xR4DRAMA: Enhancing situation awareness using immersive (XR) technologies

Spyridon Symeonidis*, Sotiris Diplaris*, Nicolaus Heise†, Theodora Pistola*, Athina Tsanousa*, Georgios Tzanetis*, Elissavet Batziou*, Christos Stentoumis‡, Ilias Kalisperakis‡, Sebastian Freitag§, Yash Shekhawat§, Rita Paradiso¶, Maria Pacelli¶, Joan Codina‖, Simon Mille‖, Montserrat Marimon‖, Michele Ferri**, Daniele Norbiato**, Martina Monego**, Anastasios Karakostas*, and Stefanos Vrochidis*

* Information Technologies Institute - CERTH, Thessaloniki, Greece
Email: {spyridons, diplaris, tpistola, atsan, tzangeor, batziou.el, akarakos, stefanos}@iti.gr

† Deutsche Welle, Bonn, Germany
Email: nicolaus.heise@dw.com

‡ up2metric P.C., Athens, Greece
Email: {christos, ilias}@up2metric.com

§ Nurogames GmbH, Cologne, Germany
Email: {sebastian.freitag, yash.shekhawat}@nurogames.com

¶ Smartex, Prato, Italy
Email: {rita, m.pacelli}@smartex.it

‖ Universitat Pompeu Fabra, Barcelona, Spain
Email: {joan.codina, simon.mille, montserrat.marimon}@upf.edu

** Autorità di Bacino Distrettuale delle Alpi Orientali, Venice, Italy
Email: {michele.ferri, daniele.norbiato, martina.monego}@distrettoalpiorientali.it

***Abstract*—In this paper we describe the xR4DRAMA system, a solution that makes use of XR capabilities to support professionals who deal with disasters, man-made crises or media productions. The key contribution of this work in progress is the increase of situation awareness, which is achieved by the innovative combination of data collection, multimedia and sensor analysis, linking data, GIS and interactive XR technologies. The proposed platform is designed to facilitate the creation of immersive environments using semantically enriched content and comprises a powerful tool that is applicable to multiple real use case scenarios.**

***Index Terms*—augmented reality, virtual reality, extended reality, situation awareness, disaster management, journalism**

I. Introduction

Decision making, the process of gathering and combining information to evaluate alternative solutions or actions, is an integral part of our everyday life and in most cases it happens instantly, in a matter of seconds. It becomes particularly complicated if we cannot completely feel and anticipate the results of the different choices. This constitutes a serious obstacle for professionals who work in quite responsible positions (e.g. first responders) and whose decisions have a huge impact on society.

The objective of xR4DRAMA is to provide an interactive system that optimises decision making and planning of an intervention by reinforcing the understanding of a given environment. More specifically, XR technologies are leveraged to: a) facilitate the provision of all relevant content required to organise necessary actions, b) generate a realistic experience of an environment, c) improve communication and collaboration between team members by creating a shared projection of an environment and d) enable continuous updates of the representation so as to cater to the dynamic changes of a situation.

The xR4DRAMA system is currently under development and aims at aggregating the outputs of many different state-of-the-art technologies. The end result will be a set of AR and VR tools enhanced with wearable sensors that is going to be assessed in two pilots related to disaster management and media production planning.

The rest of the paper is organised as follows: Section II expands on situation awareness and the levels that can be accomplished within xR4DRAMA, whereas section III describes the use of the platform in the two pilots. Next, the technical details are provided, starting with the system architecture (Section IV), and then continuing with an outline of the back-end and front-end modules (Section V and VI respectively). Afterwards, in Section VII, we briefly explain the benefits of the selected technologies. Finally, section VIII concludes the paper and discusses the next steps of our work.

II. Situation awareness

The term "Situation Awareness" (SA) originally relates to the work of human factors practitioners who explore the behaviour of professionals in specific (challenging) situations, for instance aviation, air traffic control or the operation of large systems. In these contexts, SA is considered as highly relevant for the quality of human decision-making and performance. Although the definition of the term SA is disputed, xR4DRAMA will follow the understanding as presented by M.R. Endsley in 1995, according to which *situation awareness is the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future* [1]. xR4DRAMA will translate these three elements (1) perception, (2) comprehension and (3) projection into three sequential levels of situation awareness. They each stand for different stages in the management of any given events (e.g. disasters, media productions), but will also be qualified by different degrees of immersion. On a first level (that we will call Level 1 SA), the focus will lie on extracting and presenting *pre-existing* information about a situation that derives from various resources and databases. The visualisation in the user interface on this first level will be rather simple, with the sole purpose of delivering a rough and nevertheless accurate overview. The next step (Level 2 SA) will integrate and visualise more detailed and specific information deriving from people on location and special sensors that will be deployed. Finally, Level 3 SA will present the location (and the related information) in an immersive manner by utilising various XR-technologies and visualisation techniques in order to provide a high degree of situation awareness. In general, all levels of SA are supposed to be accessible to all stakeholders in a specific scenario, albeit in graduations and on different devices. The ultimate goal is to create a shared, distributed and consistent Situation Awareness.

