Article

An Extended Reality System for Situation Awareness in Flood Management and Media Production Planning

Spyridon Symeonidis [1,*], Stamatios Samaras [1], Christos Stentoumis [2], Alexander Plaum [3], Maria Pacelli [4], Jens Grivolla [5], Yash Shekhawat [6], Michele Ferri [7], Sotiris Diplaris [1,*] and Stefanos Vrochidis [1]

1 Centre for Research and Technology Hellas, Information Technologies Institute, 6th Km Charilaou-Thermi, 57001 Thermi, Greece; sstamatis@iti.gr (S.S.); stefanos@iti.gr (S.V.)
2 up2metric P.C., Michail Mela 21, 11521 Athens, Greece; christos@up2metric.com
3 Deutsche Welle, Kurt-Schumacher-Straße 3, 53113 Bonn, Germany; alexander.plaum@dw.com
4 Smartex, Via della Croce Rossa 82, 59100 Prato, Italy; m.pacelli@smartex.it
5 Department of Information and Communication Technologies, Pompeu Fabra University, Roc Boronat, 138, 08018 Barcelona, Spain; jens.grivolla@upf.edu
6 Nurogames GmbH, Schaafenstraße 25, 50676 Cologne, Germany; yash.shekhawat@nurogames.com
7 Autorita di Bacino Distrettuale delle Alpi Orientali, Cl. Seconda del Cristo, 4314, 30121 Venice, Italy; michele.ferri@distrettoalpiorientali.it
* Correspondence: spyridons@iti.gr (S.S.); diplaris@iti.gr (S.D.)

Abstract: Flood management and media production planning are both tasks that require timely and sound decision making, as well as effective collaboration between professionals in a team split between remote headquarter operators and in situ actors. This paper presents an extended reality (XR) platform that utilizes interactive and immersive technologies and integrates artificial intelligence (AI) algorithms to support the professionals and the public involved in such incidents and events. The developed XR tools address various specialized end-user needs of different target groups and are fueled by modules that intelligently collect, analyze, and link data from heterogeneous sources while considering user-generated content. This platform was tested in a flood-prone area and in a documentary planning scenario, where it was used to create immersive and interactive experiences. The findings demonstrate that it increases situation awareness and improves the overall performance of the professionals involved. The proposed XR system represents an innovative technological approach for tackling the challenges of flood management and media production, one that also has the potential to be applied in other fields.

Keywords: extended reality; virtual reality; augmented reality; interactive technologies; immersive technologies; disaster management; media planning

1. Introduction

Often, decision making can be a challenging and tedious process because it requires individuals to be both rational and beneficial to the situation. Furthermore, it could require quickly weighing a particular circumstance's benefits and drawbacks and accurately determining the best course of action. Hence, decision making can be emotionally taxing, as it often requires individuals to make difficult choices affecting their lives, or the lives of others, in a matter of seconds.

Situation awareness plays a critical role in decision making. It entails being aware of the current environment and the different elements that may affect the decision-making process, such as the present conditions, the available resources, and the possible outcomes. A high level of situation awareness allows decision makers to make informed choices rapidly and effectively and to anticipate potential risks and hazards, allowing them to take corrective action if necessary. Effective situation awareness also allows decision-makers to react appropriately and adapt to changes in a dynamic environment.

In this work, we present an innovative approach to situation awareness and subsequent support for decision making based on exploiting recent advances in extended reality (XR) and artificial intelligence (AI). The rise of XR technologies is gradually impacting various industries, including gaming, entertainment, healthcare, education, and more. XR is set to revolutionize how people experience both the virtual and real worlds by offering users a realistic and immersive experience. Augmented, virtual, and mixed reality have enabled users to interact with digital content in new and innovative ways.

The main objective of this work is to enhance the situation awareness of the first responders dealing with a flood incident and the journalists preparing for media production. The proposed solution, which is developed in the framework of the XR4DRAMA H2020 EU-funded project (https://xr4drama.eu/ (accessed on 28 May 2023)) (and from now on will be referred to as the XR4DRAMA system), uses XR technologies to (a) support the acquisition and delivery of all pertinent information required to organize appropriate tasks; (b) generate a realistic experience of a remote area (using, for example, virtual reality) in order to anticipate a scheduled event or occurrence as comprehensively as possible; (c) generate a shared projection of an environment (a "distributed situation awareness") for team members, hence facilitating communication and collaboration; and (d) provide the ability to continuously modify the projection as an incident unfolds or as information is updated in order to re-assess the course of action.

The rest of the paper is organized as follows. Section 2 describes similar efforts that have been conducted in the past. Section 3 presents the architecture and the methodological approach of the proposed system. Next, Section 4 shows the software and hardware tools that first responders and journalists use to increase situation awareness, whereas Section 5 details the intelligent modules that operate in the background of the system to generate relevant content. After that, Section 6 explains the evaluation setup and reports its results, and finally, Section 7 summarizes the conclusions of the paper.

2. Related Work

Over the past few years, VR and AR have been gaining popularity as domains, with increasing applications in various fields and a growing market value (https://www.grandviewresearch.com/industry-analysis/virtual-reality-vr-market (accessed on 28 May 2023), https://www.grandviewresearch.com/industry-analysis/augmented-reality-market (accessed on 28 May 2023)). These technologies are also being actively researched, further contributing to their advancement.

In particular, Shin et al. [1] proposed a light waveguide LCD based on the flexo-electric effect, achieving ultrahigh transmittance and a high contrast ratio for transparent displays. Another study [2] presented a new optical system for VR displays, achieving compactness, light weight, and high light efficiency. Additionally, a novel approach [3] for achieving efficient micrometer-scale red-emitting nanowire LEDs was proposed, leading to higher external quantum and wall-plug efficiency for submicrometer-scale red LEDs. These advancements hold the potential to improve VR and AR displays in a broad range of applications.

In the field of disaster management, a lot of efforts have been made to leverage innovative interactive technologies. Training software and simulations were extensively used to achieve a better preparedness of professionals and citizens; however, the training was typically offered through traditional screen-based devices such as computers, monitors, or gaming consoles. Serious games have also become popular training and behavioral analysis tools, and examples are present across industries, for example, terrorist attacks [4] and earthquake evacuation [5].

With the advent of XR, a lot of disaster management research has shifted to these types of solutions as they have the possibility to fully immerse the users and unlock more advanced capabilities for increased situation awareness. For example, a study [6] by the department of geography at Simon Fraser University, Canada, showcases the use of XR for flood risk management. A mixed reality (MR) tool simulates a physical environment

along the Fraser River, allowing users to visualize flood risk zones before floods occur. This example utilizes the Unity software development engine and the HoloLens MR device. In [7], an augmented reality (AR) application is proposed for on-site content authoring and flood visualization. It complements existing flood management tools and combines real-time population of building models, interactive flood visualization, and live sensor readings, such as water level, humidity, and soil moisture. Serious games and immersive virtual reality (VR) are combined in the Flood Action VR framework [8], which exploits real-time and historical weather, disaster, and geographic data and is designed to promote public awareness about disasters, train first responders, and support the decisions of decision makers and scientists. Itamiya et al. [9] reviewed XR solutions for disaster education and found that people with local knowledge can experience the dangers during a disaster through immersive spatial environments. Two XR solutions were proposed for flood disasters: a smartphone-based AR app for tsunami education and an AR app using Unity, ARKit, and ARFoundation to display floodwaters and smoke on a 3D landscape. The latter was used in evacuation drills and raised participants' situation awareness.

All of the previously described disaster management solutions support a restricted scope of the disaster management process. On the other hand, our suggested approach provides a comprehensive toolset that can be used both before and during a catastrophe, facilitates the work carried out in both the control room and in the field, and involves in an integrated way both the experts dealing with a flood and the affected population.

Ever since the new wave of XR tools emerged, media organizations have been experimenting with the technology. For example, the Fader Enterprise project [10] was about building an interactive, browser-based 360/VR storytelling platform, V4Design [11,12] aimed at reusing/recycling content for 3D/immersive worlds, and MediaVerse [13] has also included immersive aspects in its quest to give users access to a wealth of media assets and co-creation opportunities. There were also smaller lab or hackathon projects such as Hidden Stories (https://innovation.dw.com/articles/hackathon-visualizing-global-inequalities-heres-what-we-came-up-with (accessed on 28 May 2023)) (in situ AR storytelling and explainer journalism), Deutsche Welle 360 World Heritage (https://www.dw.com/de/360-app-spielen-und-deutsches-welterbe-erleben/a-47711548 (accessed on 28 May 2023)) (interactive/immersive detective game featuring UNESCO sites), or ARticle (https://www.dw.com/en/water-footprint-augmented-reality-shows-water-cost-of-beef/a-52218248 (accessed on 28 May 2023)). The authors of [14] examined the effects of immersive VR journalism using a unified framework and design guidelines that can help designers create news products with better user experience and media effects. They applied this method to an immersive VR news report on "Beijing: Preventative Action for Severe Acute Respiratory Syndrome (SARS)" and evaluated its user experience and media effects through user feedback in two different storytelling formats: VR news without interaction and VR news with interaction. Their findings indicate that interactivity can significantly direct the user's attention towards important information that a news reporter intends to convey, resulting in a significant positive impact on the media effects of VR news. Additionally, interactivity can transform the user from a passive observer watching news from an omniscient perspective to an active participant experiencing the emotions of the parties involved in the news story. The common denominator of all these research and development (R&D) efforts is the desire to do two things: (a) conceive and tell stories in XR and/or (b) create tools to make this kind of XR storytelling easier and cheaper.

On the other hand, our proposed solution uses XR (along with various other innovative technologies) to achieve a different purpose. It is a tool for planning and supervising, a part of the management infrastructure. The XR4DRAMA system does not simply aim to make journalism itself more immersive but rather offers journalists immersion—and situation awareness—so they are better prepared and save time when completing a job in the field. Its uniqueness lies in the fact that it provides a comprehensive range of services (besides XR) for media people organizing their next production assignment, such as map/street

view/Earth view/navigation tools, audiovisual archives (including 3D model and 360 photo data), and task and team management features.

3. Framework

Figure 1 illustrates the high-level architecture of the proposed XR system. The XR4DRAMA system consists of four software tools that communicate and exchange data with each other in order to support the interaction requirements of the professionals in a team and the citizens affected by an incident. These tools are categorized according to user type. In particular, the professional groups are the ones who work in the control room and remotely manage a flood response or a media production operation and the ones that carry out their tasks in the field, such as the first responders. The authoring and collaborative VR tools are designed for the former group, whereas the AR application is designed for the latter group. For situations where citizens are actively involved, the awareness mobile application is conceived as the communication point between them and the professionals. Additionally, the workers in the field are equipped with a smart vest system that tracks their physiological signals and enables continuous monitoring of their status by the entire team.
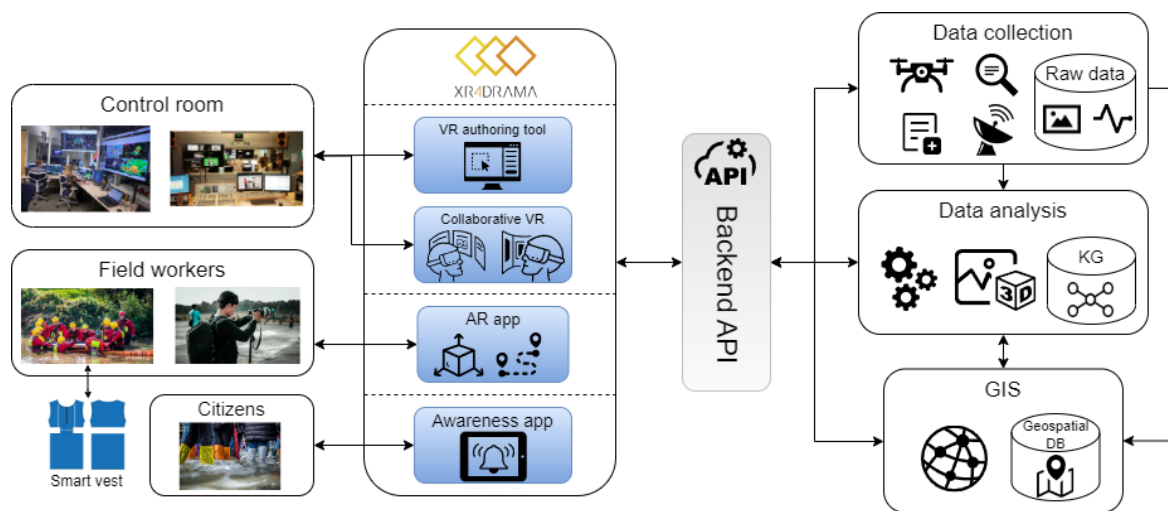

Figure 1. The XR4DRAMA system's high-level architecture.

