beAWARE

beAWARE

Enhancing decision support and management services in extreme weather climate events

700475

D6.2

Data-Source Integration Framework

Dissemination level:	Public
Contractual date of delivery:	Month 15, 31 March 2018
Actual date of delivery:	Month 15, 31 March 2018
Workpackage:	WP6: Public Safety Answering Point
Task:	T6.2 Data Source Integration Framework
Type:	Report
Approval Status:	Final
Version:	V1.0
Number of pages:	49
Filename:	D6.2_beAWARE_Data-Source Integration Framework_2018-03-31_v1.0.docx

Abstract

This deliverable reflects the work performed in task T6.2 dedicated to ensure robust and reliable integration of data sources. In particular, the T6.2 objectives are to develop a suitable support framework to create an integration layer that can be deployed on top of varying physical infrastructures to expedite the integration of existing and new data sources. By providing this abstraction from the underlying data sources, it will provide a framework to intelligently manage the data and control flows between the beAWARE platform and the physical environment. To support the integration of existing sensors, camera, video and other structured data, a set of reusable data source connectors will be developed. This deliverable will describe the Data Source Integration Framework layer.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 700475

Co-funded by the European Union

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 700475

History

Version	Date	Reason	Prepared by
0.1	December 1, 2017	Deliverable outline	Y. Mordecai (MSIL)
0.2	January 15, 2018	Deliverable draft	Y. Mordecai (MSIL)
0.3	February 15, 2018	Partner Contribution Integration – Cycle 1	Y. Mordecai (MSIL)
0.4	March 15, 2018	Partner Contribution Integration – Cycle 2	Y. Mordecai (MSIL)
0.5	March 19, 2018	Pre-Submission Review	B. Kantsepolsky (MSIL)
0.6	March 28, 2018	Peer-Reviewed Version	Y. Mordecai (MSIL)
1.0	March 29, 2018	Submitted Version	Y. Mordecai (MSIL)

Author list

Organization	Name	Contact Information
MSIL	Yaniv Mordecai	yaniv.mordecai@motorolasolutions.com
MSIL	Boris Kantsepolsky	boris.kantsepolsky@motorolasolutions.com
MSIL	Michal Shany	michal.shany@motorolasolutions.com
MSIL	Israel Altar	israel.altar@motorolasolutions.com
IBM	Benny Mandler	mandler@il.ibm.com
CERTH	Anastasios Karakostas	akarakos@iti.gr
CERTH	Stelios Andreadis	andreadisst@iti.gr
CERTH	Emmanouil Michail	michem@iti.gr
IOSB	Jürgen Moßgraber	juergen.mossgraber@iosb.fraunhofer.de
IOSB	Philipp Hertweck	philipp.hertweck@iosb.fraunhofer.de
IOSB	Hylke van der Schaaf	hylke.vanderschaaf@iosb.fraunhofer.de
IOSB	Tobias Hellmund	tobias.hellmund@iosb.fraunhofer.de
UPF	Stamatia Dasiopoulou	stamatia.dasiopoulou@upf.edu
UPF	Jens Grivolla	jens.grivolla@upf.edu
AAWA	Michele Ferri	michele.ferri@adbve.it
AAWA	Daniele Norbiato	daniele.norbiato@adbve.it
AAWA	Francesca Lombardo	Francesca.lombardo@adbve.it
AAWA	Giovanni Tomei	Giovanni.tomei@adbve.it

Reviewers

Partner	Name	Contact Information
MSIL	Boris Kantsepolsky	boris.kantsepolsky@motorolasolutions.com
CERTH	Anastasios Karakostas	akarakos@iti.gr
CERTH	Gerasimos Antzoulatos	gantzoulatos@iti.gr
HRT	Iosif Vourvachis	iosif_vour@yahoo.gr

Executive Summary

beAWARE is a highly-interconnected system of systems and interdependent modules that feed each other with information regarding evolving and ongoing extreme climate emergencies. This includes a control center solution, a mobile application, various media and social media analysis services, sensor analysis service, weather and climate forecast services, crisis classification and reasoning services, multilanguage analysis and generation services, and supporting infrastructure.

beAWARE integrates data from multiple sources, including field reports, sensor measurements, weather data, media, analysis results, and social media posts. In addition, the system has to generate information and insight as well as actions and instructions within the system and for its users and stakeholders. For this reason, a suitable data source integration framework has been defined to govern the data and information exchange architecture and realization.

The purpose of the data-source integration framework is therefore to assure and support interoperability among various stakeholders and operational actors – authority officials, emergency management officers, analysts, first responders, and citizens. This is done by a robust, dynamic, and flexible infrastructure that allows interconnectivity among the various subsystems, modules, and services. On top of this infrastructure, the system enables and implements multiple end-to-end interactions, leveraging data exchange, storage, access, acquisition, and visualization mechanisms.

This document is the main reference for the data exchange architecture of the beAWARE platform as implemented toward beAWARE's first prototype, due June 2018. The framework has been defined to be sufficiently robust to accommodate possible updates and extensions of data exchange and collaboration, as the project progresses towards the final version, due December 2019.

This document provides: a) a **public** overview of the operational user requirements entailing impact related to data and data-source management and integration; b) a **public** detailed description of the common data integration infrastructure, including messaging, data access and querying, and file storage services; and c) a **confidential**, detailed specification of each module's connectivity requirements, inputs, outputs, and data access mechanisms.

The operational section is public because it relies on publicly-available operational user requirements, which constitute common knowledge of public interest. The operational requirements have been analyzed for impact on data-source and data user integration,

interaction, and interoperability. The main modules and entities were laid out to support these requirements. This process and its outcomes are of use as a general reference for the community.

The section discussing the generic infrastructure supporting data exchange and integration, including messaging, media storage, databases, etc., is also public since it relies on publicly-available open-source products, and constitutes a benchmark that we advocate to the emergency management technology community.

In the applicative layer, the beAWARE modules' utilization of the generic infrastructure accounts for the various constraints, challenges, technical requirements, and data content. This information remains confidential, in order to prevent direct exposure of technical details, data structures, and access control mechanisms, and in order to protect the beAWARE partners' intellectual property.

This deliverable is the product of a year-long effort to define, organize, manage, design, implement, and demonstrate the data-source integration framework, as part of task T6.2 in WP6, under the responsibility of MSIL. This effort was heavily supported by IBM in the generic infrastructure aspects, and involved all the other technical partners – AAWA, CERTH, FMI, IOSB, and UPF. Each technical partner helped derive and define its technical connectivity and functional data-exchange requirements. The consortium as a whole worked together to ensure end-to-end flow of information through a series of teleconferences, face-to-face meetings, partner-to-partner discussions and end-to-end integration sessions. The operational partners – HRT, PLV, and FBBR – contributed to this effort by reviewing the system requirements and overall architecture throughout the project, and ensuring their compliance with the operational requirements. In addition, PLV and HRT provided useful comments and reviews on this deliverable prior to its submission.

Abbreviations and Acronyms

alt	altitude
API	Application Programming Interface
ASR	Automatic Speech Recognition
CAP	Common Alert Protocol
CDR	Central Data Repository
CRCL	Crisis Classification
CSV	Comma-Separated Values
DB	Database
GIS	Geographic Information System
GPS	Global Positioning Satellite
HTM(L)	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
ID	identifier
IoT	Internet of Things
IP	Internet Protocol
JSON	JavaScript Object Notation
KBS	Knowledge-Base Services
lat	latitude
long	longitude
MRG	Multilanguage Report Generation
MSB	Message Bus
MTA	Multilanguage Text Analysis
ODI	Open Data Interface
PCI	Peripheral Component Interconnect
PSAP	Public Safety Answering Point
Pub/Sub	Publish—Subscribe
REST	Representational state transfer
SDK	Software Development Kit
SMA	Social Media Analytics
SoS	System of Systems
SQL	Structured Query Language
TCP	Transmission Control Protocol
TR	Technical Requirement
UDP	User Datagram Protocol
UI	User Interface
UR	User Requirement
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
USB	Universal Serial Bus
WFS	Weather Forecast Services
WP	Work Package
XML	Extensible Markup Language
XSD	XML Schema Definition

Partner Acronyms

AAWA	AUTORITÀ DI BACINO DEI FIUMI ISONZO TAGLIAMENTO LIVENZA PIAVE BRENTA BACCHIGLIONE
CERTH	CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS
FBBR	FREDERIKSSUND-HALSNÆS: FIRE- & RESCUE SERVICE
FMI	FINNISH METEOROLOGICAL INSTITUTE
HRT	HELLENIC RESCUE TEAM
IBM	IBM ISRAEL - SCIENCE & TECHNOLOGY LTD.
IOSB	FRAUNHOFER-GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V.
MSIL	MOTOROLA SOLUTIONS ISRAEL LTD
PLV	AYUNTAMIENTO DE VALENCIA
UPF	UNIVERSITAT POMPEU FABRA

Table of Contents

List of Figures

List of Tables

1 Introduction [PUBLIC]

1.1 Scope

This document constitutes a reference for the data integration and exchange framework for beAWARE.

