Critical Action Planning over Extreme-Scale Data

D2.2 - Initial Use Case Evaluation, Pilots, Demonstrators and Simulation Models and Tools

Version 2.0

Documentation Information

Contract Number	**101092749**
Project Website	https://crexdata.eu/
Contractual Deadline	M18, 06.2024
Dissemination Level	Public
Nature	Report
Author	Arnau Montagud (BSC), Miguel Ponce de Leon (BSC), Jens Pottebaum (UPB), Dimitris Zissis (UAEGEAN), Marcel Ebel (UPB), Georgios Grigoropoulos (Kpler)
Contributors	Giannis Spiliopoulos (UAEGEAN), Elias Xidias (UAEGEAN), Piergio Castioni (URV), Thalia Diniaco (BSC), Beatriz Eguzkitza (BSC), Lisa Haga (FMI), Jasmina Schmidt (DCNA), Christina Rechberger (DCNA), Ivana Kruijff (DRZ), Merlin Stampa (DRZ), Enes Derin (FDDO)
Reviewer	Elias Alevizos (NCSR), Charilaos Akasiadis (NCSR), Manos Pitsikalis (NCSR)
Keywords	Use case evaluation, pilots, demonstrator, forecasting, online data anlaysis

CREXDATA has received funding from the European Union's Horizon Europe programme under grant agreement number 101092749.

Change Log

Version	Author	Date	Description Change
V0.1	Marcel Ebel (UPB)	02/05/2023	First Structure
V0.15	Jens Pottebaum (UPB)	06/05/2024	Comments
V0.2	Marcel Ebel (UPB)	14/05/2024	Adjustment of the structure & EmCase partner responsibilities
V0.25	Arnau Montagud (BSC)	17/05/2024	Added health case Sections and drafted the introduction
V0.3	Jens Pottebaum (UPB)	22/05/2024	Added emergency case current draft
V0.4	Dimitris Zissis, Giannis Spiliopoulos, Elias Xidias (UAEGEAN)	23/05/2024	Added maritime use case
V0.5	Miguel Ponce de Leon (BSC), Piergio Castioni (URV)	26/05/2024	Added epidemiological use case scenario
V0.6	Arnau Montagud (BSC)	29/05/2024	Merging the different parts in the template
V0.7	Jens Pottebaum (UPB)	01/06/2024	Incorporate EmCase content
V0.75	Arnau Montagud (BSC)	03/06/2024	Final edits, added executive summary
V0.8	Elias Alevizos, Charilaos Akasiadis, Manos Pitsikalis (NCSR)	11/06/2024	Document revision
V0.9	All authors	25/06/2024	Integrated the changes required from the revision
V1.0	Antonios Deligiannakis (TUC)	28/06/2024	Final Version

Contents

Executive Summary

The CREXDATA project reaches a key milestone, with the completion of the interim evaluation of its use cases, pilots, demonstrators, and simulation models and tools. This evaluation, conducted using a common methodology, has ensured consistency across the different use cases.

The initial evaluations have shed light on the effectiveness of the use cases and highlighted potential areas for improvement. The pilots and demonstrators have proven the practical applicability of our research, serving as tangible proof of the project's real-world impact.

The simulation models and tools have been used in testing hypotheses and validating solutions, allowing for performance assessment in a controlled environment. The final scenario definitions for each use case have provided a clear and detailed description of the situations the project aims to address, guiding our development work.

The initial conclusions drawn from the evaluation results provide valuable feedback for future work, highlighting the strengths of our approach and areas needing further development. As we move forward, we will continue to refine our solutions, guided by the feedback from our evaluations and the needs of our use cases.

In summary, this report is a significant milestone in the CREXDATA project, providing a comprehensive overview of our progress to date and setting the direction for our future work.

1 Introduction

This report provides an interim evaluation of the use cases, pilots, demonstrators, and simulation models and tools developed within the scope of the CREXDATA project at the 18-month juncture and is a continuation of previous D2.1.

The CREXDATA project is an initiative aimed at enhancing data exchange and interoperability to forecast interventions in a fast and precise manner. The use cases form the foundation of this project, offering practical scenarios that drive our research and development. This report presents an interim evaluation of these use cases, identifying their strengths and areas for improvement.