All the end-user software tools are supported by a set of technologies that work in the back-end and are accessed through a centralized interface, the back-end's application programming interface (API). The back-end consists of a multifaceted module set that is grouped into three main tiers: (a) the data collection tier that deals with the acquisition of data from various sources, such as the smart vest, the user-generated multimedia, as well as content obtained from social media, the open Web, and satellite services; (b) the data analysis tier that processes the raw data from different modalities (text, audio, image, video) to generate more knowledge and create realistic 3D model representations; and (c) the geographic information system (GIS) tier that makes use of a geospatial database with 2D and 3D content to allow all relevant data to be suitably placed in maps and in 3D space.

The results shown in the XR4DRAMA platform are dependent on two different parameters, the input data and the analysis modules, as shown in Table 1 and detailed in the following sections.

The architecture of the back-end modules adopts a microservice approach consisting of a modular collection of services communicating asynchronously. Each service is associated with each one of the innovations highlighted and presented individually in Section 5. This way, the services are designed to be independent and self-contained, and failures in one service do not necessarily affect the availability or performance of other services. This makes it easier to isolate and recover from failures, which can improve the overall resilience and availability of the system.

Table 1. Input data and analysis results.

Input Data Sources	Analysis Modules
Smart sensors (Heart rate, Breath rate, Inertial measurement unit)	Visual analysis
Web content	Audio analysis
Social media posts	Text analysis
Satellite data	Stress level detection (Fusion from sensor-based and audio-based data)
GIS services	Semantic integration
POIs and tasks created/edited by the professionals (in the control room and in the field)	Text generation
Multimedia uploaded by the professionals	3D reconstruction
Citizen reports (audio, text, photo, video)	GIS functionalities (e.g., navigation)

Table 2 depicts the minimum hardware requirements for all four software tools and the system's back-end.

Table 2. Minimum requirements for all tools.

	Authoring Tool	VR Collaborative Tool	AR Application	Citizen Awareness App	Backend
Operating System	Windows 10	Windows 10	Android 8.0 and newer or iOS 14 and newer	Android 6.0 and newer	Windows 10 or Linux
CPU model	Intel i5/i7/i9	Intel i7/i9	Snapdragon 636 or higher	Snapdragon 636 or higher	Any model with 1.8 GHz or higher and at least 4 threads
RAM	16 GB or higher	32 GB or higher	3 GB or higher	3 GB or higher	16 GB or higher
Hard Disk Drive	Minimum 5 GB	Minimum 5 GB	Minimum 200 MB	Minimum 50 MB	Minimum 100 GB
Graphics Card	Dedicated card	NVIDIA GTX 2080 or above	Not applicable	Not applicable	NVIDIA GPU with at least 8 GB VRAM
VR Headset	Not applicable	HTC Vive	Not applicable	Not applicable	Not applicable
Internet connection	YES	YES	YES	YES	YES
ARCore *	Not applicable	Not applicable	YES	Not applicable	Not applicable
DepthAPI	Not applicable	Not applicable	YES	Not applicable	Not applicable

* List of applicable devices: https://developers.google.com/ar/devices (accessed on 28 May 2023).

The supported functionalities of XR4DRAMA contribute to three levels of situation awareness (SA) that will build on each other:

- Level 1 SA: An appealing, easy-to-digest representation and visualization of different content and information about a location, such as geographical data, sociographic information, cultural context, and images and videos from the Web. It is important to note that all the data presented at this point are retrieved and visualized automatically by the system.
- Level 2 SA: An enhanced representation and visualization that especially builds on recent content and information originating from people in the field (first responders or location scouts) who use a variety of tools and/or sensors in order to capture data that are most relevant to the individual use case scenario. In other words, the system processes these data "from the field" and combines them with the data about the

location that were already gathered remotely from accessible web and cloud services. This representation is available to remote management in a distant control room as well as to staff in the field.

- Level 3 SA: This level denotes a more sophisticated, complex, and comprehensive virtual representation of a particular location, similar to a simulation of an event occurring in that environment. Here, users leverage rather mature XR representations that can be perceived by sophisticated tools, such as VR/XR head-mounted displays, which contribute to a higher level of (relevant) immersion and thus better situation awareness. The primary distinction between Level 3 and Level 2 SA is the ability to interactively test specific strategies and methods, e.g., simulating various possible camera movements in the media production use case.

The XR4DRAMA system is implemented using two software development cycles that follow the V-model [15] specifications. Each cycle consists of the following phases: (a) user requirements analysis, (b) system/technical requirements analysis, (c) module design and implementation, and (d) integration. Each phase is mapped to a dedicated testing session, with the pilot tests reported in Section 6 focusing on the user requirements aspect.

In the following sections, each individual XR4DRAMA system front-end tool (user application) and back-end module is described in further detail.

4. End-User Tools

4.1. Software

4.1.1. Authoring Tool

The authoring tool is a desktop application that is the entry point for professionals on the XR4DRAMA platform. Using this tool, a team in the control room can create, view, and manage projects (projects are associated with missions, flood management, or media production planning) that are characterized mainly by their exact location, the participants, and the start/end dates (Figure 2).
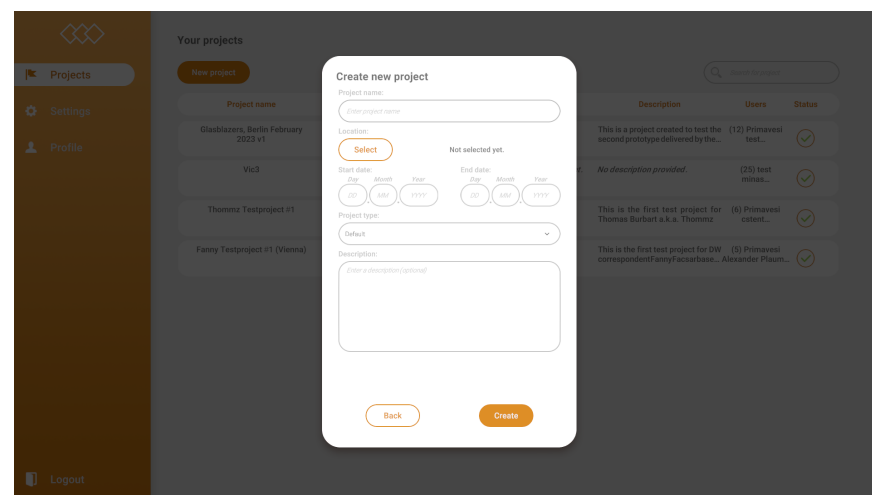

Figure 2. Authoring tool. Project creation panel.

The tool provides and enhances situation awareness for those who are geographically far from the real site of a project in many different ways. Inside the project view, one can visualize different map views of the location, including all the points of interest (POIs) and regions of interest (ROIs), contacts, and media (e.g., 3D models, 360 images, legal documents, etc.) served by the back-end of the platform. The control room team has full control over the POIs; they can add, delete, and edit their attributes and attach media files to them. Moreover, they can create and assign tasks to any team member working in the field to gather further information related to their objectives. They can also upload media files either to be analyzed for additional metadata (which facilitates searching for these

files according to the detected attributes) or to be sent for the reconstruction of realistic 3D models.

With regard to the flood management use case, the authoring tool provides additional functionalities which are necessary for the real-time monitoring of the dynamic situation and the status of the first responders in the field. In particular, it visualizes satellite, raster, and digital elevation model (DEM) data and water levels retrieved by targeted measurement stations. It highlights danger zones that should be avoided by the public and treated with extreme caution by the professionals, and it illustrates the stress levels and location of the first responders (Figures 3 and 4). In addition, from the authoring tool, one can broadcast warning messages to all citizens that use the citizen app of the XR4DRAMA platform. Lastly, all citizen reports that are recognized by the platform as being associated with an emergency are presented as POIs, which include the analysis results from the back-end of the platform. The POIs created by citizen reports automatically modify the area of the danger zones accordingly.
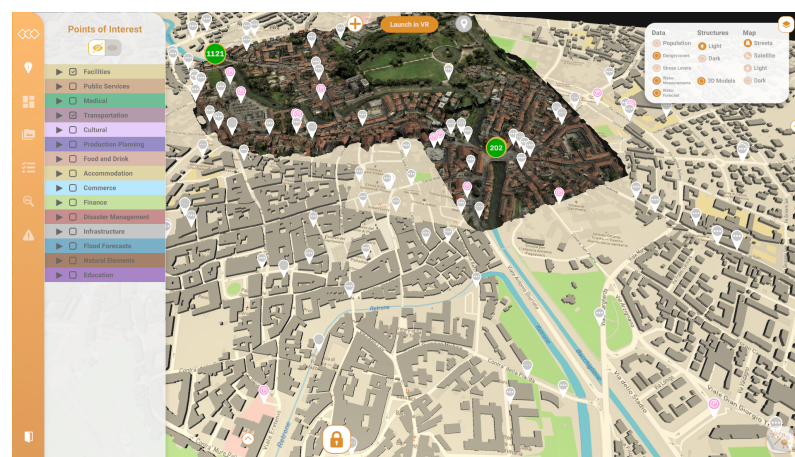

Figure 3. Authoring tool. Map view with POIs, 3D models, and water level sensor measurements.
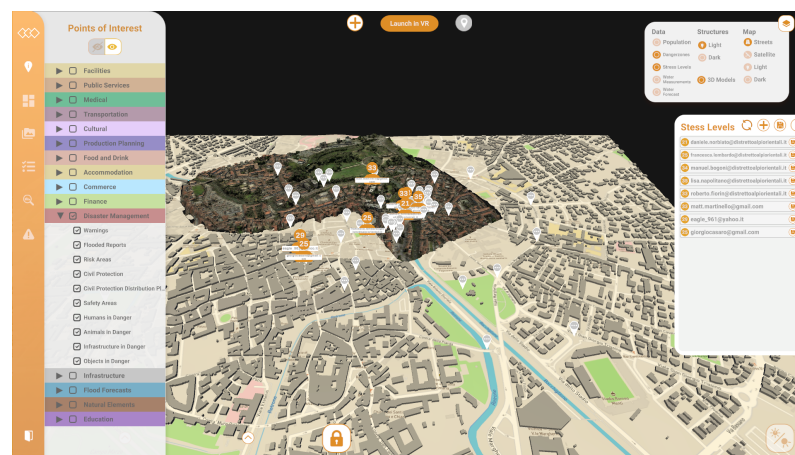

Figure 4. Authoring tool. Stress level monitoring of first responders.

4.1.2. Collaborative VR Tool

The collaborative VR tool is launched through the authoring tool and allows users to view the location in an immersive environment. It enables people to visit the target area of a planned operation in Virtual Reality to enhance their situation awareness, enabling them to assess the plan's viability and identify missing elements or previously overlooked possibilities, such as desirable shot angles or convenient vantage points. The way in which the environment is examined can be altered (based on the available data sources) to suit user preferences. It can be observed from the (virtual) ground level or from an aerial perspective, a bird's eye view (Figure 5).

Although its primary source of input is the authoring tool, it also incorporates additional data sources from the platform, such as POIs and reconstructed 3D models (instead of broad building outlines), which update with the generation of new data.

Several VR users may access the environment simultaneously, and voice communication and gesture synchronization are supported, enabling verbal and gestural communication and collaborative modification and interaction with the plan's components (such as the editing of POIs). Furthermore, users may explore the area as if they were physically there, e.g., by visiting potential target places, attempting drone flights for planned video recordings, etc.

This application can be used with any PC-based VR headset that supports OpenVR or OpenXR. It offers a low synchronization latency of 100 ms by using an architecture that is optimized for swift and smooth data transfer and allows for flexible server locations. Users have the option to co-locate on the same local network or connect from different locations via the internet.
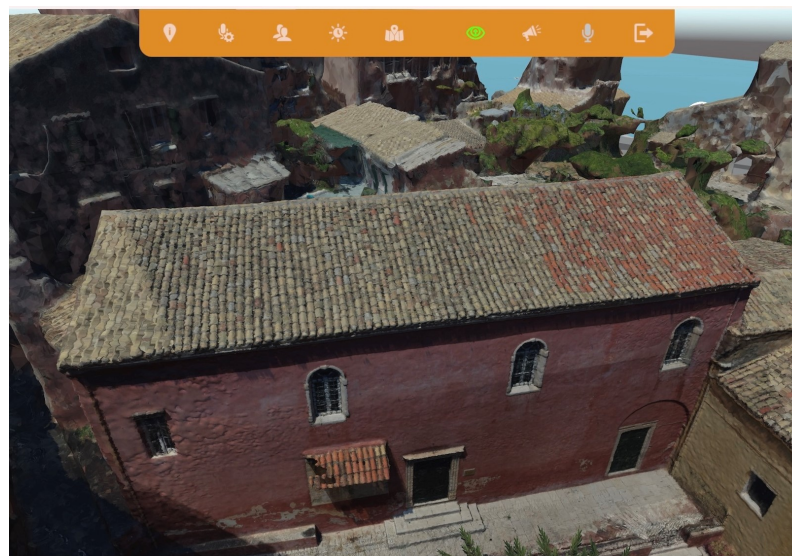

Figure 5. The 3D model view in VR (bird's eye view).