1.2 "beAWARE"

The beAWARE Project is an EU-funded collaboration (#700475) of partners from several countries in Europe to deliver a prototype disaster management system for extreme weather conditions. The Project is focused on Flood, Forest Fire, and Heatwave scenarios, and is intended for deployment and testing of these scenarios in Vicenza (Italy), Valencia (Spain), and Thessaloniki (Greece), respectively. The end users are the Alto Adriatico Water Authority (AAWA), Valencia Local Police (PLV), and Hellenic Rescue Team (HRT), respectively. In addition, the Frederiksborg Fire Department (Frederiksborg Brand OG Redning, FBBR) contributed to the operational requirements for the Fire scenario.

The technical partners involved in the project include: Centre for Research and Technology – Hellas (CERTH) who is also the coordinator of the project; Motorola Solutions Israel (MSIL), who is also the technical manager of the project; IBM Israel Haifa Research Labs (IBM); Finnish Meteorological Institute (FMI); Fraunhofer Institute for Optronics, System Technologies and Image Exploitation (IOSB); and Universitat Pompeu Fabra (UPF).

The beAWARE system is an end-to-end solution for collecting information from multiple data sources – such as end users, social networks, sensors, and data providers – analyzing it, predicting and assessing emergencies, alerting the public, and managing first responders' activities.

1.3 The need for a Data-Source Integration Framework

The multitude of data originators and consumers in flexible configurations within the BeAWARE platform mandates a robust framework for data connectivity, integration, processing, and exchange among the involved parties, modules, and services.

The data-source integration framework includes three layers, following the outline of the Model-Based Interoperability Engineering (MoBIE) framework [1]:

- An Interoperability Layer, in which the multiple operational stakeholders can collaborate, coordinate, construct and transfer knowledge, and exchange useful information in order to carry out the missions incurred by natural emergency and disaster.
- An Interconnectivity Layer, in which the services and systems supporting the operational stakeholders can jointly process and streamline content, information, and data, in order to support operational procedures and business processes enabled by these systems.
- An Intercommunication Layer, in which the multiple input-output interfaces of the constituent systems can send and receive standardized data structures as part of the informational transactions and services provided by the constituent systems.

The partners engaged in the development of the individual subsystems will address the development of appropriate mechanisms in their respective working area to comply with this data integration framework.

Similarly, the end users will further build the operational transactions and protocols based on the common language and guidelines described and specified in this document and as part of ongoing work.

The high potential for dynamics in the data exchange among the parties calls for solid and yet robust guidelines and formalisms to ensure the flexibility and extendibility of the platform to accommodate the expected dynamics.

This report is aligned and coordinated with D7.2 – System Requirements and Architecture, and adheres to the BeAWARE principle of unified “publish—subscribe” messaging bus. Still, this report does not focus on the technicalities of data messaging but on the overall principles and conventions that ensure appropriate and constructive utilization of assets such as the message bus and message protocol by the technical partners and their provided systems.

1.4 Outline

This document is structured as follows:

- Section 2 [Public] discusses the requirements for a data-source integration framework.

- Section 3 [Public] discusses the data source integration infrastructure, which enables the interoperability among the beAWARE modules and subsystems.
- Section 4 [Confidential] discusses the module-level data access mechanisms, receivable and producible data artifacts applicable to each module, and the interactions with other modules.
- Section 5 [Public] concludes this report.

2 Data-Source Integration Requirements [PUBLIC]

2.1 Scope

In this section we review and analyze the requirements for data integration and interaction among the beAWARE modules. In addition, we derive data integration and interaction needs and functionalities at the beAWARE-level and at the module-level.

2.2 User Requirements for Data-Source Integration

Stakeholders of the operational scenarios that beAWARE is required to respond to have defined various operational and functional requirements for the system, in order to support the roles, responsibilities, and activities of human agents during the occurrence of the scenario. The complete list of initial User Requirements is defined in Deliverable D2.1 [2], which was published in M6.

Several user requirements concern the exchange of data and information among the authorities, decision makers, control center operators, crisis analysts, first responders, the public, and automated analytic services.

For the purpose of this report, the initial user requirements provided by beAWARE partners are considered as a general reference for an overall understanding of user needs, expectations, intentions, and constraints with respect to the data source integration framework. The possible impact of each user requirement on data integration and information-based interaction has been analyzed throughout the course of the project and the architecting of the framework. The results of the multiple teleconferences and emails exchange discussions are summarized in the following sections. The impact is specified in the right column in each of the user requirements tables below.

User requirements whose impact resembles that of previously analyzed user requirements refer to the original user requirements from which the impact was derived (for instance, UR#202 points the reader to UR#128 for impact). The goal was to minimize multiple definitions of the same impact, and to consolidate use cases as derived from various user requirements. For example, providing a suitable mechanism for transferring public alerts from the control center to the citizens is the same for all the user requirements that refer to generating public alerts, regardless of the type of hazard they intend to warn about, the specificity of the applicable location or area, and the content of the alert.

We emphasize that the impacts derived from the user requirements are still oriented as much as possible to the problem domain, and are mostly interoperability-driven, so that they reflect operational needs for connectivity and data exchange as means to support interactions among operational users, e.g. to allow information sharing between the authority and the public, or the first responders and the control center. Thus, these impacts are still not allocated to specific modules or data access mechanisms, and serve together as the basis for the definition of this architecture.

2.2.1 User Requirements for the Flood Scenario

In this section we summarize the analysis of data source integration and connectivity impacts for the user requirements defined for the Flood scenario (SCN#1). As explained, the impacts are specified on the rightmost column for each user requirement, and requirements with an impact similar to the one already defined for previously analyzed requirements refer to the user requirement which has the same impact. Since this is the first scenarios' user requirements that were analyzed, most of the user requirements have a distinct impact in the connectivity and data exchange aspects. A negligent portion of the requirements refer to module-intrinsic capabilities that do not impact the overall connectivity or require special data exchange mechanisms. We have kept these requirements for reference. A few requirements are sufficiently similar to be covered by a similar connectivity or integration feature, such as the public alert or metric visualization.

