resources Agency chatbot (China)	A chatbot in Chinese that provides information about water resources during emergencies	Natural Disasters (Water)	Line	Pre-programmed data provided by the government	Decision trees	Close	Provides information per request	No	A basic chatbot with limited options
SA State Emergency Service (Australia)	Allows sending messages to emergency services	General Disasters	Facebook messenger	Pre-programmed messages	Decision trees	Open, but with minimal options	Provides fundamental information	No	The chatbot does not respond very fast to messages and seems like it is not functioning

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Rescue.io	General chatbot for emergencies	General Disasters	Mobile App	Pre-programmed messages	Decision trees	Open, but with minimal options	Provides basic information (numbers). Mainly used to collect information	911 and other emergency services	Used mainly for alerting the emergency services about an emergency that needs 911 assistance
Penn Medicine Triage Tool Algorithm	A chatbot for emergency health issues	Health	Web	Triage protocols	Decision trees	Close	Gives essential medical advice and recommendations for emergencies if needed	No	Although it can advise on emergencies, it does not offer the connection with a live agent, but just a recommendation to go to ER or to call an ambulance
Warning chatbot (EverBridge)	Dissemination channel that allows the users that receive warnings to send back questions	General emergencies	Mobile	Unknown	NLP	Open	Gives specific warnings to the users in a specific location and allows them to ask for relevant information	Emergency services	A location-based chatbot capability that integrates warnings with user's feedback

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

10.2 Appendix B: Reviewing document

Reviewing Solutions

In this document, we analysed existing AI-enabled chatbots, focusing on societal resilience. We included only chatbots that deal with several dimensions of emergencies and disasters in any phase of the emergency cycle. The chatbot analysis focused on the chatbot itself, additional documents and other relevant material identified.

The chatbots were analysed according to the following criteria:

- Name of chatbot
- The operational Organisation/Authority that developed/implements it
- Type of crises (Natural disasters? Health? Fires? Other?)
- Which datasets are used to train the chatbot? (If published)
- The technology of the chatbot:
    - Platform (e.g., Facebook messenger, IBM Watson)
    - Identified algorithms in use (if published, e.g., NLP)
    - Close/open scenarios (does the bot allow free writing or just choosing out of options)
    - Other services in use? (e.g., cloud services)
- The approach of the chatbot to societal resilience:
    - Does the chatbot actively provides information?
    - Does the chatbot only respond to questions?
    - Is the chatbot connected to other services (e.g., emergency services, health centres)?
    - Does the chatbot refer the user to other sources (e.g., see a doctor) or transfer the chat to human-assisted help?
    - Other (open question to add information that is not covered in the other categories)
- What advantages does the chatbot have? What opportunities does it offer? (Open question)
- What disadvantages does the chatbot have? What are the risks of using that chatbot? (Open Question)
- Gaps/weaknesses identified in the chatbot's operation

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

10.3 Appendix C: List of blueprints

Blueprint	**Microsoft Azure conversational bot**	**Google chatbot**	**IBM Reference Diagram**	**Amazon Web Service (AWS)**	**Facebook Messenger**
Component	Bot framework service (BFS)			Genesys Cloud	
Description	This service connects your bot to a communication app such as Cortana, Facebook Messenger, or Slack. It facilitates communication between your bot and the user.			A suite of Genesys cloud services for enterprise-grade communications, collaboration, and contact centre management. Genesys Cloud is the platform where you can access the Lex-Kendra chatbot solution.	
How can it fit emergency chatbots?	Necessary component - allows the communication process between the bot technology and the user interface			The arranging platform of the chatbot	
Component	Azzure app service	Cloud Pub/Sub		AWS	
Description	The bot application logic is hosted in Azure App Service.	A fully managed, real-time publish/subscribe messaging service that sends and receives messages between independent applications is the "glue" that holds the analytic components together. All transcripts (from voice calls or chats) are sent to Cloud Pub/Sub as the first step before analysis.		Amazon Web Services, a cloud computing platform that provides various cloud services such as computing power, database storage, and content delivery. AWS hosts Genesys Cloud.	

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

How can it fit emergency chatbots?					
Component	Language Understanding (LUIS)	Cloud Natural Language	DeepQa	Amazon Lex	PyTorch
Description	As part of Azure Cognitive Services, LUIS enables your bot to understand natural language by identifying user intents and entities.	This service reveals the structure of a text message; you can use it to extract information about people, places, or in this case, to detect the sentiment of a customer conversation.	IBM's NLP capability to analyse conversations and provide answers to questions	A service in AWS that uses machine learning to build conversational interfaces (chatbots).	PyTorch is an open-source deep learning framework built to be flexible and modular for research, with the stability and support needed for production deployment. It enables fast, flexible experimentation through a tape-based autograd system designed for immediate and python-like execution.
How can it fit emergency chatbots?	Allows to understand requests for information before or during emergencies and to classify the requests			to build the conversational scenarios of users with the emergency chatbot	Allows to understand requests for information before or during emergencies and to classify the requests
Component	Azure Search	Dialogflow Enterprise Edition		Amazon Kendra	
Description	Search is a managed service that provides a quick searchable document index.	A tool for building AI-powered conversations across multiple channels. Does not work on NLP algorithms but demands content experts and UX designers to build robust virtual agents for simple scenarios.		An intelligent search service in AWS that is powered by machine learning. The Lex-Kendra chatbot uses Amazon Kendra to search for answers to the customer's questions.	