The project consists of three use cases:

- **Weather Emergencies Use Case**: This use case aims to significantly improve situational awareness in weather emergencies, enabling informed decisions by civil protection authorities. It considers ranked future worlds with explicit uncertainties to avoid disaster impacts. The validation of this use case is performed in reproducible test bed scenarios and field trials.
- **Life Science Use Case**: The goal of this use case is to integrate large-scale machine learning with epidemiological and multi-scale simulation models. This integration will help develop a generalizable and flexible analytical platform for supporting decision-making processes, such as designing strategies for health crisis responses and treatment optimization.
- **Maritime Use Case**: This use case aims to develop a solution to use data originating from on board a vessel while fusing it with sensor data to create reliable digital twins of the involved assets. It will develop weather and emergency routing forecasting solutions for involved vessels relying on big data and AI technologies.

Alongside the use cases, we have developed several pilots and demonstrators that will include different technologies developed in the project (Table ***1***). These implementations serve as evidence of the project's impact, demonstrating the applicability of our research in real-world contexts. This report provides an overview of these pilots and demonstrators, emphasising their key features and contributions to the project.

Simulation models and tools form a significant part of our work. These tools are crucial for testing our hypotheses and validating our solutions. This report outlines the models and tools we have developed, detailing their function, operation, and significance within the project.

This report also includes the final scenario definitions for the use cases. These scenarios offer a detailed description of the situations our project aims to address. Additionally, we present initial conclusions drawn from the evaluation results across use cases, providing feedback that will guide our future work.

In summary, Deliverable 2.2 offers a comprehensive update on our project's progress, detailing our use cases, pilots, demonstrators, and simulation models and tools. It signifies a key milestone in our project, and we anticipate the continued progress of CREXDATA.

A glossary of terms has been added for the benefit of readers (**Table *2***).

Table *1*: Uptake of technologies in Use Cases (cf. [D2.1])

	Emergency Case		Health Case	Maritime Case
	Dortmund	Austria		
T2.4 Simulation and Tools	(x)[3]	(x)[3]	X	X
T3.2 Graphical Workflow Specification	(x)[2]	(x)[2]	X	X
T4.1 Complex Event Forecasting	(x)[3]	X	X	X
T4.5 Text Mining for Event Extraction	X	X		
T4.2 Interactive Learning for Simulation Exploration	(X)[2]	(X)[2]	X	X
T4.3 Federated Machine Learning	X	X	X	
T4.4 Optimized Distributed “Analytics as a Service”	(x)[2]	(x)[2]		X
T5.1 Explainable AI	(x)[2]	(x)[2]	X	
T5.2 Visual Analytics supporting XAI	(x)[2]	(x)[2]	X	
T5.3 Visual Analytics for Decision Making under Uncertainty	(x)[2]	(x)[2]	X	(X)[1]
T5.4 Augmented reality at the field	X	X		(X)[1]
T5.5 Uncertainty Visualization in Augmented Reality	(x)[2]	(x)[2]		(X)[1]

[1] support of potential TUC, FR contribution, [2] scheduled for M19-M36, [3] initial exploration in first trials

Table 2: Glossary

Abbreviation	Expression	Explanation
UC	*Use Case*	Applications of CREXDATA technology resp. the CREXDATA system in real-world scenarios. Within the project three use cases are defined: weather induced emergencies, health and maritime.
	Pilot (site)	Conceptual term to describe a set of stakeholders within their context like spatial environment, equipment, data sources etc. For each Use Case, several Pilot (sites) can be specified (for instance, Dortmund and Austria in the emergency case).
	Application Scenario	Procedural and structural description of potential uptake of CREXDATA technologies in Use Cases (for instance, flooding and forest fires in the emergency case).

Abbreviation	Expression	Explanation
	CREXDATA system	Output of WP3, integrating technologies created in WP4 and WP5 without use case-specific customizations. It includes customization and configuration functionality, esp. through graphical workflow management.
	Demonstrator (system)	Technical system based on the CREXDATA system, which is customized and configured for specific Use Cases, Pilots and/or Application Scenarios. The Demonstrator might include additional components both as data sources and sinks (for instance, legacy systems of end users or the ARGOS system in the emergency case).

2 Common Evaluation Methodology

CREXDATA aims at developing tools and methods that are ultimately useful for a wide range of use cases. To ensure the effectiveness and relevance of our approach, we have seeked the expertise and evaluation of a wide range of experts. Their backgrounds and relationships with modelling will provide valuable perspectives on the project's objectives and specific aspects of the different use cases. Their feedback will serve two main purposes:

- to validate that the requirements of a high-end analysis platform are fulfilled, ensuring that the use cases meet the needs and requirements of the experts, validating the relevance and applicability of the platform; and
- to evaluate the usability by assessing the ease of use of the system, ensuring that the platform is not only effective but also user-friendly and accessible.