Table 2-1. Initial User Requirements – Flood Scenario [2]

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_101	Type of visualization	Display information to authorities in a web-gis platform (citizen and first responders reports by calls, apps, social media)	Transfer position data for first responders, and social media reports
UR_102	Map of the AMICO Flood EWS results	Display reliable and trustful flood forecasts, potential dangerous situations and the forecasted level of risk to the authorities, based on the results of the EarlyWarning System AMICO (improved with the assimilation of Satellite data (snow cover, soil moisture, etc.) and Meteorological forecasts data with a finer spatial resolution provided by FMI)	Transfer flood forecast metrics and flood risk assessment metrics to authorities

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_103	Flood warnings	Provide authorities/citizens with automatic warnings on river levels overtopping some predefined alert thresholds, based on forecast results	Transfer public alerts from authorities to citizens.
UR_104	Send/receive emergency reports	Allow citizens to send text, images, audio and video messages from their mobile phones (for the different operative systems) and from their social media account to the authority during bad weather conditions when the GPS signal is low.	Transfer multimedia-enriched incident reports from citizens' mobile devices to authorities
UR_105	Send task reports	Allow First Responders to send reports about their assignments from their mobile phone to local authorities	Transfer multimedia-enriched incident reports from first responders' mobile devices to authorities
UR_106	Visualize video cameras	Display streamed video from video cameras to the authorities/citizens	Transfer video streams from cameras to authorities
UR_107	Localize video, audio and images	Provide authorities with the ability to localize videos, audio and images sent by citizens from their mobile phones	Transfer citizen report position data to authorities
UR_108	Localize task status	Provide authorities with the ability to localize first responders reports regarding the status of their assigned tasks	Transfer team status and position report to authorities
UR_109	Localize tweets	Provide authorities with the ability to localize Twitter messages concerning a flood event	See UR_101
UR_110	Localize calls	Provide authorities with the ability to localize Phone Calls to an emergency number concerning a flood event	See UR_105
UR_111	Detect flooded elements from video	Provide authorities with the ability to detect and count flooded elements (e.g. cars and people inside the river)from video and images sent from mobile phones and social media	Transfer video/image analysis reports and metrics to authorities
UR_112	Detect elements at risk from reports	Provide authorities with the ability to detect the number of elements at risk and the degree of emergency from text sent by the mobile app or social media	Transfer social media analysis reports and metrics to authorities

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_113	Detect element at risk from calls	Provide authorities with the ability to detect the number of element at risk and the degree of emergency from emergency calls	Transfer audio call analysis reports and metrics to authorities
UR_114	Detect water depth and velocity	Provide authorities with the ability to detect water depth and water velocity from video and images sent by the mobile app and social media	See UR_111
UR_115	Real time flood mapping	Display flooded areas in real time to authorities/citizens	Transfer area position data (polygons) to authorities
UR_116	Warning people approaching flood areas	Provide authorities with the ability to warn people in danger with warning messages, when they are approaching a flooded area	See UR_103
UR_117	Manage assignments in case of new emergencies	Provide authorities with the ability to manage first responder assignments	Transfer tasks from authorities to first responder teams
UR_118	River overtopping	Provide authorities/citizens with the ability to know if the river level is overtopping predefined alert thresholds	Transfer sensor analysis reports and metrics to authorities
UR_119	Manage assignments based on river level overtopping	Provide authorities the ability to assign task to first responder teams related to the overtopping of predefined river level thresholds	See UR_117
UR_120	Map of rescue teams and task evaluation	Display to authorities the location in time of first responder teams in all the municipality and provide the ability to evaluate in real time the execution of the assigned tasks with a global visualization of the activities performed	See UR_108
UR_121	Detect rainfall volume and duration	Provide authorities with the ability to detect rainfall volume and duration from videos (fixed and mobile cameras, social media and the mobile app)	See UR_111

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_122	Rainfall warnings	Provide authorities/citizens with the ability to know in real time if the rainfall intensity is overtopping predefined alert thresholds	See UR_118
UR_123	Detect embankment exceeding	Provide authorities with the ability to detect from video, automatically (fixed and mobile cameras, social media and mobile app) , if a river embankment is overtopping and/o breaking	See UR_111
UR_124	Embankment warnings	Provide authorities/citizens with the ability to know in real time if a river embankment is overtopping and/or breaking; the comprehensive and reliable real-time information about the situation, especially the breach enlargement and discharge, the spatial and temporal development of the inundation and the damages	See UR_118
UR_125	Traffic warnings	Provide authorities with the ability to send warnings to citizens in order to avoid interferences inside the area involved by civil protection activities	See UR_103
UR_126	Map of Satellite data and weather forecasts	Display updated satellite images and weather forecasts.	Transfer satellite images and weather forecasts to authorities
UR_127	Filters	Provide advanced filters in the data management platform (visualize and list information selected by filters/query)	N/A
UR_128	Evaluation of the level of risk	Provide authorities with the ability to evaluate the forecasted level of risks (based on all the available dataset)	Transfer forecast and analysis metrics to authorities
UR_129	Automatic translation from a foreigner applicant	Make the communication between people with different languages easier	Transfer text as needed to automatic translation services
UR_130	Traffic Status	Display to the authorities the current traffic situation so that they can decide where to direct the first responders or inform them of which routes to avoid	Transfer traffic data to authorities

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_131	Traffic warnings	Provide authorities with the ability to send warnings to citizens in order to avoid a certain area that is jammed with traffic	See UR_103

2.2.2 User Requirements for the Fire Scenario

In this section we summarize the analysis of data source integration and connectivity impacts for the user requirements defined for the Fire scenario (SCN#2). Again, the impacts or pointers to previously analyzed requirements with a similar impact (including user requirements for SCN#1) are specified on the rightmost column for each user requirement. Since this is the second scenario to be analyzed, and since this scenario is a bit more restricted in scope, most of the user requirements are already covered by previously defined connectivity or integration features.

Table 2-2. Initial User Requirements – Fire Scenario [2]

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_201	Detection of people and goods in danger	Display information authorities/first responders to detect people, cars and buildings in danger.	See UR_128
UR_202	Detection of critical aspects	Provide authorities/first responders with information in order to detect the following kind of situation, process, material or condition that can cause a wildfire or intensify its damaging impacts: drought, air temp. and weather aspects, fuel accumulation spots, crowds, etc.	See UR_128
UR_203	Study of the smoke behaviour	Provide authorities/first responders with information on the smoke behavior (vertical/inclined, column, smoke color…).	See UR_111
UR_204	Identification of the fuel being burned	Provide information to authorities/first responders to know the type of fuel being burned by the colour and the shape of the smoke	See UR_111
UR_205	Analysis of advancing fire	Provide authorities/first responders with an analysis of the advancing fire (flame progression, height and length).	See UR_115

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_206	Specific weather data	Provide authorities/first responders and citizens with specific weather data of the Devesa place, as it has a specific microclimate that might be different from other places.	See UR_126
UR_207	Aerial images	Display authorities/first responders to visualize aerial images of the smoke and the trajectory flames. It will provide information about the extension and the damages (kind of damages, and so on), the tracking of the fire, vehicles and people around the spot, in order to find out possible suspects or victims. Furthermore, if these aerial images provide thermal information it can be used for looking over the fire perimeter once it has been extinguished, in order to locate sleeper fire and avoid possible reproduction.	See UR_111, UR_104
UR_208	Access to road traffic cameras	Allow authorities/first responders to have access to the cameras located at CV-500 (La Devesa main road) and CV-5010. Although there are already installed cameras, they are managed and visualized by autonomic resources.	See UR_106
UR_209	Electronic traffic panels	Display authorities/first responders to display in electronic traffic panels useful information and evacuation instructions in case.	See UR_103
UR_210	Mobile application	Enable citizens to communicate a fire alert, detected neglects or other risk situations and even send visual data through a mobile application.	See UR_104
UR_211	Location of vehicles and personnel involved	Display authorities/first responders to visualize GPS location and/or real time footage of vehicles and personnel on the incident site. Transmitted to an online map where the coordination centres can follow both the development of the incident and the location and amount of resources. The online map will also provide the possibility of interacting with the police and other agencies involved	See UR_108
UR_212	Traffic warnings	Sending warnings to citizens in order to avoid interferences inside the area.	See UR_103

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_213	Recommendations	Sending recommendations to citizens.	See UR_103
UR_214	Warnings	Sending warnings of pre-emergency alerts to citizens by authorities	See UR_103
UR_215	Evacuation orders	Ordering evacuations of citizens at risk.	See UR_103
UR_216	Internal sharing of information	Sharing data (images, videos, geolocation, reports) regarding the forest fire among authorities & first responders	Transfer multimedia-enriched incident updates from authorities to first responders
UR_217	Twitter analysis and warning	Warning authorities/first responders about Twitter messages concerning the forest fire event.	See UR_112
UR_218	Automatic detection system	Having an automatic detection system of the forest fire, which is connected to firefighters and police officers	See UR_128
UR_219	Coordination and communication between different resources	Provide communication between authorities and first responders, in order to improve their coordination.	See UR_216, UR_108, UR_117
UR_220	Improvement of the signal for telephones and emergency communication	Provide authorities/first responders with an accurate coverage of telephone mobile lines and emergency communication due to there is currently a lack of signal in some spots of the area.	See UR_115
UR_221	Geolocalitation of telephone calls	To geolocalize a mobile phone citizen call by sending a request permission message to the citizen, who would accept to be tracked temporarily.	See UR_104
UR_222	Filter of the emergency messages	Transfer emergency calls by writing (only minor emergencies or only information call). The aim is to save operator time and avoid losing emergency calls	See UR_104
UR_223	Automatic selection of the level of emergency	This can be done with the operator's supervision. The aim is to save time and avoid losing emergency calls	See UR_128