III. Pilot use cases

A. Pilot Use Case 1 - Disaster management

During floods, decision makers and first responders face a stressful scenario. A plethora of tools exists to support them, which is oriented to the management of floods after they occur. However, for the majority of people involved in these crisis events, the most important and helpful information is a representation of the situation before the flood and this content is mostly unavailable. Towards tackling this issue, VR environments can contribute to replicating the principal features of a generic flood.

xR4DRAMA will allow participants to be trained in a safe situation also making use of physiological sensors to monitor response and health data. In order to simulate a natural hazard, the VR scenario will define a hazard/threat predefined by Autorità di Bacino Distrettuale delle Alpi Orientali (AAWA) that makes use of all information in its GIS database. Using this approach, the status of first responders can be monitored, to better assign tasks during emergencies. Technologies developed in the project will also be used in an exercise to demonstrate that the preparation of first responders with specific training based on AR and VR simulations is highly important to learn how to act safely and efficiently during an emergency. In a real-life scenario the xR4DRAMA toolkit can provide first responders with important information and reproduce in real time the actual event inside the control room to allow for well-informed and efficient decision making.

The Disaster Management Use Case will be structured in 3 different phases. The first one represents the preparations before an emergency: the platform will acquire automatically information from sensors, from the web and from other sources in order to reproduce the expected scenario some days in advance. This will allow decision makers inside the control room to verify all emergency procedures. The second phase will operate during the emergency where there will be necessarily a two-way communication between first responders and the control room in order to acquire information from the rescue teams, update the situation awareness and share new data with the teams. During this phase the system will be controlled and guided by operators. The last stage will be mainly focused on the VR, to support a strong communication between the control room and a rescue team involved in rescue operations. The control room, thanks to VR, could simulate the operation in a safe environment in order to provide to the team every information and instruction useful to conclude the mission in safe conditions.

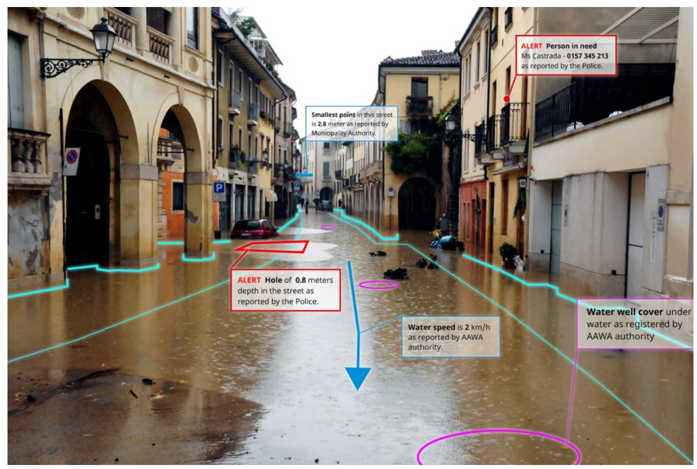

Fig. 1. Draft of possible information provision to first responder via AR in a flooding scenario.

B. Pilot Use Case 2 - Outdoors media production

In Pilot Use Case 2, we will test which kind of immersion and visualisation will create the most suitable and helpful situation awareness in a media production environment. Media production planning is a complex and very often complicated process. It utilises numerous different approaches, strategies and tools depending on the specific kind of production. Production management needs to decide on size and qualification of production staff, the necessary equipment, production schedule, travel arrangements, legal requirements, and many more issues. These decisions require experience, but also detailed knowledge about the general circumstances of the production, and specifically about the location where the production will take place. Is filming possible? Is it permitted? What are the lighting conditions like? Can we produce aerial footage? Is the location accessible? Do we need special equipment? All these questions need to be considered, and mis-planning might easily result in higher costs or even lower production quality. This knowledge - or situation awareness - is particularly relevant for production managers who have not visited and experienced the production location in person.

The second use case in xR4DRAMA will therefore focus on *media production planning*. Level 1 SA will enable production management to decide upon the general suitability of a location and the very fundamental requirements, e.g. regarding staff and equipment, that need to be met. Level 2 SA will be based on more detailed input from so-called location scouts, staff that knows about the specifics of a media production, but is most of all able to provide special and relevant (”insider”) information about the production site. Level 3 SA, finally, will present and visualise the location in an immersive XR-environment that will allow (remote) production management to not only get a ”real feel” for the location but also to (virtually) test certain production methods. These might include camera angles and movements, the use of filters in special lighting conditions or simply the choice of a certain time of the day when circumstances are most likely to facilitate the best possible production output.