With this, this project wants to involve main players of the field to demonstrate the potential of using modelling, AI technologies, and extreme-scale multimodal data to help in action planning in varied emergencies, let it be weather, health or maritime.

2.1 Validation: Fulfilment of Requirements

We contacted different expert users from public European research and high-education centres, Public Protection and Disaster Relief agencies (PPDR) as well as from industries to validate whether the requirements and characteristics of the platform and use cases matched their needs. These expert users are experienced professionals that cover different aspects in each of the use cases.

All these users were contacted and consented to participate in our questionnaire (Section 9) and their anonymised responses can be found in Section 10. From the interviews and questionnaires, we captured their interest in using data-streams and forecasting and the scarcity of current tools to use them.

They considered that the project fills a need, and they were willing to use these tools and framework, as can be seen from their evaluation on the project's objectives and their answers to the questions about their willingness to use real-time data (see Section 10). Most of them agreed on the importance of the present project and the impact its outcomes would have on their current workflows, as reported by the KPIs of the project. They are willing to incorporate these tools and data on their day-to-day work and are positive on the impact these would have in their fields.

In the weather-induced emergency case, requirements were stated both from a clear demand perspective and a technology-driven perspective. In the pilot sites in Dortmund and Innsbruck, different storylines were used with related CREXDATA technologies and and overlapping core of the system. Ten experts engaged with the CREXDATA team at the pilot sites. The prioritized test cases with individual components could be executed successfully, further use cases were incorporated to explore opportunities for later stages in the project. End users confirmed the fulfilment of requirements, being able to test the specialities of WP4/WP5 algorithms, models and tools. Both the value in single technologies as well as in an integrated setup, for instance presented with an Augmented Reality interface became perceivable.

For the Health crisis use case, most of the expert users had an interest on having a framework that could facilitate the development of forecasting techniques on real-time data.

In addition, they agreed on the usefulness of having a graphical user interface to be able to program analyses with little coding knowledge and their will to incorporate their own tools to such a software.

Note that some considered several aspects of the overall framework as unnecessarily complex (for instance using it on HPC clusters with the added complexity of streaming data out and to it) and one considered that the GUI was not necessary for them but suited for less-technical users.

For the Maritime Use Case, the majority of expert users contacted in this round participated in the Aegean Ro-Boat Race, as competing teams. For them the need for such a system capable of improving their situational awareness and assisting in forecasting future vessel behaviours, while avoiding dangerous conditions is vital. Their main interest was on the real time aspect, while mentioning that they would prefer timeliness over accuracy in some circumstances (e.g. alert in case of a collision).

2.2 Usability Evaluation: Questionnaire

In the pursuit of having a flexible platform for real-time critical situation management of emergency crises, the usability of the developed system is of paramount importance. A system, no matter how technically advanced, must be user-friendly and intuitive to ensure its effective utilization by expert users. To assess this, we employed the System Usability Scale (SUS), a reliable tool to measure perceived usability [1]. SUS is a questionnaire of ten items to which participants need to answer using a five-point Likert scale with verbal anchors at the extremes.

This Section presents the results of the SUS questionnaire completed by our expert users. These questionnaires were done on the present modules and use cases, which are not end-to-end functioning systems. Their feedback provides insights into the system's intuitiveness, efficiency, and overall user-friendliness. These insights will guide us in making necessary improvements to the system, ensuring it is not only technically robust but also user-centric. For instance, at this stage we evaluated usability and understandability of visualisations. The challenge Crexdata has in this regard is to develop appropriate visualisation of test data exposing different facets of movement and operation in spatial and temporal contexts.

Overall, the questionnaire results have been analysed to obtain an aggregated score for the usability of a product. The results of the SUS tests were average with a mean score of 58.7 (EmCase), 39.167 (Life Science) and 60.36 (Maritime).

We will repeat these questionnaires once the use cases and the architecture are completed by the end of the project.

3 Weather Emergency Use Case

According to the case study-related methodology introduced in Deliverable D2.1 (p. 23) (see Figure 1), the overarching research questions were further operationalised in the conceptual setup of the Weather Emergency Use Case (EmCase) and, more specifically, for application in pilot sites. Three guiding research questions were formulated:

a. Does information generated by artificial intelligence algorithms need to be visualized differently from information generated by traditional situational awareness and assessment?
b. Does it make a difference whether the uncertainty of information provided to enable situational awareness is based on algorithms or witnesses on the ground?
c. What influence does an algorithm-based recommendation have on action planning and decision?