4.1.3. Location-Based Augmented Reality Application

The AR application supports location-based data collection and retrieval as well as field task management to facilitate outdoor work. Moreover, it overlays the digital information available on the XR4DRAMA platform on top of the real world, which enhances the situation awareness of the field users, either the first responders or the location scouters and the filming crew. To support the computationally demanding XR features, the application targets high-end mobile phones and tablets operating on both major ecosystems, i.e., the iPhone operating system (iOS) and Android. The AR application is bilaterally and in real-time connected to the platform. Thus, the users obtain the latest information from the platform, and at the same time, the control room receives updates from the field.

To be more specific, once logged in, the users of the AR application can view all projects they have registered for; they can view and edit detailed project information, search and manage its POIs and ROIs as the main data entities, and use case-specific data such as the danger zones and a user's stress level. One can also preview and handle the information that the AI services retrieve from the web. Furthermore, the application offers access to tasks that are created in the control room, including details such as the task creator and assigned actor, headline, description, priority, comments, status, and type. Professionals in the field can process the tasks and complete them by attaching a text message and/or audio report (Figure 6a–d). Furthermore, the AR application offers access to the XR4DRAMA navigation service, which provides a convenient and effective way for users to find their

way between project POIs or arbitrary locations, while ensuring their safety by avoiding routing through danger zones.
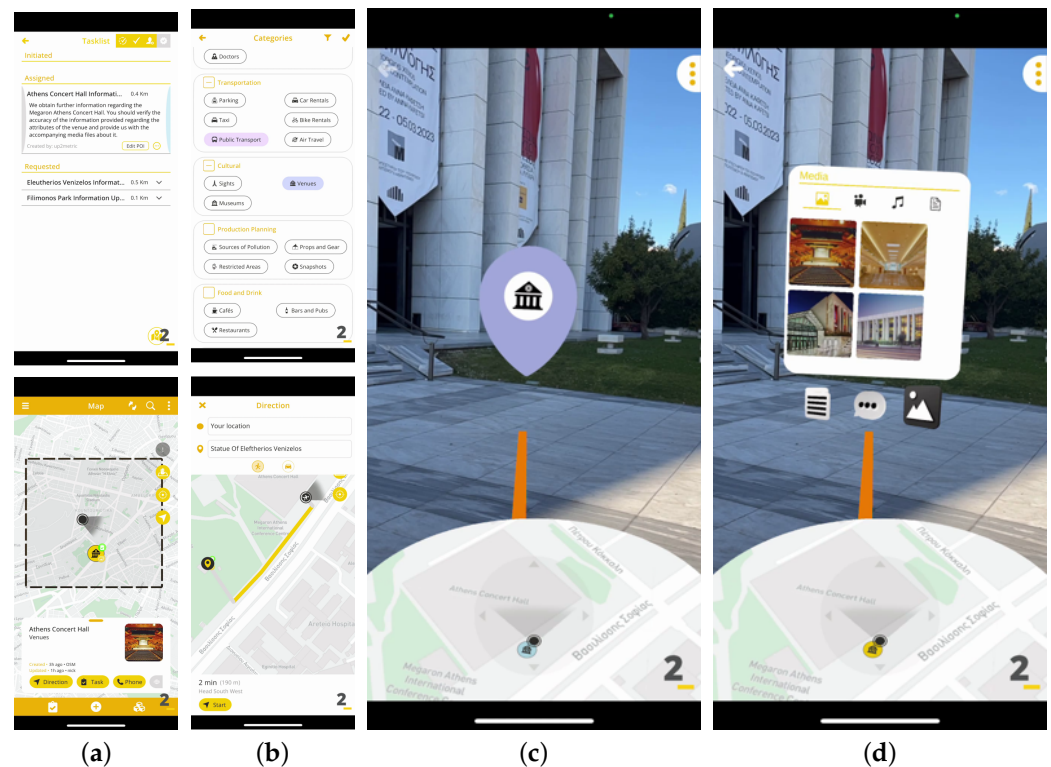

Figure 6. AR application screenshots: (**a**) Task management and a POI with its information panel on the 2D map. (**b**) POI categories and the navigation feature. (**c**) A POI overlaid on the real world and the mini-map navigation map. (**d**) A POI info panel and the the mini-map navigation map.

The most essential feature of the AR application is the AR mode, which supports functionalities that augment the real world. In that mode, they can overlay POIs, navigation to POIs (using virtual lines), flood simulations, and 3D models. An example of a relevant 3D model for floods is the temporary and modular water barriers (also named "aquadikes") illustrated in Figure 7a–c. In Figure 8a–c, different media planning setups, which are uploaded by the journalists on the platform, are displayed in the real world. A measuring tool is also available that allows the user to measure distances in the real world by creating measuring lines with multiple points.

4.1.4. Citizen Awareness App

The citizen awareness app is a mobile application designed for the wider public. This application aims to create the habit of a citizen becoming cautious about a possibly harmful situation and facilitate the authorities' jobs by acting according to their plans for flood management. When a catastrophic event takes place, the affected people seek help using traditional means such as calling local emergency numbers or looking for information online from civil protection and TV news. While this approach is standard practice, it can lead to delays and inconsistencies because an emergency request is often approximately described, and critical information such as the exact global positioning system (GPS) location or severity of the situation is missing in the description. A crucial function of the citizen awareness application is that it enables the user to report an ongoing situation or make an emergency geolocated request via text, audio, image, or video directly to the authorities that handle flood management. The emergency report is analyzed by one of the analysis modules of XR4DRAMA, which speeds up the decision-making process of

the authorities managing the flood disaster. Figure 9a shows how the main screen of the application looks, and Figure 9b shows an example of the emergency report using an image.
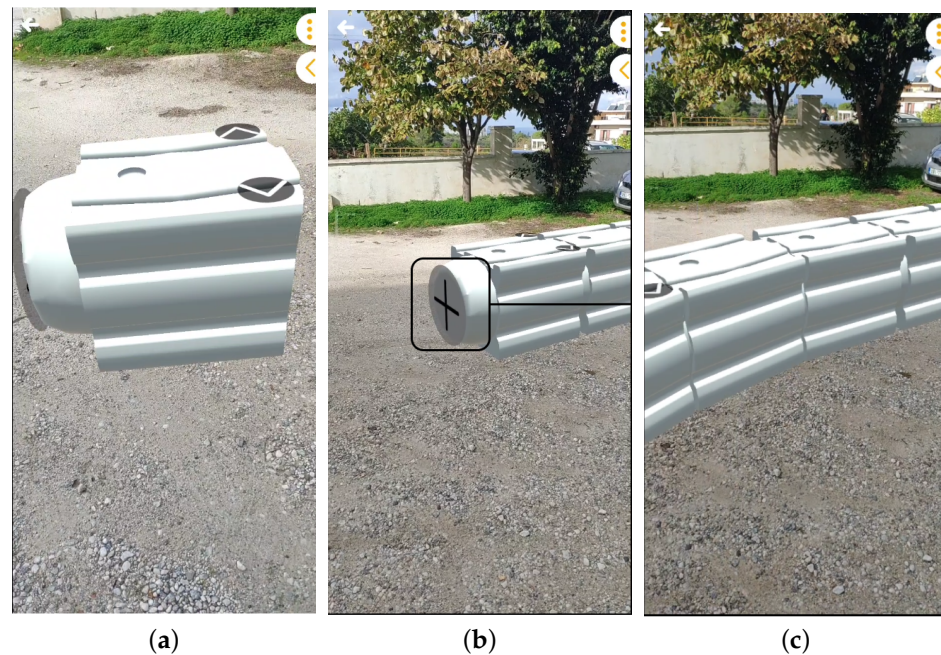

Figure 7. AR application screenshots: Placing, in the AR application, aquadikes for the flood management pre-event. (**a**) presents the initial aquadike (digital object) placed in the real world; the arrows on the aquadike facilitate the object placement. (**b**) shows how the interactable object can be reproduced (plus sign) in the scene to produce a chain of aquadikes, as in picture (**c**).
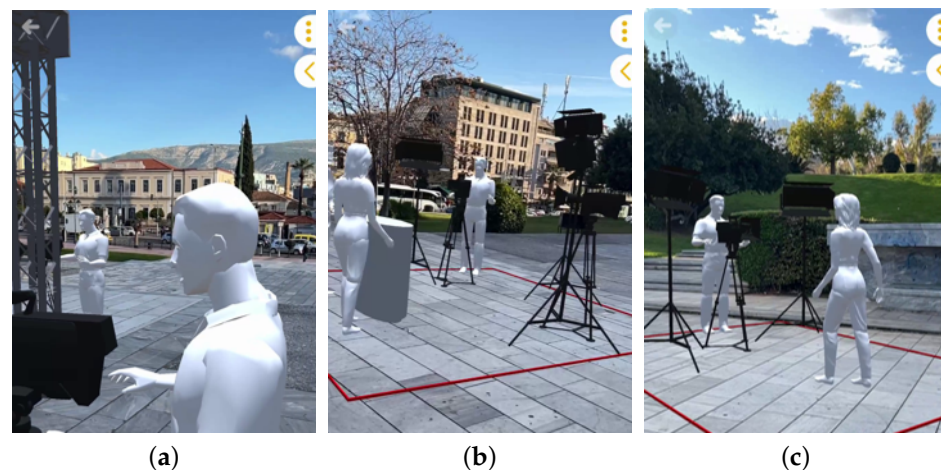

Figure 8. AR application screenshots: (**a**–**c**) present various media planning setups embedded in the real world via the AR application, as retrieved from the XR4DRAMA platform.

Furthermore, the citizen awareness application detects the context of the user based on location, time, and proximity to the event and accordingly informs and alarms the user about likely threats, from location-based threats to evacuation alerts and more. The application map includes information generated by the authorities during crisis management, such as disaster management-related POIs (e.g., safety areas, flood reports, danger zones, civil protection distribution places, etc.), thus effectively improving the situational awareness of the user. Figure 9c shows an example of a text notification created by the authorities and received on the application's device. Figure 9d illustrates the application map with

the information registered by the authorities, such as disaster management POIs and danger zones.
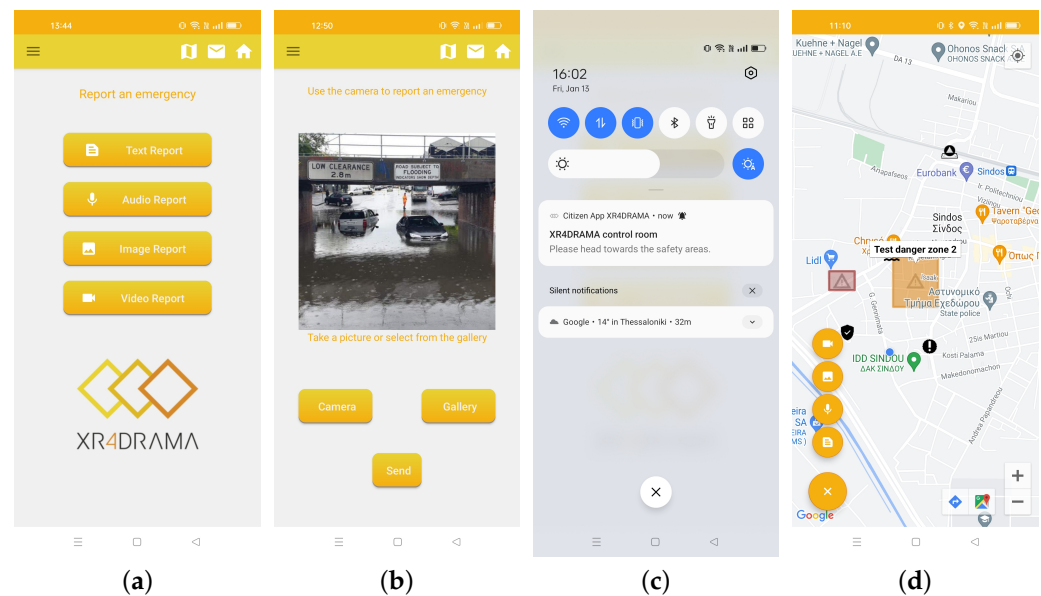

Figure 9. (**a**) Citizen awareness app main screen. (**b**) Emergency report using an image. (**c**) Text notifications from the authorities. (**d**) Application map.