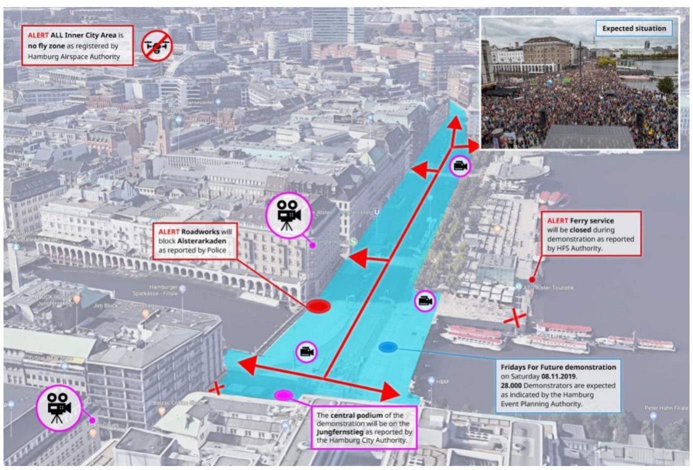

Fig. 2. Draft of possible information provision / planning interface in the case of a demonstration in Hamburg city center.

IV. xR4DRAMA ARCHITECTURE

The architecture of the xR4DRAMA system is illustrated in Figure 3. It consists of 5 layers, each one of which indicates a high-level step of a pipeline that starts with raw data and results to advanced immersive representations visualised in the user tools. Multiple modules are encompassed in each tier so as to support complementary requirements of the system. This structure helps us to clarify the data flow as well as the role of components based on the section they are in the architecture.

The users of xR4DRAMA interact with the front-end of the system (blue layer in Figure 3), which includes various data visualisation solutions targeted for different groups. The field actors make use of the AR and the mobile application, whereas the people in the control room mostly work on the VR and the desktop authoring tool. The **XR tools** communicate with the back-end using a single point of access, a centralised interface (API). This interface is the **middleware** responsible for relaying the user requests and the information generated by the back-end modules. Lastly, the back-end of the system is divided into the rest 3 tiers. **Data Collection** accumulates the required raw content (multimedia and sensors), Data **Analysis and Reconstruction** processes the content to extract additional information, metadata and 3D models, and **Linking tools** organise and interconnect the raw data and analysis results into semantically enriched items.

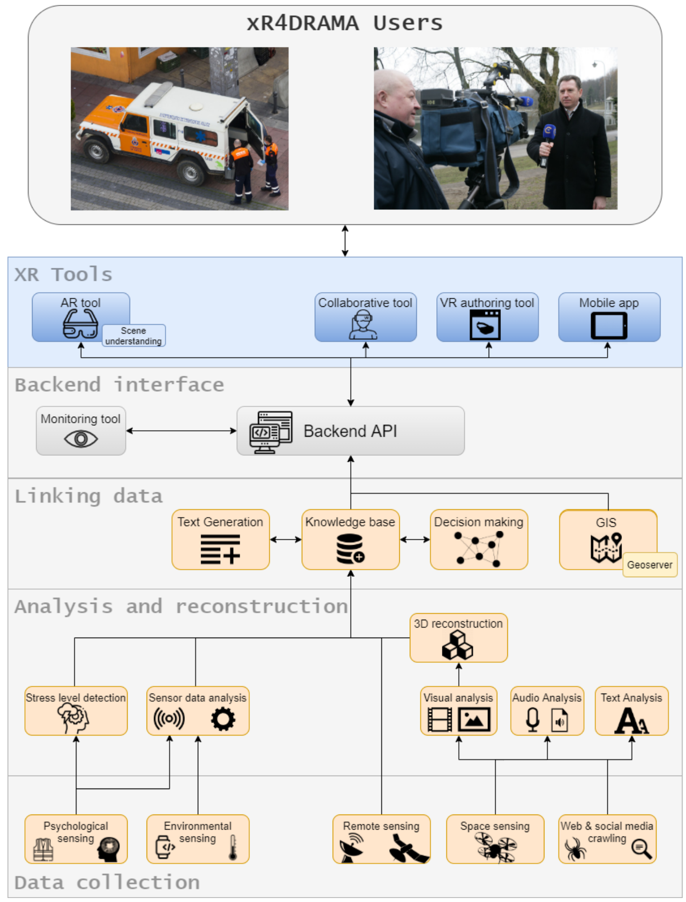

Fig. 3. Architecture of xR4DRAMA.

V. BACK-END MODULES

A. Data collection

1) Data acquisition from Web: The first data collection module fulfils the requirements for automatically fetching and providing open online Web data to the xR4DRAMA system. The target resources are: a) **traditional websites**, such as blogs, b) **social media**, for example Twitter and c) **open repositories** that provide access using an application programming interface (API). Several different mechanisms are applied in order to accommodate for different types of multimedia sources.

The data acquisition process from Web is performed in four phases. For each one of these, the followed approach and the employed functionalities may be different according to the use case of the module.