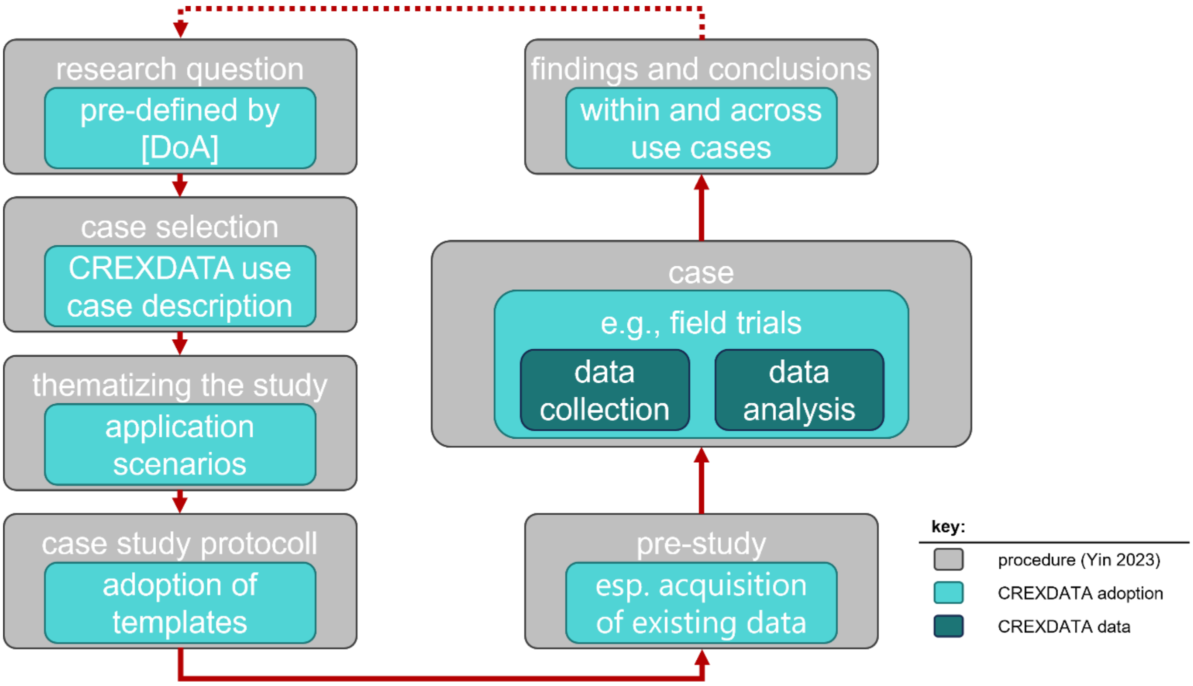


Figure 1: Weather emergency Use Case procedure (D2.1)

In line with both other use cases, the scenario description in Section 3.2 details domain-specific assumptions and environmental parameters. By that approach, the presented research questions are further broken down to specific test scenarios and, even more close to operational settings, test cases per technology. Thus, case selection for each technology is performed. Based on these specifications, Section 3.3 presents the specific conditions in pilots in correspondence with the demonstrator system. As presented in [D2.1], the demonstrator system subsumes CREXDATA components as well as domain-specific components like ARGOS. The study is thematized by the specific environments of Innsbruck and Dortmund. The case study protocol is briefly introduced, describing observational and survey conditions at evaluation locations. As kind of a pre-study, extensive data acquisition was conducted in Innsbruck and Dortmund to provide real data for the various technology partners in WP4 and WP5. After introducing the initial simulator component in Section 3.4, results and interpretations of this first evaluation phase are provided in Section 3.5.

3.1 Introduction

The application of CREXDATA in the EmCase was elaborated further on within the EmCase team, but also reflected with external experts through presentations and workshops. For instance, an event in Leoben (Austria) organized by DCNA was used to present and discuss CREXDATA results within a wider community of practitioners. Similarly, a poster was presented at the annual conference of the German Fire Protection Association (vfdb) to stimulate broader discussions. Further examples are provided in [D6.2]. Figure 2 presents the approach to apply CREXDATA technologies on top of data from a situation (e. g., from robotic systems), available data archives and forecasts (e. g., from meteorological models). Predictions are envisaged by applying new technologies, making them visible through situational maps in different command posts in terms of responsibility and size, as well as through Augmented Reality in decisive situations close to an incident.