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_224	Automatic translation from a foreigner applicant	Make the communication between people with different languages easier	See UR_129
UR_225	Quick search of events and applicants	Data storage, in order to improve indexation of information relative to events and applicants	Transfer search requests and results between authorities and data sources

2.2.3 User Requirements for the Heatwave Scenario

In this section we summarize the analysis of data source integration and connectivity impacts for the user requirements defined for the Heatwave scenario (SCN#3). Similarly, the impacts or pointers to previously analyzed requirements with a similar impact (including user requirements for SCN#1) are specified on the rightmost column for each user requirement. Since this is the third scenario to be analyzed, most of the user requirements are already covered by previously defined connectivity or integration features, most of which were identified for SCN#1. A small portion of the requirements that were not identified as having connectivity or data exchange impacts have been marked as “N/A”.

Table 2-3. Initial User Requirements – Heatwave Scenario [2]

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_301	Real time weather forecast	Provide the authorities with real time weather forecast in relation to the progression of the heatwave phenomenon	See UR_126
UR_302	Automatic warning	beAWARE system to generate and provide the authorities with an automatic warning when an imminent heatwave phenomenon is forecasted	See UR_128

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_303	Risk assessment for a forest fire	Provide the authorities with a risk assessment regarding the probability of a forest fire to occur during or in the upcoming period after a heatwave. The relevant authorities will have an assessment of a fire risk based on the weather forecast during a heatwave and especially during the following days	See UR_128
UR_304	Heatwave intensity	Provide the authorities with a risk assessment regarding the intensity of the phenomenon in the city.	See UR_128
UR_305	Possible locations for incidents	Display to the authorities visual information about possible locations in the city (or outside the city) where a situation that will require rescue team intervention is more likely to develop (for example, based on past experience, traffic jam and/or accidents will be more likely to occur at a main street intersection/ public park/ entrance to hospitals or banks… etc.). In such cases a decision might be made to send rescue teams in advance to shorten response time if/when an incident occurs	Transfer predicted incidents to authorities
UR_306	Number of people affected	Provide the authorities with an estimation of the people that might be affected from the phenomenon and in which areas	See UR_118
UR_307	Power needs	Provide the authorities with an estimation on the power needs during a heatwave based on its foreseen progression	See UR_118
UR_308	Infrastructure overload	Provide the authorities with an estimation of damage/overload to the city's infrastructure (phone lines, electricity, water, etc)	Transfer infrastructure data to authorities
UR_309	False Alarms	Provide to the authorities with a procedure to confirm necessity of rescue teams so they are not sent needlessly to one place instead of somewhere else where they are needed more urgently, therefore the ability to handle false alarms.	N/A

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_310	City-wide overview of the event	Provide the authorities with a city-wide overview of the event – allow decision making authorities an overall view of all incidents handled at any point in time/ see where all rescue teams are located in real-time to allow them to make informed decisions regarding who to send where… etc	N/A
UR_311	Information Storage	Provide the authorities with access to all historical information by providing storage for all information for future lessons-learned purposes, so that after the heat wave situation is over, decision making authorities can review the information gathered and handled during the event, and set-up better procedures to handle future events more efficiently	Transfer data to a central archive
UR_312	Warning citizens	Provide citizens with warnings through the beAWARE app, of an imminent heatwave and a list of proactive measures and how to reduce its effects	See UR_103
UR_313	First responders status	Provide to the authorities the current status and location of all first responders when they are performing their tasks	See UR_108
UR_314	Assign tasks to first responders	Allow authorities to assign additional tasks to those first responders who are available or even instruct those who are able to assist other responders	See UR_119
UR_315	Traffic Status	Display to the authorities to monitor the current traffic situation so that they can decide where to direct the first responders or inform them which routes to avoid	See UR_130
UR_316	Capacity of relief places	Show the authorities the current state of the available capacity of all relief places provided to the public	Transfer asset data to authorities
UR_317	Areas with power outage	Display to the authorities the areas where there is a power outage.	Transfer power status to authorities
UR_318	Trapped citizens	Allow authorities to know if there are people trapped (e.g. in an elevator) and display where	See UR_118
UR_319	Trapped elders at home	Allow authorities to know if there are elder people trapped in houses without an A/C and display where	See UR_118

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_320	Hospital availability	Show the authorities the current availability of the hospitals.	See UR_316
UR_321	Affected area	Provide the authorities with a prediction of the affected area	See UR_115
UR_322	Information for incident status from Social Media	Provide information to the authorities regarding potential risks in case there is a situation in the city (eg car accident, etc.) gathered from social media	See UR_101
UR_323	Information for Hospital Status from Social Media	Provide information to the authorities regarding overcrowded hospitals and places offered to the public with a/c, gathered from social media	See UR_101
UR_324	Information for existing situation in the from Social Media	Provide information to the authorities regarding existing traffic conditions all over the city grid gathered from social media	See UR_130
UR_325	Suggested places for relief	Provide information to citizens regarding the suggested places for relief through an app.	See UR_103
UR_326	Type of visualization	Display information to the authorities/ citizens all the in a web-gis platform	N/A
UR_327	Send emergency reports	Allow citizens to send text, images and video messages from their mobile phone (for the different operative systems) and from their social media account to the authority.	See UR_101
UR_328	Send task reports	Allow First Responders to send reports about their assignments from their mobile phone to local authority	See UR_105
UR_329	Visualize video cameras	Display streamed video from video cameras to the authorities/citizens	See UR_106
UR_330	Localize video and images	Provide authorities with the ability to localize videos and images sent by citizens from their mobile phones	See UR_107
UR_331	Localize task status	Provide authorities with the ability to detect the location of first responders	See UR_108
UR_332	Localize tweets	Provide authorities with the ability to localize Twitter messages	See UR_109
UR_333	Localize calls	Provide authorities with the ability to localize Phone Calls to an emergency number concerning citizens who are trapped	See UR_110

UR#	Requirement name	Requirement description	Data Sources and Data Integration Impact
UR_334	Manage assignments in case of new emergencies	Provide authorities with the ability to manage first responder assignments	See UR_117
UR_335	Map of rescue teams and task evaluation	Display to authorities the movements of first responder teams in all the municipality and provide the ability to evaluate in real time the execution of the assigned tasks with a global visualization of the activities performed	See UR_120
UR_336	Traffic warnings	Provide authorities with the ability to send warnings to citizens in order to avoid a certain area that is jammed with traffic	See UR_103
UR_337	Location of vehicles and personnel involved	Allow authorities/first responders to visualize GPS location and/or real time footage of vehicles and personnel on the incident site. Transmitted to an online map where the coordination centres can follow both the development of the incident, and the location and amount of resources. The online map will also provide the possibility of interacting with the police and other agencies involved	See UR_105
UR_338	Warnings	Allow authorities to send warnings of pre-emergency alerts to citizens.	See UR_103
UR_339	Evacuation orders	Allow authorities to order evacuations of citizens at risk.	See UR_103
UR_340	Internal sharing of information	Allow authorities and first responders to share data (images, videos, geolocation, reports)	See UR_117
UR_341	Twitter analysis and warning	Allow authorities/first responders to be warned by Twitter messages concerning traffic jam, availability of places of relief, potential hazards or people in danger	See UR_101
UR_342	Coordination and communication between different resources	Provide communication between authorities and first responders, in order to improve their coordination.	See UR_105

2.3 Operational Roles

Based on the description of the scenarios and user requirements, a set of operational user roles has been defined, in order to cover all the participants of the actual scenarios. The operational roles have been defined and listed in Table 2-4.