In the first phase, the requirements definition, the input of the module is manually specified in the form of URL addresses or textual queries. The following stage is an optional one and is considered only when we need to discover new resources. Web crawling or web and social media search is used depending on the input type of the previous step. The third phase is content extraction, where web scraping or retrieval from an API is executed upon the indicated and discovered resources. The goal in both operations is to retain only the information that is meaningful for the system. In the final step, the obtained content is integrated into the platform. The heterogeneous multimedia files and metadata are parsed and stored using a unified representation model in order to be easily reused by the rest of the components.

2) Remote sensing: In xR4DRAMA, the remote sensing service is facilitating the need for rapid availability of initial terrain and visual information for large scale environments, based on the satellite platforms' data availability. The remote sensing module supports the need for immediate visual data that cover the area of an incident. More specifically, this module enables the generation, or retrieval of an initial rough 2.5D/ 3D terrain via satellite data before sending agents to the field and without the need of deploying in-situ 3D recording means. Satellite imagery and elevation data are fetched prior to a field mission from Copernicus (Sentinels satellites and in-situ environmental sensors), open access hub web services, and Data and Information Access Services (DIAS) that are used for centralised access to Copernicus data via cloud platforms. More specifically, sobloo service, which is one of DIAS, provides very-high resolution imagery (<50cm resolution) from Pléiades satellites and Airbus streams existing world Digital Elevation Models (DEM). The user is able to seamlessly exploit any of the aforementioned services in the xR4DRAMA platform to make an initial draft model of the area.

3) Space sensing: The immersiveness of the user experience in the platform very much relates to the completeness and the realistic feeling of the 3D models in VR. To support this need the xR4DRAMA platform allows the user to collect and feed one's own drone images. For the purposes of the project demos images are collected, via exploiting quadrocopter and fixed-wing drones. The unmanned Aerial Vehicle (UAV) is deployed in the area of the incident and captures the appropriate images to reconstruct the 3D models that are visualised in the VR modules of the authoring tool, and also run the visual analysis; these increase the fidelity of terrain created by the remote sensing module.

Vertical and oblique aerial drone imagery are collected to elaborate the visual data acquired by remote sensing.

4) Physiological sensors: To detect the physiological data in the frame of the project, a garment with integrated electrocardiogram (ECG) electrodes and a respiratory movement sensor is used. Textile sensors are made by seamless knitting technology and integrated in an elastic textile frame that is the sensing part of an inner customisable garment. The ECG electrode conductive component is stainless steel in form of fiber that is spin with polyester. The ECG acquired by the two electrodes offers a lead very close to the 1-lead ECG. Instead, the respiration sensor is made with a conductive yarn based on polyamide coated with carbon, the whole structure is elasticised with Lycra® elastomer. This structure confers its piezoresistive proprieties to the carbon sensor that is able to detect the movements of the thorax due to respiration activity, i.e. recording the change in the user's chest circumference. The use of a textile detection system allows to maximise comfort for users and to easily position the sensors, without the need for specialised personnel [2].

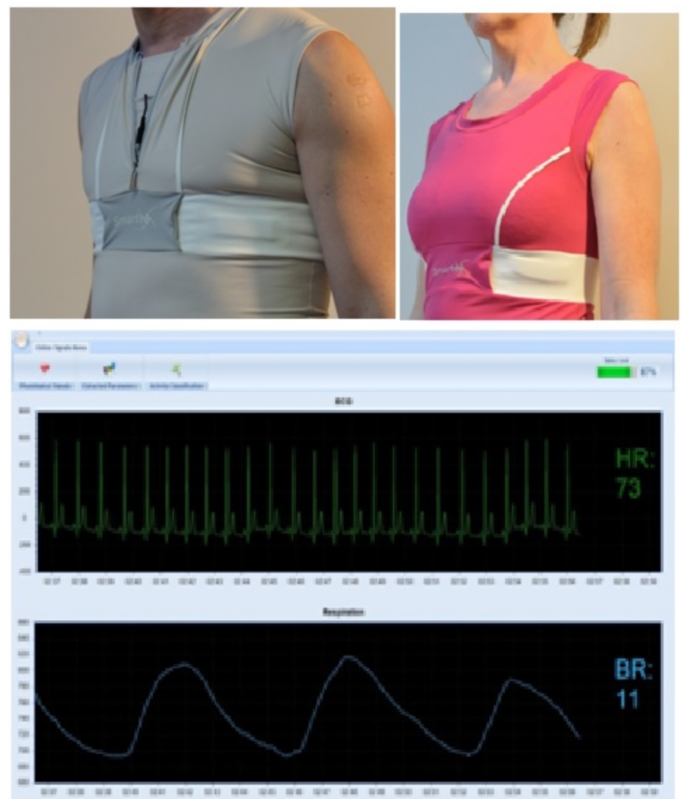

Fig. 4. On the top, an example of the textile sensing platform for man and for woman and on the bottom, an example of the ECG and Breathing Rate (BR)

The sensing area is connected to a portable device (Figure 5) for the acquisition, processing and transmission of the signals. In the portable device, accelerometers, gyroscope and Inertial Measurement Units (IMUs) sensors are also integrated. The electronic device is placed on a pocket at the level of the sternum of the thorax to detect the movements of the trunk and the posture.