Figure 2: Presentation of the CREXDATA concept to EmCase stakeholders [2], [3]

3.2 Scenario description

The explorative phase of CREXDATA was supported by informal “storylines”, established in the form of storyboards. Two storylines are created in intense stakeholder meetings to detail the general “pluvial urban flooding” scenarios. In the first storyline, the evacuation of a shopping mall in Dortmund (Germany) with heterogenous visitors is required. In the second storyline, water management with barriers is needed to assess and to avoid threats to traffic and buildings in Innsbruck (Austria). Bridging system design and system Verification & Validation (V&V), test scenarios are derived in the “case selection” phase. Each test scenario subsumes specific test cases. Test cases refer to specific data processing workflows, prospectively designed for implementation using Altair’s RapidMiner Studio. Field trials are scheduled with the dedicated purpose of evaluation. For the CREXDATA project itself, the mapping of evaluation activities to workflows through test scenarios and test cases ensures traceability of results, enables actual conclusions from application to data processing foundations, and simplifies communication among all partners [3].

3.2.1 Conceptual setup by test scenarios and test cases

In order to create a common understanding of the terminology, the terms test case and test scenario are systematically defined for the development of cyber-physical systems. Existing definitions from various disciplines such as software engineering, mechanics, etc. were analysed. Various characteristics of test cases and test scenarios were identified from the literature of the individual domains. These different definitions and their characteristics are merged and further developed towards the following definitions in the context of developing cyber-physical systems. The resulting definition for the term *test case* is as follows:

> *A Test Case is a procedure that serves to verify a property of a test item that has been specified as a requirement. The description of a Test Case must include a set of preconditions, inputs and expected results. It may include required test methods and test supports.*

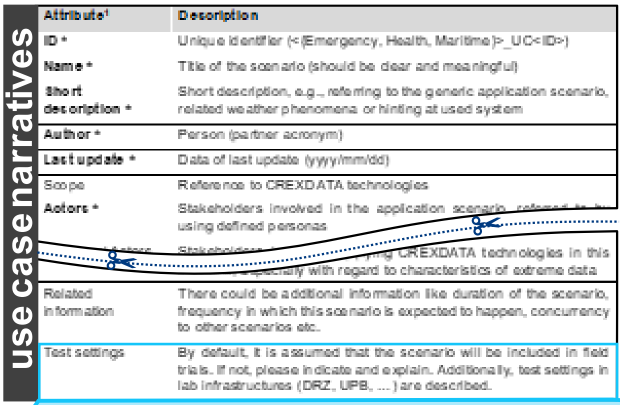
The resulting definition of the *test scenario* is:

> *A Test Scenario is a combination of two or more test cases ensuring execution of multiple test cases under the same environmental conditions. A Test Scenario must include common environmental conditions, Test Case inputs and expected Test Case results. It may include common test methods and test supports.*

In order to generate test cases and test scenarios based on the definitions presented in Section 3.1.2.1, templates for test scenarios (Table 3) and test cases (Figure 3) were developed that can be used as an auxiliary artifact in the creation and planning of verification activities for the CREXDATA system.

Table 3: Template Test Scenario

Test_Scenario_Name	Title of the test scenario
Test_Scenario_ID	Unique identifier for each test scenario
Test_Scenario_Owner	Responsible person, department or institution
Test_Case_IDs	Included test cases in the test scenario
Test_Item(s)	Included test items in the test scenario
Test_Scenario_Activities	Sequence of the test scenario and, if applicable, triggers that initiate various activities (can be modeled in an activity diagram)
Test_Scenario_Environmental_Parameter	Global parameters that apply to two or more test cases



List of Test Cases

Test_Case_ID	**Test_Case_Name**
TC-001	ARGOS demonstration Innsbruck
TC-002	ARGOS demonstration Dortmund
TC-003	CEF sewer network
…	…

Test_Case_Name	**Title of the Test Case**
Test_Case_ID	Unique Identifier of each test case
Test_Item	Item(s) to be tested
Test_Procedure	Methodical procedure of the test
CREXDATA_system_level	System level of the test item
Test_Case_Owner	Responsibility of the test case
Use_Case_underlaying	Incorporated use case
Detailed_Description	
Environmental_Conditions	
Test_Case_Environmental_Parameter	Global parameters of the test case (e.g., environmental temperature)
System_State	State(s) of the system within the test case
User	Relevant user and stakeholder
Input Parameter	Necessary inputs for the test case and for the test item
Test_Case_Sequence	Activities of the test cases that are executed
Expected_Result	Expected artifacts, values or system states

Figure 3: Test Case template (with the link to use case narratives in D2.1)

3.2.2 Test cases per technology

In order to evaluate the development progress of the various technologies, a pluvial flood situation represents the EmCase. Various test cases are generated with regard to use cases specified in Deliverable D2.1 and corresponding specific technologies. These test cases are incorporated in test scenarios which are carried out in terms of trials in Dortmund and Innsbruck. The test scenarios each contain a different combination of test cases. The test cases are planned using the templates presented above. The test cases refer to the results of the development progress of individual subsystems that will be used when applying the CREXDATA system in pluvial flooding (see details in Deliverables D4.1 and D5.1). Four of the test cases are not setup in terms of evaluation, but for deepening the exploration of possible use cases and requirements: With respect to TC_004, TC_005, TC_009 and TC_010, initial solution approaches for corresponding technologies are illustrated and used to discuss further development steps together with end users. Detailed specifications for the test cases enlisted in Table 4 are provided in appendix Section 11.1.