Table 2-4. Operational Roles

ROLE	LOCATION	INTERFACE	RESPONSIBILITIES	SCENARIOS
Emergency Manager (H)	Command & Control Center	PSAP	Oversee the emergency situation	Flood, Fire, Heatwave
Incident Manager (H)	Command & Control Center	PSAP	Manage and prioritize incidents	Flood, Fire, Heatwave
Operations Manager (H)	Command & Control Center	PSAP	Assign tasks to teams and monitor their progress	Flood, Fire, Heatwave
Analyst (H)	Command & Control Center	AMICO	Oversee Flood Forecasting	Flood
First Responder (H)	Field	FRAPP	Report on new and existing incidents, execute assigned tasks, report task progress, receive operational alerts	Flood, Fire, Heatwave
Citizen (H)	Field	SCAPP	Report on new and existing incidents, Receive public alerts	Flood, Fire, Heatwave
Sensor (M)	Field	SENSAN	Provide sensor measurements	Flood
Social Network (M)	Cloud	SMA	Generate tweets	Flood, Fire, Heatwave

2.4 Modules

Based on the analyzed user requirements, we have formed a list of modules that need to collaborate through the beAWARE system. Although some of these modules were conceptually defined at the inception of the project, some emerged as critical and necessary during our joint work and analysis. For instance, a Media Hub, which manages requests for media analysis for multiple media types, has been identified as a critical node in the overall architecture in order to create a robust capability to handle incident reports with multiple media attachments of various types.

Table 2-5. Modules

FULL NAME	OWNER	ALIAS	TYPE
AMICO Flood Prediction Services	AAWA	AMICO	Subsystem
Automatic Speech Recognition	CERH	ASR	Backend Module
Central Data Repository	IBM	CDR	Storage Service
Central Message Bus	IBM	MSB	Messaging Service

FULL NAME	OWNER	ALIAS	TYPE
Crisis Classification	CERTH	CRCL	Backend Module
Image Analytics	CERTH	IMGAN	Backend Module
Knowledge Base Services	CERTH	KBS	Backend Module
Knowledge Base Repository	IOSB	KB	Storage Service
Media Hub	CERTH	MEDHUB	Integration Service
Mobile App. for Citizens / First Responders	IOSB	APP	User Application
Multilingual Report Generator	UPF	MRG	Backend Module
Multilingual Text Analysis Service	UPF	MTA	Backend Module
Public Safety Answering Point	MSIL	PSAP	User Application
Sensor Analytics	IOSB	SENSAN	Storage and Analysis Service
Social Media Analysis Services	CERTH	SMA	Backend Module
Video Analytics	CERTH	VIDAN	Backend Module
Weather Forecast Services	FMI	WFS	Data Service

2.5 Operational Entities

Based on the analyzed user requirements, we have formed a list of operational entities identified for exchange among the various data originators and recipients (modules) in beAWARE.

Table 2-6. Operational Entities

Operational Entity	Description	Originators	Recipients
Weather	Weather forecast provided by FMI based on the HIRLAM meteorological data analysis system, in a 7-km grid resolution	WFS	AMICO
Flood Forecast	AMICO imports meteorological forecasts provided by FMI, to be used as input for the Flood Forecasting Model. Then the model runs hydrological-hydraulic simulation to obtain as results the time series of forecasted water level in each river section. Every new result of AMICO is imported by SENSAN and made available for the other beAWARE modules	AMICO	SENSAN, CRCL

Operational Entity	Description	Originators	Recipients
Pre-Emergency Flood Early Warning	During the pre-emergency phase, CRCL checks from AMICO's result imported in SENSAN if the predicted water level exceeds some fixed threshold in one or more river section and estimates the severity level of the forecasted crisis. If this exceeding occurs, CRCL generates and proceeds early warning messages to PSAP in order to alert it for extreme weather conditions and dangerous situations before or during a crisis	CRCL	PSAP
Flood Crisis Level (UC 102)	During the emergency phase, CRCL estimates severity level of every incident report from optional fields in the app filled by end users (category of incident report, estimate of the water level, elements in the flooded area), from social media post which are defined relevant by SMA, and from real measured river levels	CRCL	PSAP
Heatwave Index	Indicator of potential for heatwave for the Thessaloniki region, based on temperatures relative to seasonal temperatures, windspeed, humidity, and other factors	FMI	SENSAN, CRCL
Fire Index	Indicator of potential for eruption of forest fire, targeted for the Devesa area in Valencia and for the Thessaloniki forest area	FMI	SENSAN, CRCL
Alert	General alert sent to the public or to specific groups of first responders (police officers, firefighters, etc.)	PSAP	APP
Tweet	Twitter posts that refer to floods, fires and heatwaves	Twitter	SMA
Textual tweet contents	Messages from Twitter posts that refer to floods, fires and heatwaves	SMA	MTA
Visual tweet contents	Images from Twitter posts that refer to floods, fires and heatwaves	SMA	MEDHUB
Incident-Related Tweet Cluster	Spatiotemporal clusters of tweets that refer to an incident	SMA	KBS
Field Report	Report sent by either citizen or first responder using the mobile app.	APP	KBS, MTA

Operational Entity	Description	Originators	Recipients
Textual Message	Textual part of a field report sent by either citizen or first responder using the mobile app.	APP	MTA
Sensor Measurement	Measurement made by a sensor	SENSOR	SENSAN
Sensor Event	Event detected by Sensor Analytics.	SENSAN	KB
Incident	The KBS module generates incident reports that contain aggregated knowledge from other modules (such as the APP, IMGAN, VIDAN, etc.) and sends it to PSAP.	KBS	PSAP
Team	Group of first responders who move and act together.	APP	PSAP
Task	Assignment of a request to handle team an incident or pre-emergency situation (e.g. evacuate trapped citizens, place sandbags, extinguish fire) to a team of first responders.	PSAP	APP
Multimedia Object	Image, video, or audio file captured by mobile app or on social media.	APP	MEDHUB, VIDAN, IMGAN, ASR
Analyzed Image/Video	Images and Videos originally extracted from tweets or posted directly through the APP, that after proper analysis show detected targets like people or vehicles in danger	IMGAN/VIDAN	KB
Transcribed Speech	Transcriptions of audio recordings provided by mobile app. Audio files are analyzed by the ASR module and text transcriptions are saved in MongoDB along with audio timestamp and language information.	ASR	MTA
Analyzed Text	Events alongside involved entities, location and temporal aspects that are extracted from (the textual parts of) tweets and field reports, as well as from transcribed spoken communications	MTA	KB; SMA (only for information extracted from tweets)
KB graph	Ontological representations of the inferred situation (incidents, vulnerable objects, severity, etc.)	KBS	MRG

Operational Entity	Description	Originators	Recipients
Verbalised Report	Textual report that delivers in the targeted language (English, Greek, Italian or Spanish), the KB inferences regarding the unfolding situation (what is happening, what entities are impacted, etc.)	MRG	KB

2.6 Data Access Modalities

Based on the analysis of User Requirements, we have concluded that several modalities of data access are required, as summarized in Table 2-7.

Each modality provides data access in a different manner.

Table 2-7. Data Access Modalities

Data Access Modality	Definition	Suitable Uses	Examples	Comments
Database	Raw, unchanging data , available in relational (SQL) or elastic (NoSQL) databases	Access to data which is mostly static, published periodically, updated in low frequency, and queryable.	Weather Forecast, Global Parameter Sets, Common Knowledge	Data may be updated every several hours and grabbed by its consumers periodically in a suitable sampling frequency
File Server	Raw, unchanging data , available in structured formats (XML, JSON, CSV, etc.)	Access to data which is mostly static, published periodically, updated in low frequency, and batch-downloadable.	For Future Use	Data may be updated every several hours and downloaded by its consumers periodically in a suitable sampling frequency
Content-based Routing	Mechanism that receives messages from data providers, determines the relevant recipients according to message content (mostly header fields) and relays the message to its intended recipients	Deterministic routing, complex filter-based routing, workflow-based routing, deterministic message structures (identifiable as "topics")	Any information entity with a single well-known recipient, or multiple potential well-known recipients based on message content	Routing rules are determined by a single authority and dictated to the routing system. Routing can be fully-controlled by data originator by determining the recipients in specifically-defined message attributes.