5) Environmental sensors: A feasibility study on the use of wearable environmental sensors addressing the use case scenarios is also going to take place. In particular, the use of wearable pressure sensors is evaluated in combination with the posture information provided with the aim to investigate the possibility to design and develop a system that is capable

Characteristics	
Power management	
Battery	Lithium-polymer, 660 mAh
Battery charging	Through USB
Storage capability	
Memory	Micro-SD card, 16 GB
Autonomy*	
Autonomy while recording	Up to 30 hours
Autonomy while streaming	More than 12 hours
Memory capacity (with 1GB)	More than 400 hours
Temperature	
Operating	0 to 50°C
Storage	-20 to 50°C
Humidity	Operational up to 95% non-condensing
Device	
Dimensions	63 x 52 x 16 mm^3
Weight	45 g

Fig. 5. The portable electronics

of detecting data in the field to monitor the stability of a human body in flood waters and to acquire parameters correlated to the vulnerability curve.

B. Analysis

1) Sensor data analysis: The sensor data analysis task is dedicated to the extraction of information from physiological and environmental sensors. The physiological sensors help us detect the stress level of the first responders and decision makers. Their analysis flow starts with the preprocessing of the signals, feature extraction and selection and finally the application of a prediction algorithm. In order to train the predictive models, data are collected before-hand, by subjects wearing the smart vest that will be used in the use case. The trained model will later be implemented on the near-real time scenario, producing results based on the optimal time window chosen.

Multi-sensor approaches are more common in stress level detection tasks. One of the basic factors for stress is the electrocardiography (ECG) monitoring. In [3] the authors performed personalised individual stress analysis using ECG signals alone. Employing fiducial and entropy features, machine learning algorithms and feature selection methods they classified stress of automobile drivers in three categories. In [4] the authors combined ECG with EMG (electromyogram) signals to detect stress in two, three and four levels. The authors followed the basic pipeline of signal processing, feature extraction and selection and finally machine learning algorithms. Different signals were extracted from ECG and EMG signals. In [5] four sensors are combined, namely ECG, GSR (Galvanic Skin Resistance), respiration and temperature, in order to detect stress levels of participants that were subjected to some stressful situations. The authors performed interviews of the participants for stress annotation.

2) Stress level detection: According to the World Health Organisation "Work-related stress is the response people may have when presented with work demands and pressures that are not matched to their knowledge and abilities and which challenge their ability to cope" [6]. The detection of stress in persons that are involved in a complex situation can be an objective measure that people dealing with an emergency are being surpassed by the situation. xR4DRAMA monitors the stress level of the people involved in a disaster management system to check if the people involved are under more pressure than the one they can assume.

Physiological sensors provide reliable data to measure stress level, but they are not always available because only some members of a team wear the sensors. Also external people, that do emergency calls, are not expected to have these sensors. We include a speech stress detector in the system to increase the range of users that can be monitored. Using speech is an inexpensive non intrusive and ubiquitous way of stress detection.

There is ongoing work on stress detection from speech, the most classical schema is based on feature extraction (e.g. pitch, jitter, energy, etc.) followed by some machine learning algorithms that can be based on KNN, SVM or AdaBoost among others [7]. Current papers tend to apply different neural network structures (e.g. [8] use convolutional while [9] use Attention based). In xR4DRAMA we extend the speech features with prosodic features to enrich the inputs to a neural network classifier to guess the user's stress level.

The stress level obtained from speech is combined with the one provided by the physiological sensors (when available) to provide a global and personal stress measurement to the managers so they can focus on the most critical subjects.

3) Visual analysis: The xR4DRAMA visual analysis services consist of the Scene Recognition and the Spatio-Temporal Building and Object Localisation modules. Their aim is to assist the 3D reconstruction and provide all extracted information to the system. More precisely, Scene Recognition extracts information about the type of scene (e.g. "outdoors", "indoors") and the type of the depicted area (e.g. village, forest) or building (e.g. hospital) in the analysed images or videos. In addition, information about the existence of flood or fire in the analysed visual content is drawn. Deep learning models trained on proper datasets are deployed. For example, in [10] the authors use Convolutional Neural Networks (CNNs) and propose a new loss tailored for scene recognition, named scene coherence loss (SCL). In the case of input videos, techniques for shot detection and keyframe extraction are used to facilitate their analysis. This module's information help us decide which images or video shots are useful for the xR4DRAMA system so that we further analyse them. Further analysis is conducted by the Spatio-Temporal Building and Object Localisation module that determines the presence of buildings and/or other structures, like construction elements (e.g. bridges, roads, pavements), and possible flood or fire textures contained in the analysed scenes, extracting the corresponding segmentation masks. Spatio-temporal video segmentation methods are deployed, like the novel spatio-temporal CNN-Based Markov Random Field (MRF) proposed in [11]. For better segmentation results, Photorealistic Style Transfer is applied to the images or video frames before the semantic segmentation. Photorealistic style transfer aims to transfer the style of one image to another but preserves the original structure and detail outline of the content image, which makes the content image still look like a real shot in different lighting, time of day or weather. In [12] the authors propose a wavelet corrected transfer based on whitening and coloring transforms (WCT^2) that allows features to preserve their structural information and statistical properties of VGG feature space during stylisation.