Table 4: List of Test Cases

Test_Case_ID	Test_Case_Name	Test_Case_Owner	Test Scenario	
			Innsbruck	Dortmund
TC_001	ARGOS demonstration Innsbruck	DCNA (HYDS)	X	
TC_002	ARGOS demonstration Dortmund	FDDO (HYDS)		X
TC_003	CEF sewer network	DCNA (NCSR)	X	( )*
TC_004	CEF flooding	DCNA (NCSR)	( )*	
TC_005	Parameter exploration in urban flooding simulation model	UPB (NCSR)	( )*	( )*
TC_006	3D mapping using UAV imagery	DRZ (TUC)	X	X
TC_007	Text mining Innsbruck	DCNA (BSC)	X	
TC_008	Text mining Dortmund	FDDO (BSC)		X
TC_009	XAI based on biometrical data	FDDO (CNR)		( )*
TC_010	Visual analytics of emergency protocols and precipitation data	FDDO (FR)		( )*
TC_011	AR routing Innsbruck	DCNA (TUC)	X	
TC_012	AR visualizing points of interest Innsbruck	DCNA (TUC)	X	
TC_013	AR routing Dortmund	FDDO (TUC)		X
TC_014	AR visualizing points of interest Dortmund	FDDO (TUC)		X
TC_015	Gradient Boosting machine learning technique	MoIFI (FMI)		

X = formative evaluation in first trials / ( ) = initial exploration in first trials*

3.3 Pilot definition and demonstrator

In Deliverable D2.1, pilots sites, lab opportunities, potential demonstrator system components as well as initial application scenarios are described. The consortium agreed on a focus on pluvial floodings for the first half of the project. Aligning pilot sites to a common scenario setting helped both to align discussion within the EmCase team and to streamline discussions between use case and technology partners, for instance, in exploring data sources. Thematizing the study for this first evaluation phase, the demonstrator system was setup according to the progress of technology evolution. Available technologies were adopted from WP4 and WP5 (cf. Deliverables D4.1 and D5.1). A brief overview is presented in Section 3.3.1, as a fundament for detailed descriptions of research in all CREXDATA pilot sites (Sections 3.3.2 and following).

3.3.1 Demonstrator system

In parallel to the integration of CREXDATA system components into an integrated system (cf. Deliverable D3.1), single components were provided by WP4/WP5 for the initial evaluation in pilot sites. Preparation concerned both the technology implementation as well as the adoption through specific use cases with corresponding specifications of data sources and provision of test datasets. Figure 4 presents the selection based on the initial system architecture sketch presented in Deliverable D2.1 (p.28). Green boxes represent demonstrator system components and datasets that are brought in by EmCase partners and adopted for testing the CREXDATA system. This concerns especially

- the ARGOS system providing the ground for data flows and an “interactive screen” in terms of a map-based Graphical User Interface (see Deliverable D2.1, pp.37-39),
- robotics components acting as sensor systems and, thus, providing data feeds (see Deliverable D2.1, pp.39-40),
- the Gradient Boosting machine learning technique developed and studied by FMI based on data from Helsinki emergency operations, and
- an exemplary application of simulators by incorporating an urban flooding simulation model (see Section 3.43.4.2).

In this first ambition, three technologies from WP4 (yellow boxes in the Figure 4) as well as Augmented Reality from WP5 (orange box) were selected to be tested with end users. For each of these technologies, specific use cases were selected in intense collaboration between requirements and technology partners. Further technologies were brought into discussions during field trials, informing the schedules for extending technologies and setting up interfaces between components in the second half of the project.