4.2. Smart Sensing Device

One of the key aspects of disaster management is ensuring the safety and protection of the volunteers that work in the field. The organization managing the disastrous event has to provide its first responders with personal protective equipment suitable for the tasks that they are required to perform. In the context of XR4DRAMA, a critical part of this equipment is a wearable smart vest equipped with an electrocardiograph (ECG) sensor, a respiration (RSP) sensor, and an Inertial Measurement Unit (IMU) with an accelerometer, a gyroscope, a magnetometer, and quaternion sensors that are used to continuously monitor the condition of the workers in the field.

The physiological data acquired using this smart sensing device are based on textile sensors fully integrated into the vest layer in contact with the skin and a data logger that can record and process data on board and transmit them via Bluetooth 2.1 for the stress detection module of XR4DRAMA. The wearable sensing device is shown in Figure 10. The system includes:

- Two textile electrodes to acquire an ECG signal;
- One textile respiratory movement sensor;
- A portable electronic (called a RUSA Device) in which an inertial platform (accelerometers, gyro, and IMU sensors) are integrated in order to acquire trunk movements and posture;
- One jack connector to plug the garment into the electronic device;
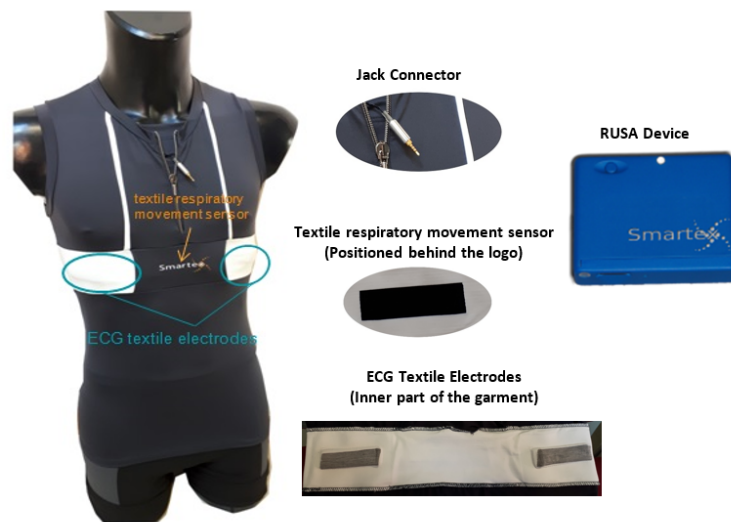- A pocket to hold the electronic device during the activity.

Figure 10. Smart sensing device architecture.

5. Core Technologies

5.1. Data Collection

5.1.1. Web and Social Media

The Web and social media contain a huge amount of structured and unstructured data that, when properly analyzed and mined, become a valuable source of insight and intelligence. However, with the immense proliferation of information and the profusion of noise, the analysis and indexing of such data constitute a highly challenging task, and sophisticated data acquisition algorithms and implementations are needed. With ever-expanding data divided into different formats, multiple codes and languages, and various categories, interconnected in no particular order, an intelligent data collection mechanism tailored to the system requirements is necessary to tackle these challenges.

For the XR4DRAMA system, we developed a module that utilizes filters to collect only subsets of interest in the relevant information. Such filters are related to specific multimedia items, i.e., Twitter posts. Specifically, three sources of data are considered: traditional websites, social media platforms, and open repositories that gain access using an API.

The goal of the web and social media data collection module is to collect useful multimedia content (e.g., text, images, videos) that can be then processed by the analysis modules. The integrated data collection process consists of four phases, each incorporating different techniques depending on the use case:

- Phase 1—Requirements: The input for the module is established during this phase. The requirements for data collection can be defined either as keywords forming textual queries or as URL addresses.
- Phase 2—Discovery: In situations where the exact web resources to be gathered are not known, discovery has to be conducted. Detection of relevant web resources can be achieved in three ways: (a) using web crawling, when the input is an entry URL; (b) using search, when the input is a textual query; and (c) streaming in real-time, instead of searching, where an existing API (e.g., Twitter) provides such capability. This phase is bypassed when the indicated URL resource in Phase 1 is the only one to be integrated. Web crawling uses a breadth-first algorithm to iteratively traverse webpages (starting from the entry URL) and extract a list of discovered unique URLs. The process is executed until a crawling depth of 3 is reached (i.e., the maximum distance allowed between the discovered and entry pages). Searching and streaming build upon existing third-party APIs to discover related content.
- Phase 3—Content Extraction: This procedure is straightforward when data collection is conducted using existing APIs, as it corresponds to a simple retrieval action. However,

if there is no available API, web scraping techniques are applied using the Boilerpipe library [16] that makes use of shallow text features and C4.8 decision tree classifiers to identify and remove the surplus "clutter" (boilerplate, templates) around the main textual content of a web page.

- Phase 4—Storage and Integration: The final stage involves parsing and storing the content that was previously retrieved using a unified representation model that can aggregate different types of multimedia. The base data model is the SIMMO (https://github.com/MKLab-ITI/simmo (accessed on 28 May 2023)) [17] one.

5.1.2. Remote Sensing

The remote sensing service of the XR4DRAMA platform aims to swiftly and automatically provide an initial 2.5D terrain and visual information for large-scale environments, i.e., rendering the control room free of the need to send actors or deploy unmanned aerial vehicles (UAVs) for 3D recording to the field. The service acquires pertinent satellite data, pre-processes them, and reconstructs 2.5D and 3D textured models. In the XR4DRAMA use case, remote sensing data are exploited to develop a rough terrain of the wider project area during the first phase of media production planning or the first hours of a flooding incident. The current implementation offers access to Copernicus open data and is based on the Sentinel hub (https://www.sentinel-hub.com/ (accessed on 28 May 2023)) web service, a platform that provides data time series, products and processing of Earth observation imagery, e.g., from the Sentinel or Landsat constellation. The spatial resolution of the terrain and its dimension (2.5D or 3D) are determined by the availability of the satellite data, either images or even pre-processed digital elevation models (DEM), and they range from 10 m for openly available satellites down to 0.5 m for commercial ones. The XR4DRAMA platform retrieves true-color satellite images (i.e., visible spectral zone), multi-band imagery at 10 m spatial resolution, and DEMs at 15 m resolution. This implementation offers basic functionalities, such as scan requests, downloading specific data, and search for available content, and processing, such as DEM simplification, altitude correction, geographic reference system preparation, and triangular 3D mesh model creation, appropriate for the Unity engine (i.e., for the authoring tool and VR application). Figure 11 shows an example of a 3D textured mesh model derived from a true-color and DEM image (see Section 5.2.6).
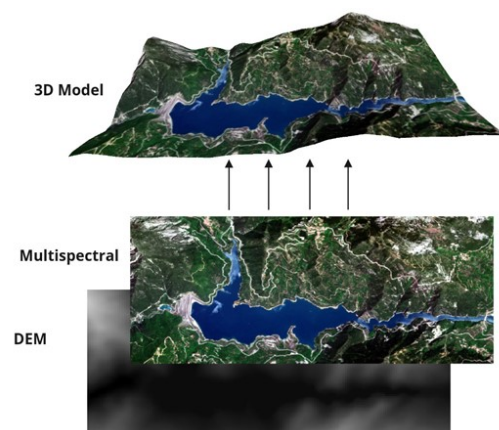
Figure 11. Example of a remote sensing-based 3D model, which exploits a multi-spectral image and a DEM.

5.2. Analysis
5.2.1. Visual Analysis

The visual analysis module processes audiovisual content (from online sources and from the users' input) to localize the exteriors of buildings, the surrounding environment, as well as objects and other valuable assets needed for media production and situation monitoring. The extracted sub-sequences are used to assist in the construction of the 3D

model of an unknown area, including the detected objects of interest. Moreover, it enriches the system's knowledge with semantic properties that characterize an image or video (e.g., the scene type). The high-level workflow of this module, presented in Figure 12, consists of:

- Shot detection: It applies only to videos and automatically detects shot transitions in videos to facilitate next analysis steps. TransNet V2 [18] is deployed for this task, a deep network that reaches state-of-the-art (SoA) performance on respected benchmarks. TransNet V2 builds upon the original TransNet framework, which employs Dilated Deep Convolutional Network (DDCNN) cells. Six DDCNN cells are used, each consisting of four $3 \times 3 \times 3$ convolutions with different dilation rates. The final sixth cell reaches a large receptive field of 97 frames while maintaining an acceptable number of learnable parameters. In the new version, DDCNN cells also integrate batch normalization that stabilizes gradients and adds noise during training. Every second cell contains a skip connection followed by spatial average pooling that reduces the spatial dimension by two.

- Scene recognition (SR): It aims to classify the scenes depicted in images or videos into 99 indoor and outdoor scene categories using the VGG16 [19] Convolutional Neural Network (CNN) framework. VGG16 consists of 16 layers, not including the max-pool layers and the SoftMax activation in the last layer. In particular, the image is passed through a stack of convolutional layers, which are used with filters of a small receptive field. Spatial pooling is carried out by five max-pooling layers, which follow some of the convolutional layers. The width of convolutional layers starts at 64 in the first layers and then increases by a factor of 2 after each max-pooling layer, until it reaches 512, as depicted in Figure 13. The stack of convolutional layers is followed by three fully connected (FC) layers. The final layer is the SoftMax layer. The first 14 layers of the VGG16 framework were pre-trained on the Places dataset (http://places2.csail.mit.edu/ (accessed on 28 May 2023)), which has the initial 365 Places categories. The remaining layers were trained on a subset of 99 selected classes of Places dataset.

- Emergency classification (EmC): This component recognizes emergency situations, such as floods or fires, in images or videos. It uses the same CNN framework as the scene recognition module with the difference that the last fully connected layer is changed to fit the emergency classification purposes. Specifically, the final fully connected (FC) layer was removed and replaced with a new FC layer with a width of three nodes, freezing the weights up to the previous layer. A softmax classifier was also deployed so as to enable multi-class recognition (where classes are "Flood", "Fire", and "Other").

- Photorealistic style transfer: A U-Net-based network [20] is paired with wavelet transforms and Adaptive Instance Normalization to generate new images that look like they are in different lighting, times of day, or weather. The result enhances the performance of the building and object localization module for challenging images featuring poor lighting and weather conditions. The implemented network is comprised of the following three sub-components: encoder, enhancer, and decoder. The encoder has convolutional layers and downsampling dense blocks. In the latter, two convolutional layers, a dense block and a pooling layer based on Haar wavelets, are included. The decoder has four upsampling blocks creating a U-Net structure in addition to the encoder previously mentioned. An upsampling layer, three convolutional layers, and a concatenation operation make up each block. In the enhancer module, multi-scale features are transmitted from the encoder to the decoder.

- Building and object localization: This component localizes buildings and objects of interest and is based on DeepLabV3+ [21] that allows for segmenting images and visual content in general. DeepLabV3+ extends DeepLabV3 [22] by employing an encoder–decoder structure. The encoder module encodes multi-scale contextual information by applying atrous convolution at multiple scales, while the simple yet effective decoder module refines the segmentation results along object boundaries.

Furthermore, a second image segmentation model is deployed for emergency localization. The model is also based on the DeepLabV3+ architecture and was trained on the ADE20K dataset [23]. Specifically, it localizes "water" or "flame" regions and, in combination with the localization of people or vehicles, it determines the exact presence of people or vehicles in danger.
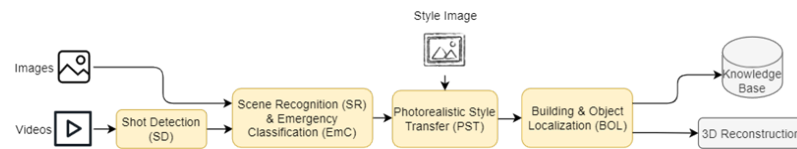

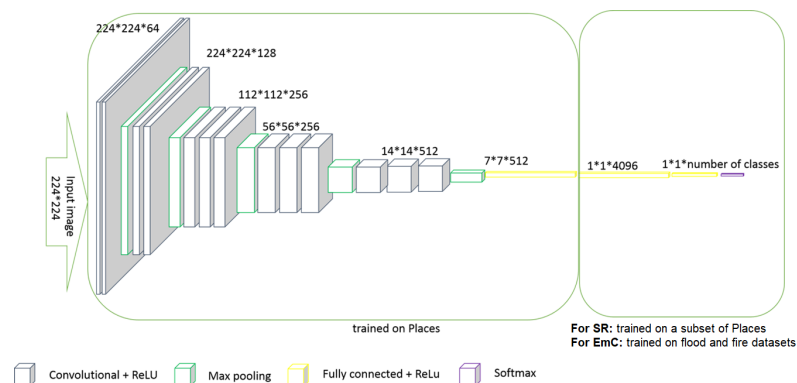**Figure 12.** The visual analysis workflow.

Figure 13. The VGG-16 architecture used for SR and EmC models.