4) Audio and text analysis: The role of the language analysis component of xR4DRAMA is to analyse audio and written information that is collected from different sources, including communication data from the different actors, social media, web repositories, etc. Speech data in audio format are converted to text by state-of-the-art off-the-shelf automatic speech recognition (ASR) module. Target languages include English, German and Italian. The resulting text output and the acquired textual material are processed by the text analysis component that provides abstract conceptual representations with disambiguated entities as outputs, which can be readily used to populate ontologies. Written language analysis includes named entity identification, entity linking, word sense disambiguation, concept extraction, co-reference resolution, and syntactic and semantic parsing.

5) Space Modelling: The scope of this component is to exploit imagery from UAVs, web resources, and satellites to progressively generate *3D models of various resolutions* from draft to high-fidelity models of urban and country areas; the different levels of detail lead to different levels of situation awareness depending on the available data and the user needs. The workflow for visual data processing includes Structure-from-Motion, simultaneous localisation and mapping, stereo-reconstruction, as well as multi-modal data fusion algorithms. More specifically, satellite images and available elevation data are used to rapidly extract and reconstruct the initial landscape for the authoring platform. The initial draft landscape from remote sensing is fused with the 3D models that are produced from in-situ drone images; the processing and the data fusion of 3D data consist of mutual registration, mesh model creation and photo texturing of the models via the available image and satellite data. The reconstructed 3D models aim at increasing situation awareness of the users, thus the geometry and texture, but also the computational complexity is optimised for the use in the XR interactive tools. The user has the option to choose among the different available sources to build the 3D model that better suits a specific case's needs.

C. Linking data

1) Multimodal fusion: This module combines the heterogeneous data coming from the different modalities involved in the scenarios, like audio, video and sensors. The combination of data varies according to the required outcome of this analysis. The first fusion task is the stress level detection from the two modalities involved, namely sensors and audio. The fusion is conducted in the results level with the use of relevant methodologies like majority voting, averaging and stacking, just to name a few. Another fusion task is the combination of the different physiological sensors embedded in the wearable, as described earlier. Finally fusion is applied to combine video, audio and sensor data, in order to receive an integrated result of the analysis. In relevant literature, neural networks have been proven very effective for that type of fusion. A useful guide can be found in [13] where the authors conducted a detailed review on deep learning in remote sensing. As a last step of the fusion processes, the results are semantically integrated and stored into a knowledge base (KB).

2) Language generation: The role of the Natural Language Generation module is to communicate to end users, in a natural human-friendly way and several languages (via coherent sentences and/or texts), the linked data stored in the Knowledge Base. Due to the lack of datasets in the crisis management domain, and to the high cost of creating them [14], applying end-to-end neural techniques is made very difficult. Instead, we tackle description generation from the fused ontological information using FORGe, a portable grammar-based generator that can take structured data as inputs [15]. The input fused data, under the form of ontological triples, is mapped to minimal predicate-argument (PredArg) templates, which are then sent to the generator. The generation consists of a sequence of graph-transduction grammars that map successively the PredArg templates to linguistic structures of different levels of abstraction, in particular syntax, topology, morphology, and finally texts. PredArg structures are similar to the *Message triples* in NaturalOWL [16], but all predicates in the PredArg structures are intended to represent atomic meanings for improving the flexibility. The first part of the generation pipeline, which produces aggregated PredArg graphs, is comparable to the ILEX system [17], while the surface realisation is largely inspired by MARQUIS [18]. In our pipeline however, two types of aggregation take place during generation, one at the predicate-argument level (in a NaturalOWL fashion), and one at the syntactic level.

The main challenge in xR4DRAMA consists in extending the generator (grammars, lexicons) to cover the idiosyncracies of the fused data, and crafting domain-specific rules, in particular for aggregation, to improve the fluency of the outputs.

3) Decision support: The aim of the Decision Support System (DSS) module is the extraction of information in the domains of media production and disaster management from the xR4DRAMA Knowledge Base (KB). The correlation of this information, underlining their similarities and enabling predictions is going to be the guideline for the decision makers.

In our case we have to cope with partial and imperfect information coming from heterogeneous sources such as multimedia (e.g. textual, visual, audio), social-media analytics, geo-spatial data and historical data from other modalities. We handle this uncertain information with the use of fuzzy and/or non monotonic reasoning approaches and methodologies (e.g, FuzzyDL reasoner) prioritise the most relevant proposed actions.