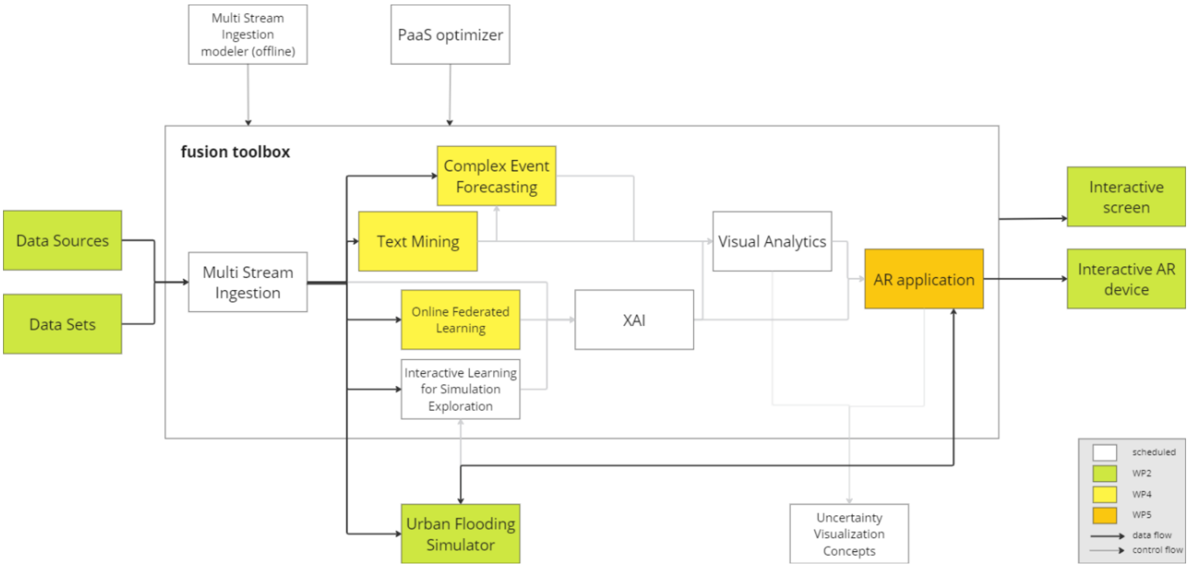


Figure 4: Focused elements of the CREXDATA system in the first evaluation phase

3.3.1.1 Complex Event Forecasting (T4.1)

For Complex Event Forecasting, a use case concerning the underground water distribution of the sewer network of the city of Innsbruck was prioritized. To this end, data pertaining to the technical structure of the sewer network was obtained from the municipal services of the city of Innsbruck. Other data pertinent to this objective include the water distribution within the sewer network and precipitation data during the heavy rainfall event AMRAS 2016. The objective of complex event forecasting is to provide as precise a statement as possible about the timing and location of sewer network overloads resulting from specific inputs (e.g., precipitation) or blockages at specific points of the sewer inlet network. This information can then be used to support decisions regarding the opening of openings or activation of pumps to relieve the sewer network during extreme weather events.

3.3.1.2 Online Federated Learning (T4.3)

Federated Machine Learning is adopted for innovative algorithms that might replace conventional photogrammetry. Research questions were broken down to three specific ones: Do situation models produced by photogrammetry tools provide useful additional support for mission planning? What is a good compromise between model quality and model preparation time, incl. flight and computation time? Does an approach based on machine learning provide added value over a standard approach? The aim is to process photos from UAVs and generate three-dimensional insights of a current situation, in best case in real-time “on the fly”. Several evolving technologies were investigated by WP4 partners and prepared for testing with end users.

3.3.1.3 Text Mining (T4.5)

T4.5 analysed the social media data surrounding the heavy rainfall event in the Amras district of the city of Innsbruck in July 2016 on Twitter. Identified postings were presented and discussed in the field trials with decision-makers from the various authorities from the city of Innsbruck as well as from the city of Dortmund in the pluvial flood in terms of relevance. On this basis, relevant social media postings (including text with pictures or videos) can help decision-makers in crisis situations to obtain a better picture of the situation and make faster and better decisions.

3.3.1.4 Augmented Reality (T5.4/T5.5)

T5.4/T5.5 developed an application for the visualization of water levels and sewer network data, as well as points of interest (POI) in the immediate vicinity of first responders with the emergency situation. The sewer network data of the city of Innsbruck was analysed. The location of the test case was selected based on the relevance of the location in flood situations and the number of POIs such as vulnerable buildings in the near of the rescue operation. In addition to the augmented reality application, the HoloLens 2 device used is also part of the test case.

3.3.1.5 Interactive Learning for Parameter Space Exploration (T4.2)

While Interactive Learning for Parameter Space Exploration is mainly focusing on other use cases besides the EmCase, the team explored opportunities to experiment with selected parameters in an urban flood simulation model in MIKE+. This is done via several differently configured input files for the simulation of the flood model (see specific use cases proposed by the EmCase team in Deliverable D4.1). The initial evaluation adopts such simulations with relevant interfaces in principle, analysing impact on current routines where simulations are hardly started within emergency response. Results shall be utilized to refine requirements collected in the initial phase.