Data Access Modality	Definition	Suitable Uses	Examples	Comments
Publish & Subscribe	Mechanism that receives messages from recipient-agnostic data providers under various topics, and redirects the messages to any recipient that subscribed to its topic.	Dynamic routing, dynamic data provider/consumer composition, dynamic topic composition, recipient-agnostic data exchange policy	Any topic with a dynamic list of multiple potential recipient-agnostic producers and recipients multiple potential recipients	Routing is determined by topic subscriptions only. Producers cannot control the recipient list for the data they provide.
Object Storage	Storage for large and static files	File sharing, file exchange	Video, image, audio, and data files	Access to the files is through URLs, file uploading is accessed through a RESTful API
RESTful API	Data is provided as a response to an HTTP request, or pushed through an HTTP request followed by an acknowledgement response	On-line querying of dynamic information, data provider exposing a generic interface for external consumers to query it, data consumer exposing a generic interface for external providers to push data to it	Sensor data request (pull), Flood Forecast (pull), real-time weather data request (pull), Google search results (pull), on-line logging service (push), messaging service (push)	Pull: API serves for data acquisition by requester from provider upon request (with or without query arguments) through data response. Push: API serves for data posting by provider to receiver upon request (which includes the data), followed by an acknowledgement or rejection response from receiver
Direct data communication	Data is provided to recipient through an ad-hoc interface with a predefined structure known to provider and receiver	Dedicated or legacy interface over, e.g. Ethernet (TCP/UDP), serial (RS232/RS422 etc.), PCI/PCIx, USB, etc.	Video streams, serial raw data streams from devices, hardware, sensors	

Data Access Modality	Definition	Suitable Uses	Examples	Comments
Data display	Visualization of data and information to a user on a computer monitor, mobile device, or dedicated display	Interaction with users	Incident reporting (mobile), Incident lists (desktop), on-map event display (both), GPS signal indication (dedicated)	

3 Data Source Integration Infrastructure [PUBLIC]

3.1 Scope

The purpose of this chapter is to describe the central mechanisms for data source integration in beAWARE, based on the previously-defined data access modalities.

For each infrastructure mechanism, we provide an overview, high-level description of its architecture, and the types of supported data and information entities, as well as the relevant use cases, modes, and data access conventions applicable to this kind of technology or application.

Since beAWARE is a complex system with many modules developed separately across Europe, it became evident quite early in the architectural process that a central messaging service must be set-up, rather than allow a mesh of peer-to-peer interfaces among the multiple beAWARE modules. Moreover, all the communication must pass through a central manager, which provides governance, routing, delivery assurance, and monitoring. For this reason, the main element of the infrastructure is the message bus (MSB).

In addition to the MSB, we have identified the need for a central media repository, that will be used for storing images, videos, audio files, and data sets, so that rather than route this kind of information as raw attachments to messages, references to the media files will be provided along with the messages, such that media access will only be on the basis of necessity, and network traffic would be minimized.

Another means of data storage, mostly for plain data objects, has been put in place in the form of a NoSQL database, which allows easy access to non-relational data. This solution is mostly used for storing social media posts and their analysis results, whose structure could be flexible and less suitable for a relational SQL database.

As all of these services rely on existing, off-the-shelf technologies, this section is public. We strongly advocate the utilization of such a combination of solutions as part of an overall architecture for a complex system of systems in general and especially for an emergency management solution like beAWARE.

3.2 Central Message Bus – MSB

3.2.1 Overview

The Central Message Bus is used as the main vehicle for exchanging information and notifications among different components in the platform. Following a micro-services approach each component is autonomous to a large extent and all interactions among components are achieved via the central message bus. Hence, each component that has information that could potentially be of interest to other components will send a message through the message bus, and each component that needs to be aware of specific notifications from other components shall subscribe to receive messages sent on a specific topic. Different components interacting via the message bus need to agree on a topic name and message format so that information is received and understood.

The communication bus is realized by using an instance of a MessageHub service[1], deployed in IBM's BlueMix cloud. The back-end is based on a cluster of Apache Kafka servers, and the interaction with the service is realized using standard Kafka clients, embedded within the various components.

3.2.2 Supported Entities

Most of the system components make use of the communication bus as this is the only way for communicating between different components.

In a representative flow a new piece of information flows into the system via the APP. The APP's back-end uses the message bus as a producer to alert all interested components as to the existence of a new input data item. The interested components, mainly the components in charge of analyzing the incoming data, receive the corresponding message via the message bus callback mechanism, by subscribing to its topic as consumers, and proceed to analyze the new data based on the content of the received message. Once the analysis is finished, once again the message bus is used by the analyzing module, now as a producer of a new topic, to inform interested components that new analyzed information is available. One of the subscribers to this kind of information is the KBS. The KBS turns as a producer to the MRG as a consumer,

[1] https://www-03.ibm.com/software/products/en/ibm-message-hub

via the bus, for a new report to be constructed. In turn, MRG, as a producer, sends the report through the message bus to the KBS, which receives it a sa consumer. Then, the KBS publishes a new message which is consumed by the PSAP, via the message bus.

3.2.3 Connectivity Diagram

Below is a generic connectivity diagram, showing the MSB as the central module. Since the MSB is agnostic to the topics it processes and to the clients that produce and consume messages based on these topics, the diagram shows them as generic – Producers 1, 2, and 3; Consumers 1,2, and 3; and Topics 1, 2 and 3.

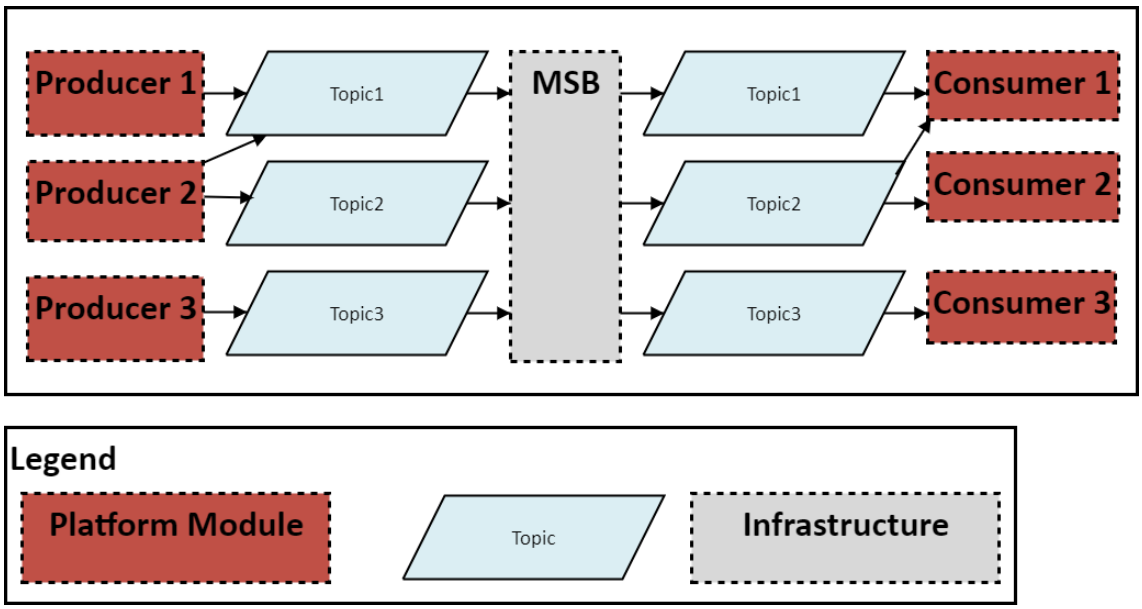

Figure 3-1. Connectivity Diagram - MSB

3.2.4 Features and Capabilities

The abstraction provided is that of a "publish—subscribe" (Pub/Sub) middleware. Pub/Sub is a messaging pattern in which senders of messages, called publishers, send messages without specifying the receivers, but instead categorize published messages into topics. In turn, potential receivers, called subscribers, subscribe to the topics in which they have interest, and subsequently only receive messages they have subscribed to, without knowledge of which publishers originated them. Thus, different components do not need to be aware of each other, or even be active at the same time. In addition, the Pub/Sub mechanism is content-agnostic and endpoint-agnostic, i.e. it does not need to know in advance who are the senders and receivers, and what messages they exchange. The central message bus infrastructure ensures that all messages published on a specific topic reach all subscribers of that specific topic.