Another visual analysis pipeline tailored to the flood management use case makes use of live streams from static surveillance cameras for the estimation of rivers' water levels and the detection of overtopping, which is crucial for the detection of flood situations. The canny edge detection algorithm [24] is used for the detection of a rod that is placed in the river and acts as a marker of known length. This algorithm applies Gaussian smoothing to reduce noise in the image and detects edges by computing gradient magnitude, applying non-maximum suppression to remove unwanted pixels and utilizing hysteresis thresholding. After the marker detection, the system calculates the distance in pixels between the marker's highest and lowest detected points (that should mark the surface of the water). Three types of alerts ("Moderate", "Severe", and "Extreme") are generated if the water level exceeds predefined thresholds.

5.2.2. Multilingual Audio and Text Analysis

The multilingual audio and text analysis module processes the raw material acquired from the web and social media data collection module or the material provided by the users for the derivation of abstract linguistic representations. These linguistic representations are used by the language generation component that will export information relevant to the users' needs and requirements. Regarding multilingual audio inputs, the Wav2Letter++ [25] automatic speech recognition (ASR) tool is used as a first step for transcribing natural spoken language into text, whereas for textual inputs, the language analysis module is directly executed. Language analysis is a complex task that requires combining various components in a series of steps, going from low-level linguistic analysis, e.g., tokenization, through higher levels of linguistic complexity, e.g., dependency parsing, to the extraction of entity–relation–entity triples. Therefore, the language analysis module implements the following algorithmic components:

- Concept extraction to translate an input sentence into concepts employing an adapted version of the pointer-generator model [26,27] (using separate distributions for copying attention and general attention) and the use of multi-layer long short-term memory networks (LSTMs), whose network architecture is described in [26]:

- Named entity extraction using spaCy 2.0 (http://spacy.oi (accessed on 28 May 2023)) to annotate 18 different entity types;
- Temporal expression identification with HeidelTime [28], a rule-based multilingual, domain-sensitive temporal tagger;
- Entity and word sense disambiguation by collecting and ranking candidate meanings of the words in the text transcript, calculating the salience of a meaning with respect to the whole set of candidate meanings, and determining its plausibility with respect to the context of its mention in the input texts [29,30]. Meanings are compared with each other using sense embeddings and with their context using sentence embeddings calculated from their English Wikipedia and WordNet glosses;
- Geolocation using a two-stage procedure: first, location candidate identification, and second, the search of normalized surface forms of detected candidates in a key-value index. We make use of two geographical databases, OpenStreetMap (https://www.openstreetmap.org/ (accessed on 28 May 2023)) (OSM) and GeoNames (https://www.geonames.org/ (accessed on 28 May 2023)), which we convert into a direct and inverted search index and extend with possible shortened versions of names. To reduce the number of candidates, we pre-filter the search index according to the perimeter of the geographical area of application;
- Surface language analysis via UDPipe (https://github.com/ufal/udpipe (accessed on 28 May 2023)) [31] which performs part-of-speech (POS) tagging, lemmatization, and parsing;
- Semantic parsing to output the semantic structures at the levels of deep-syntactic (or shallow-semantic) structures and semantic structures [32].

The main domain-specific information extraction part is then performed through the application of grammars developed using the MATE framework [33].

The entire analysis pipeline incorporating all these components is built using UIMA (https://uima.apache.org/ (accessed on 28 May 2023)) and DKPro (https://dkpro.github.io/ (accessed on 28 May 2023)).

5.2.3. Stress Detection

The stress detection module processes the physiological signals of workers in the field measured by the smart vest they are wearing and the audio messages recorded during their communications with the control room. These signals go through the stress detection module of XR4DRAMA, measuring their stress levels during the operation, thus monitoring their condition and raising situational awareness.

There are two separate stress detection methods employed for each specific source of data, one for the physiological signals and one for the audio messages. Nevertheless, a fusion stress detection method that works when both sources of information are available is also considered. Physiological signals are measured at all times, while audio signals are produced only during communications. If there is no audio signal, only the sensor-based results are considered. All three developed stress detection methods estimate stress levels in real-time and return a value between $[0, 100]$, with higher values indicating more stressful conditions. The individual stress level results for each worker in the field can be seen in the AR application. The list of results for all workers in the field is also visible at all times in the control room on the authoring tool screen, so the conditions of the people involved are continuously monitored. The methods were trained on a custom dataset comprising participants of different genders who wore the smart vest and had a microphone on them. The dataset was designed to simulate challenges ranging from calmness to high stress. After each challenge, the users were requested to rate their stress levels on a scale of 0 to 100. The details of the three developed methods are described below:

- Physiological signals stress detection: Data gathered from the sensors are analyzed in order to extract statistical and frequency features that are fed into a trained model for the final continuous value of stress level detection. We deployed the same developed method described in [34]. In particular, a total of 314 features were extracted, consisting

of 94 ECG, 28 RSP, and 192 IMU (16 per single-axis data) features. ECG features encompass statistical and frequency features related to the signal, R-R intervals, and heart rate (HR) variability. The hrv-analysis (https://pypi.org/project/hrv-analysis/ (accessed on 28 May 2023)) and neurokit toolboxes [35] were utilized to analyze the ECG features. Respiratory features include statistical and frequency features of the signal, breathing rate, respiratory rate variability, and breath-to-breath intervals, extracted using the neurokit toolbox. IMU features consist of basic statistical and frequency features, including mean, median, standard deviation, variance, maximum value, minimum value, interquartile range, skewness, kurtosis, entropy, energy, and five dominant frequencies. After the feature extraction step, we utilized the Genetic Algorithm (GA)-based feature selection method [36] and finally deployed the Extreme Gradient Boosting (XGB) tree [37] modified for regression to predict the stress values.

- Audio signals stress detection: A set of acoustic features is extracted from the audio signals using [38], known as the eGeMAPs feature set. The eGeMAPs (extended Geneva Minimalistic Acoustic Parameter Set) feature set is a collection of acoustic features extracted from speech signals. It includes various low-level descriptors (LLD) such as pitch, energy, and spectral features, as well as higher-level descriptors such as prosodic features, voice quality measures, and emotion-related features. These features capture different aspects of the speech signal and can be used for tasks such as speech emotion recognition, speaker identification, and speech synthesis. The extracted features are fed to a Support Vector Machine (SVM) [39] trained for regression.

- Fusion stress detection: The proposed fusion method is performed at the decision level of both sensor-based and acoustic methods; in other words, it follows a late fusion approach. Both unimodal results are fed into a trained SVM with the radial kernel, which showed the best-performing results during training. The results indicate that the fusion-based method improves the performance of the stress detection task compared to its unimodal counterparts.

5.2.4. Semantic Integration and Decision Support

The semantic integration components represent, aggregate, and interlink the raw and analyzed data from the different components into knowledge graphs.

The main role of this component is to effectively organize and generate knowledge related to disaster management and media planning, with a particular emphasis on (a) reports and incidents (for instance, details derived from images and messages); (b) geographical and time-related information, including coordinates, place names, and timestamps; and (c) activities, tasks, and strategies for disaster management and media planning [40,41]. It maps multimodal information (i.e., textual and visual, including their analysis results, and stress level measurements) into ontology structures, specifically Turtle (https://www.w3.org/TR/turtle/ (accessed on 28 May 2023)) (Terse RDF Triple Language) triple stores, and it automatically creates and updates POIs based on a sophisticated POI management mechanism. In this way, POIs are dynamically enriched (according to their current status) with valuable information such as how many people/objects/animals are in danger, the affected objects, the emergency level and textual comments related to what people (social media users or users of the XR4DRAMA platform) say about the POI.

Additionally, this module is equipped with a decision support mechanism whose contribution is to reinforce the disaster management procedure. Specifically, based on the information in the knowledge graph, and using a rule-based approach, it detects and indicates danger zones and updates their status, including their severity levels (of danger). It also automatically creates tasks for disaster management based on a predefined civil protection plan.

Lastly, an information retrieval mechanism is supported in order to enable the sourcing and filtering of the content from the text generation component (described below) according to flexible criteria.

5.2.5. Text Generation

The text generation module is a Natural Language Generation system that satisfies the users' needs regarding the verbalization of information in the form of multilingual reports.

For disaster management, the module generates a title and a description of the emergency situation that caused the creation of the POI (using the semantic integration module) and creates a timeline report conveying a slightly shorter description of the emergency situations that occurred during that period, ordered by the time at which they happened.

For media production, data available through web crawling of sites like Wikipedia or OpenStreetMap, as well as text analysis of comment messages sent by citizens through FourSquare are exploited. The output is a three-section report consisting of the introductory paragraph of the location's Wikipedia page (with a direct link to the source material), a brief summary with weather information in textual or visual form (also including the source's link), and a list of sentences regarding the quantity and type of POIs already stored on the platform server.

The text generation process is divided into three subtasks: content selection, discourse planning, and linguistic generation. This division of text generation into specific tasks conforms to the ideas of the Meaning–Text Theory [42], which advocates for a precise and independent modeling of each level of language description (semantics, syntax, topology, morphology). The generation is performed in successive steps by mapping one level of representation onto the next, and the two submodules work together.

Content selection involves the decision of what information coming from the knowledge base of XR4DRAMA needs to appear verbally in the generated reports based on the user's needs. The role of discourse planning is to discursively organize the contents of the inputs to be generated, in particular to establish discourse relations between utterances to be produced by the agent. Linguistic generation converts the contents into well-formed text. Discourse planning (also referred to as "sentence packaging", "text planning", or simply "aggregation") and linguistic generation were tackled by improving UPF's grammar-based generator, FORGe [43]. The improvement consisted of increasing its coverage both for English and Italian (although a large part of the rule engine is multilingual) by extending our linguistic resources and by experimenting with the implementation of a complementary module using neural networks. The generator consists of 2483 rules across the two languages, covering seven layers of transformations starting from structured information as an input, through different abstractions representing semantic and linguistic information, to plain text made for human readers as the final output. Around 76% of the rules are language-independent and could therefore be applied to other languages beyond English and Italian. The others deal with language-specific linguistic phenomena and thus need to be adapted to each individual language.

Additionally, neural text paraphrasing was applied to the outputs of the FORGe system mentioned above in order to improve the quality of the generated texts by using a hybrid approach: adding more fluency through the neural paraphrasing tool without losing the content accuracy that our rule-based system provides. This work (described in [44]) follows the approach introduced by Thompson and Post [45], using a general domain multilingual neural machine translation model but discouraging it from generating n-grams present in the input by down-weighting these tokens without completely disallowing them (i.e., adding a softer constraint), which provides outputs lexically biased away from the inputs, generating non-trivial paraphrases.

5.2.6. 3D Model Reconstruction

The 3D reconstruction service of the XR4DRAMA platform uses sophisticated photogrammetry and computer vision techniques to automatically generate 3D models in the form of textured triangular meshes and serve them to the authoring tool and the collaborative VR application. These 3D models drastically contribute to situation awareness as they represent highly realistic views of real-world objects, buildings, and environments. The service consists of two separate sub-services: the *image-based* 3D reconstruction from a

series of overlapping images recorded from a drone, a handheld camera, or a mobile phone and the *satellite-based* one (see Section 5.1.2). The image-based 3D reconstruction employs a typical approach based on feature extraction and matching, structure-from-motion, multiview stereo-matching, model fitting, and texture mapping [46], and the implementation was based on OpenCV [47], AliceVision framework libraries [43], and Meshroom [48].

In more details, feature extraction involves automatically extracting keypoints that are distinct and invariant to various image distortions and creating descriptors for each feature to uniquely define them and use the descriptors to compare features across images and estimate the homologous points. The SIFT (Scale-Invariant Feature Transform) [49] algorithm is an effective way of detecting and describing keypoints for reliable sparse image matching. Furthermore, this sparse feature matching reveals which images are overlapping, i.e., they depict the same world space. Then, structure-from-motion estimates the camera interior and exterior calibration parameters and a sparse point cloud. The estimation of the cameras' poses usually starts from the pair of images with the largest number of matches in a process known as incremental reconstruction. First, the relative position of the two images is initialized, and their homologous points form the first set of points of the sparse point cloud. Then, new images are added incrementally, and additional tie points are added progressively to the sparse point cloud. Following the pose estimation, a dense cloud is reconstructed by exploiting Multi-View Stereo (MVS) algorithms to accurately describe the scene. In the model fitting step, the triangulation of the model takes place, which involves planar subdivision of the dense cloud into triangles, projection of points onto a horizontal plane, and mapping 2D triangles to 3D. The Delaunay triangulation method is used to create triangles while avoiding elongated triangles. This method maximizes the smallest angle of the triangle by ensuring that the circle circumscribed by the triangle's points does not have any other points. This ensures that all points are connected to the two closest ones, creating a unique triangular network. At the final step, texture mapping adds photographic texture to the reconstructed model to achieve visually accurate results, after fusing the texture from multiple overlapping images and considering the model's spatial resolution. The reconstructed 3D models are georeferenced using image metadata, i.e., the GPS data (latitude, longitude, and altitude), and online digital elevation model (DEM) servers, and a 3D similarity transformation between the SfM estimated cameras' poses and the GPS measured ones is also estimated. Figure 14 presents examples of 3D models created via the image-based 3D reconstruction service of the XR4DRAMA platform.