The ontological structures that are developed constitute the foundations of the reasoning mechanism and on top of it the query formulation takes place. The query information consists of all the data extracted from users regarding entities, relations and objects.

4) GIS: It is the geospatial database with 2D and 3D content and a reference frame for the underlying localisation platform that allows all relevant data to be suitably placed in 3D space. The geographic information system (GIS) connects to all relevant processed data and manages them in the geo-referenced system via geospatial queries, in order to support the XR interaction, mainly the AR in outdoor cases, but also the VR authoring tool. Information is served to users either on demand, or via the spatial coherence and other set of rules implemented through spatial queries. All available 2D information and 3D reconstructed areas/buildings of interest are organised in layers to be efficiently accessible to the AR module and to support awareness in the VR by providing its 3D space and terrain.

VI. Front-end tools

A. AR tool

The requirement to provide seamless interaction to the user in the field is supported by the Augmented Reality (AR) module; AR augments the physical world via applying layers of information based on the user needs. This information is streamed from available data sources in xR4DRAMA platform: open geodata, the stakeholders' "archives", the data captured during the response process by xR4DRAMA tools, i.e. news coverage, social media streams, visual data, prediction models and the results from the analytics in DSS. AR devices allow to position first responders findings in real 3D space by seamlessly merging real 3D space with the virtual data connected to the GIS and, hence, the xR4DRAMA platform. These data to superimpose on the real world are narrative information (e.g. history, protocols), real-time streamed data and 3D information, such as infrastructure, obstacles and restricted areas etc. Initially, the AR tool is developed for an Android tablet using ARcore to efficiently handle mobile app resources. As a further improvement the tool will be deployed in Hololens headset that can, among other advantages, provide real-time information and a view from the actor's viewpoint. Stakeholders in the control room can use this info for support, supervision, and advice to the on-site personnel. In the media production use case, this kind of interactive walk-through can enhance the planning process.

B. VR authoring tool

The authoring tool serves as the main front-end for end-users to create and manage a project, as well as to concretely plan a mission and add metadata. Starting from the map-based selection of the envisioned target location, the tool generates basic geo-referenced 3D visualisations of the surroundings (retrieving, e.g., terrain and/or building representations from open sources) to provide a context for planning, as well as showing further relevant information about the site generated by the other platform tools.

Based on the 3D visualisation, the tool allows for the user-friendly and efficient specification of further details to plan the mission, such as labelling relevant locations, highlighting target sites or areas to avoid, or adding metadata that could not be retrieved automatically (e.g., building types and their relevance to the mission, important locations such as power outlets or public restrooms, or adaptations to the environment geometry such as road blocks or passageways not present in the source data).

The output of the authoring tool is available on the platform for other components to consume, and especially serves as the primary input for the collaborative VR tool.

C. Collaborative VR tool

The collaborative VR tool provides users with the opportunity to visit the target location of a planned mission in Virtual Reality in order to facilitate situation awareness, allowing users to check the feasibility of the plan, and to detect missing details or previously unnoticed opportunities (such as favorable shot angles or convenient vantage points). Depending on data sources and user preferences, the environment can be explored in actual size for increased situation awareness or as a diorama for an improved overview.

While it takes its input primarily from the VR authoring tool, it integrates other data sources from the platform as well, e.g., displaying detailed 3D reconstructions (instead of broad building outlines, cf. Section V-B5) where available, and updating with the release of new data.

The environment can be entered with several VR users at the same time, and supports both voice communication and gesture synchronisation, thus facilitating verbal and gestural discussion and collaborative modification of and interaction with the plan's components (such as modifying objects added in the authoring tool, or adding new annotations). In addition, users can explore the environment together as if on site, e.g., by visiting different potential target locations, trying out drone flights for possible video shots, etc.

The tool is compatible with all PC-powered VR headsets via OpenVR/OpenXR, and users are able both to co-locate in the same local network and to distribute to different locations, connecting via internet. Low synchronisation latency ($<$100ms) is supported by allowing for flexible server locations and a latency-optimised synchronisation architecture.

D. Mobile application

The mobile awareness application's aim is to reduce the amount of poorly informed people in the area when a disastrous incident takes place. The mobile app is able to detect the user context based on variables such as location, time and proximity to dangerous areas, and informs the user about the event current status, possible threats and alerts. Moreover, a multimedia data exchange is achievable between the user and the control room, with the user making a proper use of mobile's sensors.

The application is a tool for citizens during emergency situations that allows them to report and access events regarding several points of interest (infrastructures, hazards etc) which have specific geographical coordinates. Furthermore, real-time feed with updating information about an incident or security instructions are also provided.