3.2.5 Deployment and Operation

The realization of the central message bus is via a cloud deployment of MessageHub, which is a managed Apache Kafka cluster deployed on the IBM cloud (BlueMix). The service itself is provided by a cluster of 5 Kafka brokers (kafka01, .. , kafka05).

The specific connection details are provided to the components running inside the platform's Kubernetes cluster (a managed set of containerized applications that make up a logical unit). The connection details include a namespace wide secret corresponding to the deployment namespace, namely prod. The secret contains an API-key and the list of Kafka brokers.

3.2.6 Data Access Conventions

Access to the message bus capability is provided by using standard client libraries which are available in multiple languages. In addition, there is a REST interface available as well for communication with the message bus.

3.2.7 Authentication

All Kafka agents (producers and consumers) must authenticate themselves with the central service in order to receive or produce messages. The authentication is done using a certificate assigned to the agent, in order to ensure that only allowed data sources interact with the beAWARE message bus.

3.2.8 Common Topic Header

All topics of messages going through the beAWARE Message Bus share a common header, which includes conventional topic, message, and origin identifiers. The attributes in the header are listed in Table 3-1.

The topic header structure generally conforms to the Common Alert Protocol (CAP) [3]. For every attribute we also specify the equivalent CAP attribute. CAP is a standardized format for distributing alerts, warnings, and notifications, especially in emergency-related systems. Since CAP is in principle an XML standard, we created a JSON template for our header that corresponds to the XML Schema Definition (XSD) of CAP alerts. More information about CAP can be be found here: **http://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.html**.

Table 3-1. Common Header for beAWARE Message Bus Topics

Attribute Name	Attribute Description	CAP Equivalent	Mandatory
topicName	name of topic of message (hierarchically allowed)	n/a	Y
topicMajorVersion	topic major version (V in “V.u”)	n/a	Y
topicMinorVersion	topic minor version (u in “V.u”)	n/a	Y
sender	system that originated the message	sender	Y
msgIdentifier	unique message identifier	identifier	Y
sentUTC	UTC time in which message was sent	sent	Y
status	The code denoting the appropriate handling of the alert message (REQUIRED)	status	Y
actionType	type of action to apply to the alert	msgType	Y
specificSender	specific enpoint/user/device/ip address who originated the message	source	N
scope	intended distribution of the alert message (REQUIRED)	scope	Y
district	specific district in which the message is relevant	restriction	Y
recipients	specific recipients' name, id, ip address, etc. within the subscriber to whom the message is intended	addresses	N
code	identifier of the reality to which the message applies	code	Y
note	textual notification for various purposes, e.g. error message (when actionType == “Error”)	note	N
references	optional identifier of previous messages to which the current message refers, or URI in which the content of the current message is stored (multiple values allowed)	references	N

3.2.9 Topics

The beAWARE message bus topics are listed in Table 3-2 below. For each topic, a unique identifier in the format TOPXXX_TOPIC_NAME has been defined. Each topic's producers (or publishers in the Pub/Sub terminology) and consumers (subscribers) are listed. While the producer is not familiar with the subscribers, the subscribers can know the producer's identity thanks to its specification in the common header "Originator" attribute. In addition, an indirectly-routed message can have a "sender"

which is not the "originator". A link to the detailed structure of each topic is available for authorized readers with access to the beAWARE WIKI repository.

Table 3-2. beAWARE Message Bus Topics

Topic ID	Description	Producers (Publishers)	Consumers (Subscribers)	Link to Template
TOP001_ SOCIAL_MEDIA_TEXT	This topic is used to send and receive recently crawled Twitter posts that contain text	SMA	MTA	TOP001
TOP002_ SOCIAL_MEDIA_IMAGE	This topic is used to send and receive recently crawled Twitter posts that contain images, represented by their ids	SMA	MTA, IMGAN	TOP002
TOP003_SOCIAL_MEDIA_REPORT	Used by SMA for communicating Twitter reports to the KBS.	SMA	KBS	**TOP003**
TOP010_ AUDIO_ANALYZED	Voice recording has been analyzed by ASR module and transcription has been stored in Raw Storage	Mobile app	MTA	TOP010
TOP017_ video_analyzed	A video has been processed and the analysis files have been stored and linked to this message	VIDAN	PSAP	top017
TOP018_ image_analyzed	An image has been processed and the analysis files have been stored and linked to this message	IMGAN	PSAP	**TOP018**
TOP021_ INCIDENT_REPORT	Report incident from mobile app	APP	KBS, MTA, MEDHUB, IMGAN, VIDAN, ASR	TOP021
TOP022_ PUBLIC_ALERT	New public alert for the mobile app	PSAP	APP	TOP022
TOP023_ TASK_ASSIGNMENT	New task created for first responders	PSAP	APP	**TOP023**
TOP028_ TEXT_ANALYSED	Extracted information has been stored in the KB	MTA	KBS	**TOP028**
TOP030_ REPORT_REQUESTED	Set of triples capturing the content to be communicated	KBS	MRG	**TOP030**
TOP040_TEXT_ REPORT_GENERATED	Verbalised report (alarm, message, update) in the recipients language has been stored in the CDR.	MRG	KBS	**TOP040**

Topic ID	Description	Producers (Publishers)	Consumers (Subscribers)	Link to Template
TOPC051_ FLOOD_FORECAST	Notification on new AMICO forecast results	AMICO	CRCL	TOP051
TOP101_ INCIDENT_REPORT	Incident Report to PSAP	KBS	PSAP	top101
TOP102_ TEAM_REPORT	Team Report to PSAP	APP	PSAP	top102
TOP103_ TASK_REPORT	Task report to PSAP	APP	PSAP	top103
TOP104_ METRIC_REPORT	Metric data to show on PSAP dashboard	CRCL	PSAP	top104

3.3 Central Data Repository – CDR

3.3.1 Overview

The central data repository (CDR) is intended to store large files that need to be shared between different platform components.

For this purpose, we deployed an instance of the IBM Object Storage in the cloud, providing unstructured cloud data storage, for easy storing and accessing content from various applications.

3.3.2 Supported Entities

Most of the platform components interact with the CDR for uploading and accessing media or data files, or to perform both operations. In a representative flow a new video file is ingested into the system via the APP. The application back-end stores the video file in the CDR and declares the existence of the new file using a message sent via the MSB. The MEDHUB, which subscribed to receive notifications on the same topic, proceeds to access the stored file and analyze it. Once the analysis is completed, a new file is stored into the CDR including conclusions reached by the corresponding analysis component, and information is shared with other system components in a similar manner.

3.3.3 Connectivity Diagram

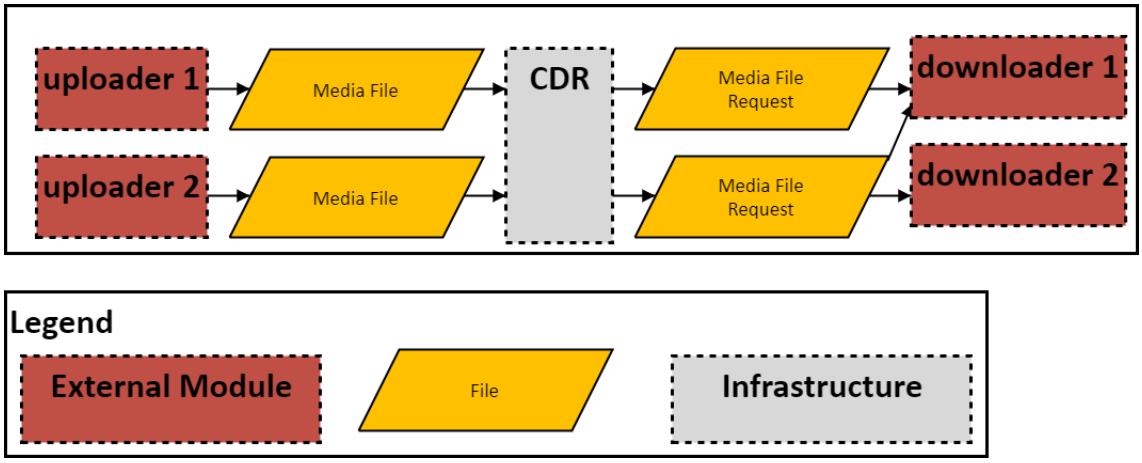

Figure 3-2. Connectivity Diagram - CDR

3.3.4 Features and Capabilities

The CDR is meant as the central secure access point for storing and accessing data by different platform components using a REST interface. It is manifested as a cloud-enabled file system, providing sharing and collaboration capabilities. The CDR serves as a scalable persistent storage layer for the beAWARE platform. Thus, a secure and protected data repository serving different platform components working with different kinds of files is provided.