Accordingly, the satellite-based 3D reconstruction employs multi-spectral image enhancement and conversions, DEM simplification, altitude correction, geographic reference system conversions, and the 3D model triangular mesh estimation. In the last step, Delaunay triangulation is performed, and the triangulated model is then textured using the RGB image. Finally, the model is simplified to decrease the large number of redundant faces.

5.3. GIS Service

The XR4DRAMA system heavily depends on location-based information obtained from various sources to support its foreseen capabilities. The geographic information system (GIS) service is the back-end component responsible for collecting and organizing all data in a geospatial database following a location-coherent format that can be easily displayed on a map and accessed by the platform components and users. This service processes georeferenced information, enabling the use of location-dependent features, such as point-to-point navigation, within the platform. The GIS service collects information from various sources, including online geo-information databases, XR4DRAMA back-end services, and platform users. It supports the front-end tools by fusing content of different modalities, including 2D and 3D visual content, audio files, and generic data such as PDF files. The GIS service provides the following functionalities: (a) the creation of projects, (b) file storage management, (c) the retrieval of information about existing projects and data categories, (d) search engine support, (e) the creation of new POIs and the editing of

existing POIs, (f) task list organization, (g) navigation to a user-specified location, (h) risk report handling to mark dangerous areas, and (i) elements at risk estimation.

Figure 14. Views of an example 3D model generated from drone footage by the 3D reconstruction service.

Two geoportals are currently exploited to populate the GIS database upon project creation: (a) the OpenStreetMap, from which geographical data for POIs, road networks, and place names are obtained, and (b) the geoportals maintained by the Vicenza municipality authority and the Alto Adriatico Water Authority, which include sandbags, shelters, risk maps, and flood maps, among others. This portal hosts the GeoServer that provides the geospatial data extracted from the XR4DRAMA platform for the flood management use case. Their extracted data include, among others, forecasting models, public services, and buildings. The extracted geospatial data are used to create a detailed and accurate view of the affected area in the event of a disaster, enabling emergency responders to plan and execute rescue operations more effectively.

The geospatial information extracted by the available geoportals during project creation is organized by the GIS service in a two-level category format. Each information element essentially contains a main category and subcategory tag, representing its thematic footprint. Figure 15 briefly depicts the available POI (and ROI) categories and corresponding subcategories.
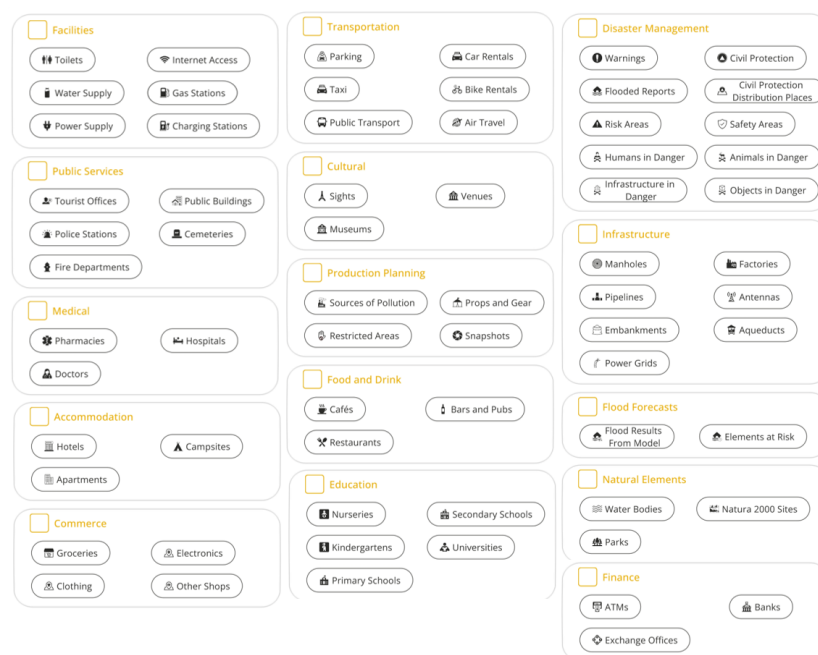
Figure 15. The POI and ROI categories and subcategories organized in the XR4DRAMA platform.

6. Evaluation and Results

This section is dedicated to the description of the two use cases and the evaluation that has been conducted in their framework. As the system's implementation followed a user-centric design approach (that focuses on understanding the needs, goals, and preferences of users and designing a system that meets those needs), the quality of the design and the system as a whole are evaluated in terms of satisfaction, usability, efficiency, effectiveness, and the supported functionalities in relation to their requirements. At the moment of the writing of this paper, the last development cycle of the system is almost finished, and an additional pilot testing round will follow. Hence, in the subsections below, we report the evaluation of an earlier version of the system, and it has to be noted that most of the recommendations provided in these last tests have already been addressed.

6.1. Flood Management Use Case

Flood management is a complicated process that requires prompt action and cooperation between first responders and authorities. The XR4DRAMA system aims to create situation awareness to support remotely working disaster managers and in situ first responders. The system treats pre-emergency, during emergency, and post-emergency situation needs to prevent a negative impact on the effective emergency response plan and actions resulting from a lack of awareness of the emergency environment and progression.

During a flood management situation, there are essentially two operating phases. The first phase consists of emergency preparations (the control room needs to be alerted on the expected effects of the event to better manage the situation, plan actions, and verify the procedures of the civil protection plans). In the second phase, just before and during the emergency, the control room must be able to view information from the territory, such as the location of any requests for people in danger and any potential critical situations, and establish a bilateral exchange of information with first responders in order to manage civil protection interventions efficiently, effectively, and in a timely manner.

The pilot was conducted in the city of Vicenza (Figure 16) and involved three different places. The first is the control room, which is placed in the Vicenza municipal operative command center, the second is in the city center of Vicenza, and the third is the S. Agostino–Park Retrone district, a suburban–rural area situated in the southern part of the Municipality of Vicenza and crossed by the Retrone river. The last two places are typical examples of

where the people in the field are operating. This test was performed by the end-users and stakeholders of the flood management scenario, and the following roles had been assigned:

- Control room operators: These operators used the authoring tool to obtain forecasts, conduct real-time monitoring of the crisis's progression, issue region-wide notifications to citizens, and establish bidirectional contact with and from first responders equipped with the AR application. The people who performed these roles stayed in the control room during the pilot. This role was carried out by members of the Vicenza Municipality and the Alto Adriatico Water Authority (AAWA).
- Civil protection volunteer teams: During the pilot, there were two teams of first responders positioned in different locations according to the storyline. The leaders of each team used the XR4DRAMA AR application to communicate with the control room; provide incident reports including text, image, and video; and receive tasks. The teams were formed by AAWA personnel.
- Citizens: Citizens used the XR4DRAMA citizen awareness mobile application to send incident reports (i.e., text, audio, image, video) and to receive notifications from the first responders. During the pilot, the participants who were assigned the role of citizens were located in specific areas of the city, according to the use case story. The role of the citizens was also carried out by AAWA personnel.
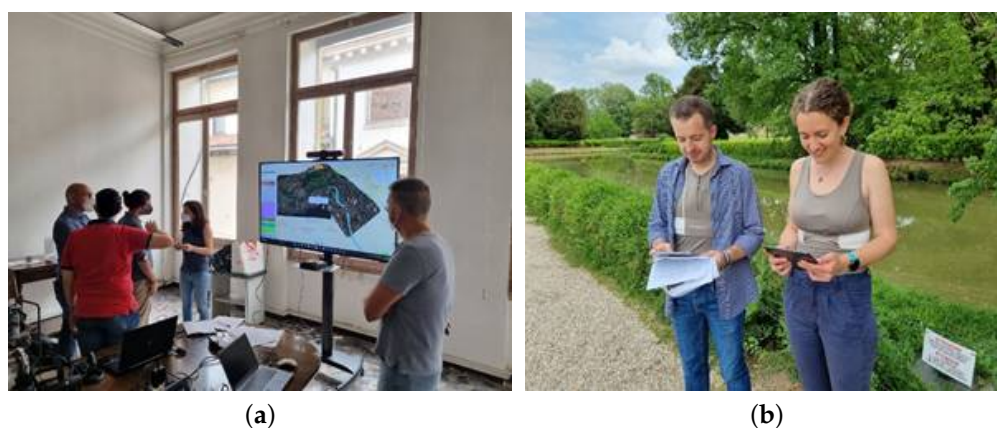

(**a**)   (**b**)

Figure 16. Flood management pilots in Vicenza: (**a**) Preparations in the control room; (**b**) first responders testing the AR application and smart vest in the field.

The first tool for evaluating the platform in this pilot was the observation sheets. During each session, these sheets were used to compile feedback and notes taken by the actors. Each actor was assigned a specific role and instructed to record details such as task performance, timing, and any issues encountered. Additionally, actors were encouraged to write down any helpful comments related to their experience with the XR4DRAMA technology. Observation sheets are structured documents that can be filled out in the field during demonstrations by observers or other users not directly involved in the project. The sheets are designed to assess the workflow of the XR4DRAMA platform in real emergency situations. They contain a set of 35 actions that cover all requested functionalities, ranging from simple tasks such as login and project creation to complex processes involving multiple front-end and back-end tools, such as the visualization of the citizen flood reports in the authoring tool.

The observation sheets' results concluded in providing feature suggestions to be added or altering existing ones based on the end user experiences on the Vicenza pilot. Most of the foreseen actions in the test were performed correctly, and for the partially performed actions, feedback was provided to guide the necessary adjustments for the final version of the system.

The second method for the platform's evaluation was questionnaires (see Supplementary Materials). In total, 37 tailored questions were provided to the platform's testers,

including standardized questions for the assessed system to ensure consistency and objectivity. Every questionnaire included a set of inquiries regarding the actors' evaluation of the structure and organization of the pilot as well as the simplicity of carrying out specific tasks using the XR4DRAMA technology. Furthermore, there were questions concerning the assessment of certain system functionalities and the actors' overall satisfaction with the implemented tools. Specific questions are provided for each tool used in the exercise, focusing on usability, alignment with expectations, and gathering any additional needs or ideas for the final system. In addition, the standardized questions covered UI friendliness, usefulness for decision making and operations, enhancement of situation awareness, clarity of information, and overall impression. This way, the questionnaires also serve as a means for technical partners to validate certain decisions.

In addition, the questionnaires gather general information about the users, including their background, age, and gender, to ensure a balanced and representative evaluation group. The evaluation involved an equal distribution of male and female participants, aged between 20 and 50 years, with diverse technical backgrounds in crisis management, research, and first responder roles. Most participants had prior experience with VR tools, which contributed to the study's success. Overall, this comprehensive evaluation process assesses the effectiveness of different tools in supporting disaster management, considering crucial factors for success in this context.

The questionnaire results observed during the Vicenza pilot in terms of the general impression and the usefulness from an operational point of view are presented for all tools in Figure 17. Figure 18a illustrates the participants' opinions in terms of whether the tools utilized by the professionals can increase situation awareness or not. Finally, Figure 18b indicates the participants' opinions on the impact of the trial as well as its realistic conditions.

The last approach employed for the assessment was a debriefing session for feedback collection. The debriefing session took place immediately following the conclusion of the pilot, during which the participants shared their thoughts and provided valuable feedback on their experience with the XR4DRAMA technology, in terms of their roles, what they liked, what they found challenging, and suggestions for future improvement. The debriefing session provided a platform for an open and interactive discussion between the participants and the organizers. It created an informal atmosphere where everyone could freely share their experiences and insights gained from using the system. The debriefing session not only allowed participants to express their individual perspectives but also facilitated a collective exchange of ideas and opinions among the group. The organizers actively listened to the participants' feedback, taking note of their suggestions and insights. This collaborative approach ensured that the final system could benefit from the diverse perspectives and expertise of the participants.