VII. Advantages overview

The back-end and front-end tools that were described bring many advantages to the xR4DRAMA system, the main of which are summarised as follows:

- It makes use of multiple data sources and it provides the ability to reuse existing online multimedia (to not always generate content from scratch).
- Monitoring of actors in the field is significantly improved by the data feed received from physiological sensors.
- Analysis is performed on a multi-modal level and, as a result, it exploits the potential of various inputs to enrich the existing knowledge of the platform and reinforce the capabilities of the Decision Support System.
- All raw and processed information are stored in the GIS to make it available for the user tools. This offers the benefit of spatially merging existing information, e.g. city infrastructure, with new data and events, and the capacity to execute geospatial queries, e.g. routing between safe points in disaster scenarios.
- It consists of several front-end tools which are tailored to the needs of different types of users that collaborate into the same team.

VIII. Conclusions

In this paper, we presented xR4DRAMA, a solution for professionals such as first responders and journalists that is still under development. This system consists of various XR tools, leverages information produced by several state-of-the-art interactive technologies, and aims at increasing situation awareness of users, which is essential for making the optimal decisions. We described two real application scenarios and every technical component that composes the platform. The next milestones are to finalise the first prototype of the system and extensively evaluate the outcome into pilots associated with the aforementioned use cases.

Acknowledgment

This work was supported by the European Commission in the context of its H2020 Program under the grant number 952133-IA.

References

[1] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human factors*, vol. 37, no. 1, pp. 32–64, 1995.

[2] D. Curone *et al.*, "Assessment of sensing fire fighters uniforms for physiological parameter measurement in harsh environment," *IEEE transactions on information technology in biomedicine*, vol. 16, no. 3, pp. 501–511, 2012.

[3] I. Bichindaritz, C. Breen, E. Cole, N. Keshan, and P. Parimi, "Feature selection and machine learning based multilevel stress detection from ecg signals," in *International Conference on Innovation in Medicine and Healthcare*. Springer, 2017, pp. 202–213.

[4] S. Pourmohammadi and A. Maleki, "Stress detection using ecg and emg signals: A comprehensive study," *Computer methods and programs in biomedicine*, vol. 193, p. 105482, 2020.

[5] Y. Shi *et al.*, "Personalized stress detection from physiological measurements," in *International symposium on quality of life technology*, 2010, pp. 28–29.

[6] Occupational health: Stress at the workplace. [Online]. Available: https://www.who.int/news-room/q-a-detail/ccupational-health-stress-at-the-workplace

[7] Y. S. Can, B. Arnrich, and C. Ersoy, "Stress detection in daily life scenarios using smart phones and wearable sensors: A survey," *Journal of Biomedical Informatics*, vol. 92, p. 103139, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1532046419300577

[8] A. R. Avila, S. R. Kshirsagar, A. Tiwari, D. Lafond, D. O'Shaughnessy, and T. H. Falk, "Speech-based stress classification based on modulation spectral features and convolutional neural networks," in *2019 27th European Signal Processing Conference (EUSIPCO)*. IEEE, 2019, pp. 1–5.

[9] H.-K. Shin, H. Han, K. Byun, and H.-G. Kang, "Speaker-invariant psychological stress detection using attention-based network," in *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. IEEE, 2020, pp. 308–313.

[10] H. Seong, J. Hyun, and E. Kim, "Fosnet: An end-to-end trainable deep neural network for scene recognition," *IEEE Access*, vol. 8, pp. 82 066–82 077, 2020.

[11] L. Bao, B. Wu, and W. Liu, "Cnn in mrf: Video object segmentation via inference in a cnn-based higher-order spatio-temporal mrf," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5977–5986.

[12] J. Yoo, Y. Uh, S. Chun, B. Kang, and J.-W. Ha, "Photorealistic style transfer via wavelet transforms," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 9036–9045.

[13] X. X. Zhu *et al.*, "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017.

[14] C. Gardent, A. Shimorina, S. Narayan, and L. Perez-Beltrachini, "The WebNLG challenge: Generating text from RDF data," in *Proceedings of the 10th International Conference on Natural Language Generation*, 2017, pp. 124–133.

[15] S. Mille, S. Dasiopoulou, B. Fisas, and L. Wanner, "Teaching FORGe to verbalize DBpedia properties in Spanish," in *Proceedings of the 12th International Conference on Natural Language Generation*. Tokyo, Japan: Association for Computational Linguistics, Oct.–Nov. 2019, pp. 473–483. [Online]. Available: https://www.aclweb.org/anthology/W19-8659

[16] I. Androutsopoulos, G. Lampouras, and D. Galanis, "Generating natural language descriptions from owl ontologies: the naturalowl system," *Journal of Artificial Intelligence Research*, vol. 48, pp. 671–715, 2013.

[17] M. O'Donnell, C. Mellish, J. Oberlander, and A. Knott, "Ilex: an architecture for a dynamic hypertext generation system," *Natural Language Engineering*, vol. 7, no. 3, p. 225, 2001.

[18] L. Wanner, B. Bohnet, N. Bouayad-Agha, F. Lareau, and D. Nicklaß, "MARQUIS: Generation of user-tailored multilingual air quality bulletins," *Applied Artificial Intelligence*, vol. 24, no. 10, pp. 914–952, 2010.