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

How can it fit emergency chatbots?	Allows to search information in documents of the organisations. For example, in order to find answers to questions, based on the information and organisation holds.	We can work for simple chatbots intended only for simple close scenarios and not based on NLP capabilities.	Allows to search information in documents of the organisations. For example, in order to find answers to questions, based on the information and organisation holds.
Component	QnA maker	Actions on Google (+Google Assistant)	
Description	QnA Maker is a cloud-based API service that creates a conversational, question-and-answer layer over your data. Typically, it is loaded with semi-structured content such as FAQs. Use it to create a knowledge base for answering natural-language questions.	Integrates the bot with the google assistant services to allow basic commands	
How can it fit emergency chatbots?	Gives the base for questions and answers before, during and after emergencies	allows using basic commands that the google assistant is familiar with in the chatbot itself	
Component	Web app	Web integrations	
Description	Allows to integrate the bot in a web platform	Based on angular (front-end) and Node.js with socket.io integration (back-end) to provide a web platform for the chatbot	
How can it fit emergency chatbots?	can provide the service of web hosting of bots for organisations that needs it		
Component	Azure Data Factory	Dialogflow Phone Gateway	AWS CCI

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Description	Data Factory orchestrates and automates data movement and data transformation.	Integration of the chatbot with phone call centres	Contact Centre Intelligence that enables the integration of AI into contact centres. The Lex-Kendra chatbot is part of the AWS CCI self-service accelerator.
How can it fit emergency chatbots?		Can integrate the work of emergency chatbots not just as web/mobile platforms, but also as part of the call centres scenarios	
Component	Logic apps	Cloud Functions	AWS CloudFormation
Description	Logic Apps is a serverless platform for building workflows that integrate applications, data, and services. Logic Apps provides data connectors for many applications, including Office 365	a lightweight compute platform for creating single-purpose, standalone functions that respond to events without the need to manage a server or runtime environment. In this case, the event will be triggered by Cloud Pub/Sub: Every time a message arrives there through the subscriber endpoint, a cloud function will run the message through two Google Cloud services (see below) before storing it in Google BigQuery.	A management tool that uses templates to write, deploy and maintains your AWS infrastructure. An AWS CloudFormation template for the Lex-Kendra chatbot is used to deploy the AWS components of the solution.
How can it fit emergency chatbots?			will be used to deploy the components of AWS in the chatbot
Component	Azure functions	BigQuery	

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Description	You can use Azure Functions to write custom serverless code that is invoked by a trigger	Google Cloud's serverless enterprise data warehouse, supporting super-fast SQL queries enabled by the massive processing power of Google's infrastructure. For example, using BigQuery you could combine your website data with your chat logs.		
How can it fit emergency chatbots?		integrating and analysing data from the chatbot with other sources of information of emergency organisations		
Component	Application insights	Chatbase	Apache UIMA (open source)	
Description	Use Application Insights to log the bot's application metrics for monitoring, diagnostic, and analytical purposes.	In addition to tracking health KPIs, it provides deep insights into user messages and journeys through various reports combined with transcripts. Chatbase also lets you report across different channels/endpoints.	Unstructured Information Management Architecture	
How can it fit emergency chatbots?			Managing and analysing the information that the chatbot needs to manage	
Component	Azure blob storage			Amazon S3
Description	Blob storage is optimised for storing massive amounts of unstructured data.			Simple Storage Service, an object storage service in AWS. Amazon S3 hosts the document repository searched by Amazon Kendra.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

How can it fit emergency chatbots?	Storing the data of emergency chatbots		Integrates with the chat feature. Hosts the documents and data from the databases of the organisations.
Component	Cosmon DB		
Description	Cosmos DB is well-suited for storing semi-structured log data such as conversations.		
How can it fit emergency chatbots?	Can store the conversations with the chatbot for future analysis or to train the chatbot to improve itself		
Component	Power BI		
Description	Use Power BI to create monitoring dashboards for your bot.		
How can it fit emergency chatbots?	Can create dashboards for the conversations with emergency bots, understand its effectiveness, answer the chatters' needs, and more.		
Component	Azure active directory	Cloud Data Loss Prevention	AWS IAM
Description	Users will authenticate through an identity provider	This service discovers and redacts any sensitive information such as addresses and telephone numbers remaining in transcripts before storage.	Identity and Access Management that controls access to AWS resources such as services or features. Use AWS IAM to set permissions to allow and deny AWS resources for the Lex-Kendra chatbot solution.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