Overall, the assessment of the XR4DRAMA system yielded a generally positive outcome. It had been noted that a few adjustments and enhancements would be necessary to ensure the final version meets all user requirements. The presence of the representatives of the municipality of Vicenza, who were responsible for emergency management, comprised a huge benefit during the evaluation process, as their input is crucial for the correct guidance of the development of XR4DRAMA from an end-user perspective. Their feedback was particularly encouraging, and they expressed positive sentiments about the platform and the underlying technologies. With respect to all the end user tools, including the smart sensing device, the feedback was positive in terms of its clarity and general impression. Users found the tool to be clear and beneficial, with many expressing appreciation for its operational and decision-making capabilities. With this in mind, upon their request, further improvements were made in the final system to achieve the expected outcome.
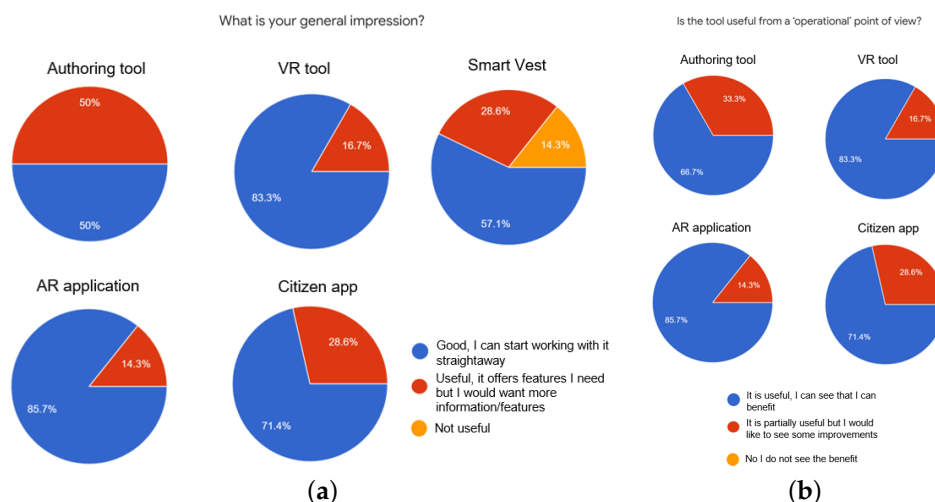
Figure 17. Questionnaire results observed during the pilots in Vicenza. (**a**) The general impression of all tools. (**b**) The usefulness of the software tools from an operational point of view.
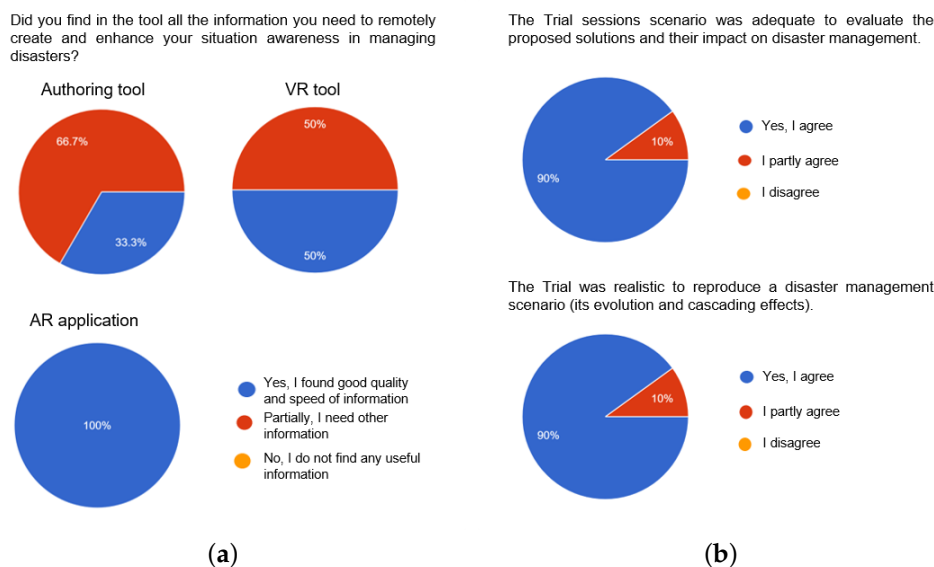

Figure 18. Questionnaire results observed during the pilots in Vicenza. (**a**) Rate of the software tools used by the professionals in terms of situation awareness. (**b**) Trial session impact and realistic conditions.

6.2. Media Production Use Case

Media production planning is a similarly complex and often complicated process. There are numerous approaches, strategies, and tools used depending on the type of production. Production management, for instance, needs to decide on the size and qualification of production staff, the necessary equipment, the production schedule, travel arrangements, and legal requirements, among other concerns. The situation awareness that we intend to create specifically aims to support production managers who have not visited and experienced the production location in person. This lack of awareness about the location and the conditions in situ often complicates their work considerably and might even have a negative impact on production quality.

The scenario in which the platform is showcased and evaluated is a documentary on the city and island of Corfu, Greece. Such documentary planning consists of three steps. In the first step, when the Corfu project is created in the system, an initial automated query is carried out with respect to Corfu. This happens by retrieving and indexing information about Corfu from openly available sources as well as from a predefined set of proprietary repositories (e.g., Earth observation (EO) databases). The results of this query are analyzed,

organized, and presented in a user-friendly way. This phase is conducted in the control room and contributes to Level 1 SA. In the second step, production management assigns a location scout to travel to Corfu and further investigates the location. Alternatively, this location scout could already live on Corfu or operate in the area. The location scout is provided with the results of the initial query about Corfu and is able to access the XR4DRAMA system via the AR application. The location scout verifies and adjusts the provided information. The location scout will also add new information, which could be text, images, videos (also from drones), and data from other available sensors (e.g., temperature, precipitation, noise). The location scout also explores several suitable shooting and filming locations in Corfu and provides relevant information and content. The final phase in planning the production on Corfu covers the process where the production management team in the control room receives all relevant information about the various locations on Corfu via the XR4DRAMA platform (also in immersive mode) and creates a detailed production plan. The basis for a good and realistic production plan is that the management team has a very good understanding of the location in question, and the system contributes to that direction in all the aforementioned steps.

For the evaluation of the platform in this use case, Deutsche Welle conducted the pilot in Corfu (Figure 19), where all three steps for the documentary production planning process were executed. The whole project's consortium also participated in the activities to observe and provide feedback on the implemented functionalities. Regarding the evaluation criteria, a checklist approach (writing down which features are properly supported in each step of the media production planning scenario) was combined with a catalog of carefully selected qualitative questions that are reflected in the walkthroughs and summarized user interviews.

The agenda of the activities in the pilot can be summarized as follows:

- Technical setup (laptops, XR equipment etc.);
- Download and initial startup of the latest project software;
- Preparation of related documents/spreadsheets to record the scenario steps, the associated requirements, the expected performance, and important notes;
- Testing of the platform for the first step of the scenario in a space used as a control room;
- Execution of second step of the scenario, i.e., location scouting, media recording, and data verification on site on the island (in Corfu City);
- Testing of the platform for the third step of the scenario in the same control room space;
- More system test runs and documentation (including post-processing of the notes kept during each scenario step execution);
- Evaluation (based on qualitative questions) and technical discussion towards further improving the platform,

The six key questions that had to be answered in each testing phase were the following:

- How stable and mature is the system?
- How much situation awareness does the system provide?
- What about user experience (UX), user interface (UI), usability?
- To what extent is the user able to fulfill their task and achieve their goals? (effectiveness)
- How much effort does the user need to invest to come up with accurate and complete results? (efficiency)
- How satisfied is the user with the system? (satisfaction)

In terms of a general assessment, the test run on Corfu showed that the consortium has managed to build a very solid, stable prototype for the media use case. The XR4DRAMA platform could already be used in a real-life setting and was on the verge of becoming a feature-rich XR production planning tool.

The system was able to provide enhanced situation awareness. Production areas could be defined in an easy and accurate fashion, and the platform retrieved several POIs and also generated basic 3D models (Level 1 SA).

|(**a**)|(**b**)|

Figure 19. Media production planning pilots in Corfu. (**a**) Immersion in the control room. (**b**) Location scouting using the AR application.

Tasks regarding research and AV documentation could be created, edited, sent, and fulfilled by the field agents, who were equipped with smartphones running the AR application and additional devices such as 360° cameras and audio recorders (Level 2 SA).

The fully immersive mode worked on a more than acceptable level, already giving users a good sense of the real, physical place of production (Level 3 SA).

The evaluators found the UX, UI, and usability to be quite solid, with the entire process headed in the right direction, and all minor design/interaction issues were meticulously documented/analyzed (and resolved in the following period).

Users commented that they had the flexibility to accomplish any task at any time. The platform allows for an extensive range of POIs to be added, modified, and enriched with various forms of content. Furthermore, users have the option to manually enhance the automatic information gathering on the dashboard. The basic block models in 3D/VR mode have the potential to be transformed into high-fidelity, photogrammetry-based assets. The XR4DRAMA platform facilitates the creation of a digital twin of almost any location.

Performing basic data correction and augmentation to form accurate and complete results could be accomplished relatively easily. It is important to note that a single person can now achieve satisfactory results without the need for an entire scouting team. This demonstrates the efficiency of the platform, which enables individuals to complete tasks that would typically require a team effort.

Lastly, users were generally quite satisfied with the platform and its comprehensive range of services. Its innovative, novel approach and user-centric design have been well received and have contributed to the overall favorable feedback.

6.3. Summary

To facilitate user comprehension with respect to the two pilots (scenarios and tests), Table 3 summarizes their details.

Table 3. Pilot evaluation details for both use cases.

Use case	Location	Phases	End User Tools	Participants	Evaluation Tools
Flood management	Vicenza, Italy	**Phase 1**: Emergency preparation **Phase 2**: Information update by field workers and emergency management	**Authoring tool** (Phase 1, Phase 2) **VR tool** (Phase 1, Phase 2) **AR application** (Phase 2) **Citizen app** (Phase 2) **Smart vest** (Phase 2)	**Control room operators** (Phase 1, Phase 2) **Civil protection teams** (Phase 2) **Citizens** (Phase 2)	Observation sheets Questionnaires Pilot debriefing session
Media Production	Corfu city, Greece	**Phase 1**: Project creation **Phase 2**: Information gathering and updates from location scouts **Phase 3**: Immersive mode	**Authoring tool** (Phase 1, Phase 2) **AR application** (Phase 2) **VR tool** (Phase 3)	**Production management team** (Phase 1, Phase 2, Phase 3) **Location scouts** (Phase 2)	Assignment checklist Qualitative questions User interviews

7. Discussion, Conclusions, and Future Work

In this paper, we have introduced a comprehensive and fully featured XR solution designed for flood management and media production planning. The platform encompasses both user-facing tools and behind-the-scenes features that generate essential data/content to enhance situational awareness.

Our proposed platform stands out for both its performance and its seamless integration of cutting-edge technologies in a user-friendly manner. It caters to a wide range of users, fostering communication and collaboration when meticulous planning and/or rapid response and action are crucial. This sets it apart from other tools that focus on a more specialized scope, only partially addressing the needs of an operation.

The evaluation demonstrates that situational awareness is indeed heightened among XR4DRAMA users, both in flood management and media production planning. This is achieved through the automatic retrieval, analysis, and generation of data, such as (but not limited to) detailed maps, 3D models, and stress level measurements. Additionally, the manual content editing features allow a collaborating group to monitor real-time changes in dynamic situations or to simulate what a specific place will (probably) look like at a specific time.

Flood managers and first responders underlined the sense of immediateness the platform provides: no more wearisome calls and messenger threads, no more lengthy explanations and misunderstandings—just one XR4DRAMA session—and all stakeholders know what is going on and can make the right decisions. This is a main differentiating factor from other related tools (such as [6–9]) that focus on providing flood visualizations in XR. These solutions, even though they contribute to situation awareness in terms of preparation and planning for such events, are unable to handle the communication and information exchange between different professional groups and the general public in an integrated fashion, and they are restricted to the preparedness aspect of an incident that may happen in the future. As validated in the pilot tests in Vicenza, XR4DRAMA covers and contributes to both the preparation phase before the emergency and the highly crucial management that takes place during a flood.

Media production planners pointed out that XR4DRAMA's multimodality plays a big role in fostering insights and knowledge. There are all kinds of assets and pieces of information, and users can check them out in one place with only a couple of clicks. Designated production sites thus become very tangible. It is important to note that all

levels of situational awareness are equally valuable, including the seemingly simpler Level 1 and 2 SA, even though they may not be as visually stimulating as the fully immersive Level 3 SA. All in all, these users of XR4DRAMA were particularly excited about the prospect of exploiting XR for a different purpose than what has previously been explored. Recent research [10–14] has focused on facilitating the creation of appealing XR stories for audiences. On the other hand, this work enables media people to use a single and complete XR solution that handles and optimizes the management of any media production, addressing its requirements for collaboration, content sharing, and realistic visualizations of remote locations. Most importantly, it supports all the planning phases and all the different collaborating teams (in the control room and in the field) in such an assignment.

The only limitation of the XR4DRAMA platform is that it is heavily reliant on an internet connection, meaning that it cannot function properly and the users cannot exploit its capabilities if there is no stable network (e.g., in extreme catastrophe situations).