Readily available SDKs and client libraries exist for a variety of popular programming languages. A management console can be used to configure the generic service for the specific access patterns. In addition, access policies can be configured based on the needs of the different system components.

3.3.5 Deployment and Operation

The realization of the CDR is via a cloud deployment of IBM's Object Storage service instance.

3.3.6 Data Access Conventions

Currently there are two main ways to access the central data repository, directly via a client library or through a helper application which is deployed as a cloud foundry application (https://object-store-app.eu-gb.mybluemix.net) inside the project space in

the IBM cloud. The convenience function exposes a REST interface for uploading, retrieving, and deleting files.

3.4 NoSQL Database

3.4.1 Overview

The NoSQL capabilities are mainly used for storing and accessing relevant tweets information. The information is stored as JSON documents, and pointers to it are passed between components through the message bus.

For the realization of the NoSQL capabilities we use an instance of a MongoDB deployed on BlueMix (https://www.compose.com/databases/mongodb).

3.4.2 Supported Entities

The main users of this capability are the SMA and MTA components. In a representative flow, the tweeter crawler feeds the SMA component with new tweets. If the SMA determines that a certain tweet is relevant, it stores the tweet in a JSON format in the NoSQL DB, and publishes the corresponding link via the message bus. The text analysis component in turn picks up this information from the message bus and extracts the relevant tweet information from the DB using the supplied link.

3.4.3 Deployment and Operation

The NoSQL capability in the platform is realized via a cloud deployment of a MongoDB instance in IBM's cloud. A Kubernetes secret, named mongo-bw2-secret, is injected into the platform namespace providing the required connection information in the form of the corresponding URI.

3.4.4 Data Access Conventions

The DB can be accessed using standard client libraries which are available in multiple languages.

4 Module-Level Data Access Mechanisms [CONFIDENTIAL]

The purpose of this chapter is to describe the mechanisms for data access in each subsystem and module in beAWARE. For each module, we provide:

- an overview of the module and its purpose or main functionality,
- a connectivity diagram that shows the inputs (on the left-hand side of the diagram) to the module from other modules and mediators (as specified in Section 3), and its outputs to other modules and mediators (on the right-hand side of the diagram),
- a list of receivable inputs including the mapping of the operational entity to a technical data or information entity, and the relevant data access mechanism which facilitates it,
- a list of produced outputs including the mapping of the operational entity to a technical data or information entity, and the relevant data access mechanism which facilitates it,
- a subset of the module's technical requirements from the full list (specified in [4]) that entail interoperability, connectivity, or data access requirements or challenges,
- a list of mechanisms and interfaces of the module for receiving input or generating output.

Altogether, all inputs into a module must be mirrored by an output from another module, and vice versa, such that eventually all the outputs by any module in the system are used by at least one other module. Inputs from external data sources such as social networks or sensors are indicated accordingly, and regarded as boundary-crossing inputs. Similarly, outputs to external data receivers such as operational users are regarded as boundary-crossing outputs.

The services and modules described next rely on dedicated innovative technologies or heavily-customized assets, and may be subject to intellectual property. In addition, specifics of the data exchange among the modules are not necessarily relevant in any context or for any application, and stem for the internal definitions and specifications of the beAWARE platform. Due to these considerations, this section is confidential, and available only to European Commission officers and beAWARE Consortium Members.

The full content of this section is available in Deliverable D6.6 – Data-Source Integration Framework – Confidential Version.

5 Conclusion [PUBLIC]

This document has presented the beAWARE data-source integration framework, aimed to provide robust interoperability among the operational agents, connectivity among the subsystems and modules, and data exchange and information-based interaction capabilities across the system.

We have started with identifying the user requirements which are relevant for interoperability and connectivity, and the common available data access modalities, methods, and solutions.

It became apparent very early in the architecting and high-level design process that the system architecture must rely on a robust and extensible connectivity and data exchange infrastructure, including assets like a message bus (Kafka), central data base (MongoDB), and central data repository (ObjectStore).

In addition, each subsystem and module has been designed to interact seamlessly with the infrastructure as well as with some of the other modules via alternative connectivity channels, for enhanced performance and flexibility. For instance, there is a direct interface between the Flood Forecasting service (AMICO) and the Weather Forecasting service to obtain up-to-date and area-specific weather forecasts

We have presented the high-level specification of the common infrastructure components, including the supported data access modalities and exchanged operational and informational entities. In addition, we have presented the technical requirements and solution specification for connectivity and data exchange services in each of the beAWARE modules and components (available in the confidential version only).

Finally, we have elaborated on the informational entities that we have identified over the course of the project. We have specified their attributes and appropriate data access modalities and utilizing functional flows. These flows provide end-to-end value by supporting information exchange and information-based collaboration among the operational users.

This document marks the outcome of the intensive work of all technical partners as part of task T6.2, with continuous feedback from end-users, in order to ensure the facilitation of a robust data-source integration infrastructure, and the intelligent and efficient utilization of the infrastructure by the modules developed by the partners. During the project it became increasingly critical to ensure the commonality and standardization of data exchange processes, govern the definition of new data topics

and ensure compliance across the board, especially for the seamless exchange and manipulation of information while preserving a common semantic language. In most cases, it was found that the Kafka-based Message Bus was the most suitable means of exchanging data among the modules, including both informative data and notifications regarding the availability of data in other channels. In some cases, RESTful APIs were set-up or exploited for on-demand data requisition and retrieval, which allows for synchronous interactions and transactions, or for asynchronous requests of data when notifications are provided. In addition, during the course of the project and the formulation of the solution, it became clear that a storage solution for large media files (audio, video, image, and dataset) is necessary, and that links to media would be provided as part of Kafka messages rather than raw attachments.

Future post-project or Phase-II extensions of the framework may include monitoring of data acceptance via functional acknowledge messages that will be generated by consumers as topicked messages in their own right in response to incoming operational topics. This would allow assurance of data availability to the relevant stakeholders, whether by the originator of the information or by a neutral monitoring authority that would oversee and later supervise the flow of information in the system. Additional enhancements may include differentiation of service levels according to the criticality and frequency of various topics, as well as the adoption of topic filtering conventions in order to allow hierarchical processing of information (e.g. separation of incident topics by category in order to support easy routing of incident updates to relevant processing services).

To conclude, the current infrastructure has proven to be an optimal solution, in terms of cost-benefit balance, and in the context of supporting the challenging integration of the beAWARE platform. We continue to monitor the correct utilization of the infrastructure by the beAWARE modules, critical extensions and enhancements, and, when the volume and throughput of data reach a critical mass, also the evaluation of reliability, performance, availability, affordability, scalability, and overall service level.

6 References

[1] Y. Mordecai, O. Orhof, and D. Dori, “Model-Based Interoperability Engineering in Systems-of-Systems and Civil Aviation,” *IEEE Trans. Syst. Man Cybern. Syst. (to Appear.*, vol. TBD, no. TBD, p. TBD, 2016.

[2] beAWARE, “D2.1 Use Cases and Initial User Requirements.”

[3] E. Jones, W. Systems, and J. Westfall, “Common Alerting Protocol Version 1 . 2,” no. July, pp. 1–47, 2010.

[4] beAWARE, “D7.2 System Requirements and Architecture.”