How can it fit emergency chatbots?	Help the chatbot to fulfil security and privacy regulations	Secure private data + storing it for necessary use	
Component	Azure key vault		
Description	Store credentials and other secrets using Key Vault.		
How can it fit emergency chatbots?	Help the chatbot to fulfil security and privacy regulations		
Component	Azure DevOps		AWS Lambda
Description	Provides many services for app management, including source control, building, testing, deployment, and project tracking.		Serverless computing service for running code without creating or maintaining the underlying infrastructure. AWS Lambda executes Amazon Kendra fulfilment and other operations for the solution.
How can it fit emergency chatbots?			
Component	VS code		
Description	A lightweight code editor for app development. You can use any other IDE with similar features.		
How can it fit emergency chatbots?			
Component	Structured Data		
Description	CRM, SQL, Tables		
How can it fit emergency chatbots?	All the data that emergency organisations hold is organised in a structured way. For example, tables of cases that are distributed under specific columns, rows and cells		
Component	Unstructured Data		

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

Description	FAQs, PDFs, Word documents
How can it fit emergency chatbots?	Data which is stored in an unstructured way, such as transcriptions of phone calls

Component	
Description	
How can it fit emergency chatbots?	

Component	Cloud Build (deployment)
Description	Build and deploy these containers to Google Kubernetes Engine with a single command in minutes.
How can it fit emergency chatbots?	

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

10.4 Appendix D: AI-enabled chatbot implementation stages

Step #	Questions to ask	Actions to do
#1 Framework	• What do we want the chatbot to be able to do? What do we not want it to do? What is possible? • Can we self-develop the chatbot or do we need an external developer? • How much budget do we have for developing the chatbot? What can be done with this budget? What cannot be done?	• Defining to scope of the chatbot. • Defining goals and objectives. • Allocating funds and calculating the budget. • Deciding whether the chatbot will be self-developed or by an external company.
#2 Channels	• How do we communication with the public? • Are there any other communication channels that the public, or segmented parts of the public, is using? • What other communication channels can be used by the organisation and integrated with the chatbot?	• Mapping the current communication channels of the organization (online and offline). • Expaning the communication options to other channels that the public is using, but are not yet accessible by the organization. • Examining which channels can be integrated with the chatbots and which can not.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

#3 The Brain

- What are the technological demands from the chatbot we wish to create?
- How do we want to communicate with our users, using the chatbot?
- What services are available to create the "brain" of the AI-enabled chatbot which better fits to our goals and objectives?

- Choosing the platform which will be used to build the body of the chatbot, which best answers our needs.

#4 Security

- What types of data is going to be stored in the chatbot's system?
- What types of security measures do we need for the data?
- What security and privacy measures are needed according to the regulations that the orgnisation is following?

- Choosing security providers.
- Creating authentication processes.
- Setting the different groups of users and their levels of access (e.g., system managers, DevOps, content creators, regular users).

#5 Datasets

- What are the datasets that the organization is already creating?
- How are they organised?
- How can the datasets facilitate different parts and processes in the work of the AI-enabled chatbot?
- Are there any other datasets that can be relevant? How can we start creating them?

- Mapping the datasets the organisation have and their types (i.e., structured versus unstructured).
- Deciding which datasets are relevant for which activities of the chatbot, and in what ways.
- Identifying new types of datasets that can be created in order to help the AI-enabled chatbot to function better.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

#6 Capabilities

- What type of input (e.g., text, image, video) is the chatbot going to process?
- What algorithms can facilitate processing the data in the datasets, the communication methods with the users and the goals of the AI-enabled chatbots?
- What technological capabilities do we need to facilitate our needs?

- Defining the ML algorithms and learning approaches which are relevant to the facilitation of our AI-enabled chatbot.
- Choosing the services and technological capabilities that can facilitate our chatbot's goals and objectives.

#7 Monitoring

- What activity of the chatbot do we need to log?
- For what purposes are we going to use the logged data (e.g., statistics, quality assurance, identifying stress situations)?
- What services are relevant to facilitate our monitoring needs?

- Choosing the services and providers for logging the information.
- Defining monitoring processes.
- Defining monitoring activities and results (e.g., QA, alerts, dashboards).
- Choosing the relevant services for facilitating the monitoring needs.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°

#8 Processes

- What are the relationships between the different components of the chatbot (e.g., is the relationship between user queries and the dataset is unidirectional or multidirectional, meaning that the queries can affect the datasets)?
- Where are there components of feedback?
- How does the chatbot learn and improve?

- Defining which components in the chatbot's system are interconnected.
- Deciding which type of feedback is crucial for the system and how to process it (e.g., only to notify about suggested changes, changing things online).

#9 Build

- Do we have answers to all the questions and needs?
- Can we draw a framework that illustrate how the chatbot works?

- Building the AI-enabled chatbots, programming all the components, according to the previous steps.

#10 Implement

- Did we finish building all the components of the chatbot to meet our needs?

- Training the AI-enabled chatbot and testing it until it reaches a satisfactory level.
- Monitoring the AI-enabled chatbot for another period of time, while it is live and answers the public's question, to track mistakes and apply improvements.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020/2014-2020) under grant agreement n°