As this research is ongoing, immediate future work involves finalizing the platform's implementation and conducting further evaluation to assess improvements, usability, and user satisfaction in the final prototype. Future directions may include refining individual platform modules, incorporating emerging technologies (such as generative AI) and adapting the platform to support different disaster scenarios (such as earthquakes or forest fires). The XR4DRAMA platform can be easily adapted to any map-based scenario, requiring little more than adjustments to the POI categories and subcategories to accommodate specific user requirements.

Author Contributions: Conceptualization, S.S. (Spyridon Symeonidis) and S.D.; Funding acquisition, S.V.; Methodology, S.S. (Spyridon Symeonidis); Project administration, S.S. (Spyridon Symeonidis) and S.D.; Resources, M.P.; Software, S.S. (Spyridon Symeonidis), S.S. (Stamatios Samaras), J.G., C.S. and Y.S.; supervision, S.D., M.F. and S.V.; Validation, A.P. and M.F.; Writing—original draft preparation, S.S. (Spyridon Symeonidis), S.S. (Stamatios Samaras), A.P., M.P. and C.S.; Writing—review and editing, S.S. (Spyridon Symeonidis), S.D. and C.S. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data analyzed in this study are available within the referenced articles, as well as in the following publicly available links: http://places2.csail.mit.edu, accessed on 28 May 2023 and http://www.openslr.org/94/, accessed on 28 May 2023.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shin, Y.; Jiang, Y.; Wang, Q.; Zhou, Z.; Qin, G.; Yang, D.K. Flexoelectric-effect-based light waveguide liquid crystal display for transparent display. *Photonics Res.* **2022**, *10*, 407–414. [CrossRef]
2. Cheng, D.; Duan, J.; Chen, H.; Wang, H.; Li, D.; Wang, Q.; Hou, Q.; Yang, T.; Hou, W.; Wang, D.; et al. Freeform OST-HMD system with large exit pupil diameter and vision correction capability. *Photonics Res.* **2022**, *10*, 21–32. [CrossRef]
3. Pandey, A.; Min, J.; Malhotra, Y.; Reddeppa, M.; Xiao, Y.; Wu, Y.; Mi, Z. Strain-engineered N-polar InGaN nanowires: Towards high-efficiency red LEDs on the micrometer scale. *Photonics Res.* **2022**, *10*, 2809–2815. [CrossRef]
4. Chittaro, L.; Sioni, R. Serious games for emergency preparedness: Evaluation of an interactive vs. a non-interactive simulation of a terror attack. *Comput. Hum. Behav.* **2015**, *50*, 508–519. [CrossRef]
5. Tanes, Z.; Cho, H. Goal setting outcomes: Examining the role of goal interaction in influencing the experience and learning outcomes of video game play for earthquake preparedness. *Comput. Hum. Behav.* **2013**, *29*, 858–869. [CrossRef]

6. Rydvanskiy, R.; Hedley, N. Mixed Reality Flood Visualizations: Reflections on Development and Usability of Current Systems. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 82. [CrossRef]

7. Haynes, P.; Hehl-Lange, S.; Lange, E. Mobile Augmented Reality for Flood Visualisation. *Environ. Model. Softw.* **2018**, *109*, 380–389. [CrossRef]

8. Sermet, Y.; Demir, I. Flood action VR: A virtual reality framework for disaster awareness and emergency response training. In *ACM SIGGRAPH 2019 Posters*; Association for Computing Machinery: New York City, NY, USA, 2019; pp. 1–2.

9. Itamiya, T.; Kanbara, S.; Yamaguchi, M. XR and Implications to DRR: Challenges and Prospects. In *Society 5.0, Digital Transformation and Disasters: Past, Present and Future*; Kanbara, S., Shaw, R., Kato, N., Miyazaki, H., Morita, A., Eds.; Springer Nature: Singapore, 2022; pp. 105–121. [CrossRef]

10. Bösch, M.; Gensch, S.; Rath-Wiggins, L. Immersive Journalism: How Virtual Reality Impacts Investigative Storytelling. In *Digital Investigative Journalism: Data, Visual Analytics and Innovative Methodologies in International Reporting*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 103–111.

11. Symeonidis, S.; Meditskos, G.; Vrochidis, S.; Avgerinakis, K.; Derdaele, J.; Vergauwen, M.; Bassier, M.; Moghnieh, A.; Fraguada, L.; Vogler, V.; et al. V4Design: Intelligent Analysis and Integration of Multimedia Content for Creative Industries. *IEEE Syst. J.* **2022**, 1–4. [CrossRef]

12. Avgerinakis, K.; Meditskos, G.; Derdaele, J.; Mille, S.; Shekhawat, Y.; Fraguada, L.; Lopez, E.; Wuyts, J.; Tellios, A.; Riegas, S.; et al. V4design for enhancing architecture and video game creation. In Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, 16–20 October 2018 ; IEEE: Manhattan, NY, USA, 2018; pp. 305–309.

13. Brescia-Zapata, M. Culture meets immersive environments: A new media landscape across Europe. *Avanca Cine.* **2021**, 1029–1033. [CrossRef]

14. Wu, H.; Cai, T.; Liu, Y.; Luo, D.; Zhang, Z. Design and development of an immersive virtual reality news application: A case study of the SARS event. *Multimed. Tools Appl.* **2021**, *80*, 2773–2796. [CrossRef]

15. Forsberg, K.; Mooz, H. The relationship of system engineering to the project cycle. In *INCOSE International Symposium*; Wiley Online Library: Hoboken, NJ, USA, 1991; Volume 1, pp. 57–65.

16. Kohlschütter, C.; Fankhauser, P.; Nejdl, W. Boilerplate detection using shallow text features. In Proceedings of the Third ACM International Conference on Web Search and Data Mining, Association for Computing Machinery, New York City, NY, USA, 4–6 February 2010; pp. 441–450.

17. Tsikrika, T.; Andreadou, K.; Moumtzidou, A.; Schinas, E.; Papadopoulos, S.; Vrochidis, S.; Kompatsiaris, I. A unified model for socially interconnected multimedia-enriched objects. In Proceedings of the MultiMedia Modeling: 21st International Conference, MMM 2015, Sydney, NSW, Australia, 5–7 January 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 372–384.

18. Souček, T.; Lokoč, J. Transnet V2: An effective deep network architecture for fast shot transition detection. *arXiv* **2020**, arXiv:2008.04838.

19. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.

20. Batziou, E.; Ioannidis, K.; Patras, I.; Vrochidis, S.; Kompatsiaris, I. Low-Light Image Enhancement Based on U-Net and Haar Wavelet Pooling. In Proceedings of the MultiMedia Modeling: 29th International Conference, MMM 2023, Bergen, Norway, 9–12 January 2023; Springer: Berlin/Heidelberg, Germany, 2023; pp. 510–522.

21. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 801–818.

22. Chen, L.C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv* **2017**, arXiv:1706.05587.

23. Zhou, B.; Zhao, H.; Puig, X.; Fidler, S.; Barriuso, A.; Torralba, A. Scene parsing through ade20k dataset. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 633–641.

24. Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *6*, 679–698. [CrossRef]

25. Pratap, V.; Hannun, A.; Xu, Q.; Cai, J.; Kahn, J.; Synnaeve, G.; Liptchinsky, V.; Collobert, R. Wav2letter++: A fast open-source speech recognition system. In Proceedings of the ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; IEEE: Manhattan, NY, USA, 2019; pp. 6460–6464.

26. See, A.; Liu, P.J.; Manning, C.D. Get to the point: Summarization with pointer-generator networks. *arXiv* **2017**, arXiv:1704.04368.

27. Gu, J.; Lu, Z.; Li, H.; Li, V.O. Incorporating copying mechanism in sequence-to-sequence learning. *arXiv* **2016**, arXiv:1603.06393.

28. Strötgen, J.; Gertz, M. Heideltime: High quality rule-based extraction and normalization of temporal expressions. In Proceedings of the 5th International Workshop on Semantic Evaluation, Uppsala, Sweden, 15–16 July 2010; pp. 321–324.

29. Casamayor, G. Semantically-Oriented Text Planning for Automatic Summarization. Ph.D. Thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2021.

30. Camacho-Collados, J.; Pilehvar, M.T.; Navigli, R. Nasari: Integrating explicit knowledge and corpus statistics for a multilingual representation of concepts and entities. *Artif. Intell.* **2016**, *240*, 36–64. [CrossRef]

31. Straka, M.; Straková, J. Tokenizing, pos tagging, lemmatizing and parsing ud 2.0 with udpipe. In Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies, Vancouver, BC, Canada, 3–4 August 2017; pp. 88–99.

32. Ballesteros, M.; Bohnet, B.; Mille, S.; Wanner, L. Data-driven deep-syntactic dependency parsing. *Nat. Lang. Eng.* **2016**, *22*, 939–974. [CrossRef]

33. Bohnet, B.; Wanner, L. Open Source Graph Transducer Interpreter and Grammar Development Environment. In Proceedings of the LREC, Valletta, Malta, 17–23 May 2010.

34. Xefteris, V.R.; Tsanousa, A.; Symeonidis, S.; Diplaris, S.; Zaffanela, F.; Monego, M.; Pacelli, M.; Vrochidis, S.; Kompatsiaris, I. Stress Detection Based on Wearable Physiological Sensors: Laboratory and Real-Life Pilot Scenario Application. In Proceedings of the Eighth International Conference on Advances in Signal, Image and Video Processing (SIGNAL), Barcelona, Spain, 13–17 March 2023; pp. 7–12.

35. Makowski, D.; Pham, T.; Lau, Z.J.; Brammer, J.C.; Lespinasse, F.; Pham, H.; Schölzel, C.; Chen, S.A. NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behav. Res. Methods* **2021**, *53*, 1689–1696. [CrossRef]

36. Siedlecki, W.; Sklansky, J. A note on genetic algorithms for large-scale feature selection. *Pattern Recognit. Lett.* **1989**, *10*, 335–347. [CrossRef]

37. Chen, T.; He, T.; Benesty, M.; Khotilovich, V.; Tang, Y.; Cho, H.; Chen, K.; Mitchell, R.; Cano, I.; Zhou, T.; et al. Xgboost: Extreme Gradient Boosting. Available online: https://cran.microsoft.com/snapshot/2017-12-11/web/packages/xgboost/vignettes/xgboost.pdf (accessed on 28 May 2023).

38. Eyben, F.; Wöllmer, M.; Schuller, B. Opensmile: The munich versatile and fast open-source audio feature extractor. In Proceedings of the 18th ACM International Conference on Multimedia, Firenze, Italy, 25–29 October 2010; pp. 1459–1462.

39. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–28. [CrossRef]

40. Vassiliades, A.; Symeonidis, S.; Diplaris, S.; Tzanetis, G.; Vrochidis, S.; Bassiliades, N.; Kompatsiaris, I. XR4DRAMA Knowledge Graph: A Knowledge Graph for Disaster Management. In Proceedings of the 2023 IEEE 17th International Conference on Semantic Computing (ICSC), Laguna Hills, CA, USA, 1–3 February 2023; IEEE: Manhattan, NY, USA, 2023; pp. 262–265.

41. Vassiliades, A.; Symeonidis, S.; Diplaris, S.; Tzanetis, G.; Vrochidis, S.; Kompatsiaris, I. XR4DRAMA Knowledge Graph: A Knowledge Graph for Media Planning; In Proceedings of the 15th International Conference on Agents and Artificial Intelligence—Volume 3: ICAART, Lisbon, Portugal, 22–24 February 2023; pp. 124–131.

42. Mel'čuk, I.A. *Dependency Syntax: Theory and Practice*; SUNY Press: Albany, NY, USA, 1988.

43. Mille, S.; Dasiopoulou, S.; Wanner, L. A portable grammar-based NLG system for verbalization of structured data. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, Limassol, Cyprus, 8–12 April 2019; pp. 1054–1056.

44. Du, S. Exploring Neural Paraphrasing to Improve Fluency of Rule-Based Generation. Master's Thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2021.

45. Thompson, B.; Post, M. Paraphrase generation as zero-shot multilingual translation: Disentangling semantic similarity from lexical and syntactic diversity. *arXiv* **2020**, arXiv:2008.04935.

46. Stentoumis, C. Multiple View Stereovision. In *Digital Techniques for Documenting and Preserving Cultural Heritage*; Collection ed.; Bentkowska-Kafel, A., MacDonald, L., Eds.; Amsterdam University Press: Amsterdam, The Netherlands, 2017; Chapter 18, pp. 141–143. [CrossRef]

47. Bradski, G. The OpenCV Library. *Dr. Dobb'S J. Softw. Tools* **2000**, *25*, 120–123.

48. Moulon, P.; Monasse, P.; Marlet, R. Adaptive structure from motion with a contrario model estimation. In Proceedings of the Computer Vision–ACCV 2012: 11th Asian Conference on Computer Vision, Daejeon, Korea, 5–9 November 2012; Springer: Berlin/Heidelberg, Germany, 2013; pp. 257–270.

49. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]