3.3.1.6 Visual Analytics (T5.2/T5.3)

T5.2/T5.3 processed FDDO's deployment data (see Section 3.5.2) further and carried out various analyses and statistics. It was analysed whether there is a correlation between the deployment data (from Dortmund) and the weather data in order to determine whether certain precipitation events lead to increased deployments. On this basis, a forecast is to be drawn up that includes, among other things, the number of expected deployments in such weather conditions and their locations. Results of this analysis are used to detail requirements for technology uptake in the second half of the project.

3.3.2 Description of pilot site Innsbruck

Organisational settings

The organisational conditions were prepared by visits and continuous information exchange with five institutions. The Tyrol National Warning Center (NWZ Tyrol; Landeswarnzentrale Tirol) and the Fire Department of Innsbruck (FD IBK) cooperate as governmental organisations. The Control Centre Tyrol (CC Tyrol; Leitstelle Tirol) and the Municipal Service of Innsbruck (IKB) are institutionalised as private companies, operating for the public administration. All of these partners, associated through DCNA, provided expertise and contributed to the specification of test scenarios. In addition, contacts were established to GeoSphere Austria (GSA), a company that takes over responsibility as the Austrian meteorological service.

Stakeholders

Depending on the test case, different stakeholders were engaged in the field trial. For TC_001, TC_004, and TC_007, A-Level personnel of the Fire Brigade Innsbruck and the Upper Austrian Fire Brigade took part in the evaluation. For TC_006, TC_011 and TC_012, additionally to the aforementioned, C-Level fire brigade personnel took part as well as a task force expert of the Tyrol National Warning Center. For TC_003, a representative of the Municipal Service of Innsbruck (IKB) who has to decide whether to administer discharges during a heavy rain or flooding event, participated in this part of the trial.

Geographical settings

The trial took place in a hybrid setting with different geographical locations. The hybrid sessions with technology partners joining in remotely were carried out in a seminar room of the University of Innsbruck, Innrain 52a, 6020 Innsbruck, Austria (TC_001, TC_003, TC_004, TC_005, TC_007). For TC_006, the district of Amras (see Chapter 11.4) was chosen as a test location for the UAS flight and its respective data collection because of a significant heavy rain event that took place in 2016. This served as an exemplary event for the preceding data acquisition in the pilot site of Innsbruck. TC_011 and TC_012 (see Figure 5) focussed on the old town of Innsbruck with many historical buildings. The area contains several vulnerable infrastructures, such as public schools and kindergartens. Furthermore, the old town is located close to the river Inn and possesses the deepest point of the sewer system which could lead to extensive flooding due to a heavy rain event in the future.

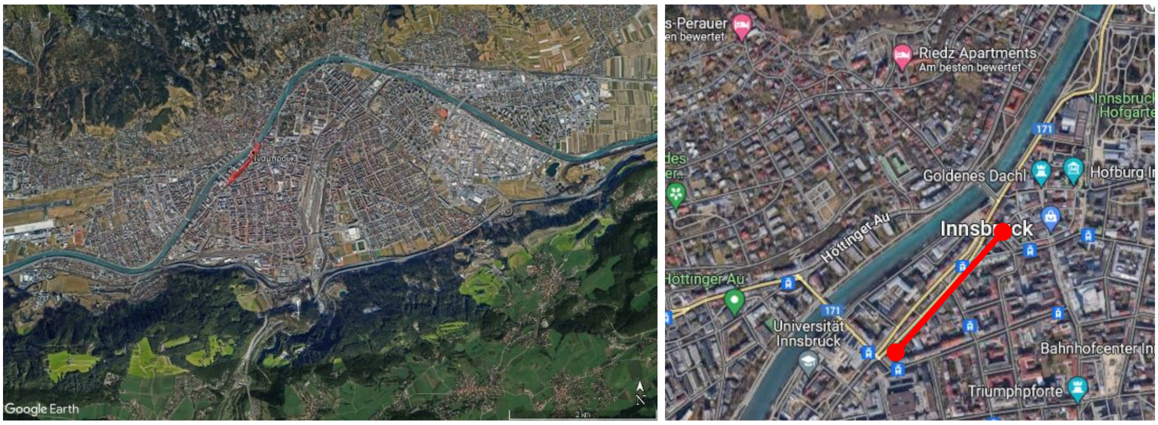


Figure 5: Geographical setting around Innsbruck University (AR track highlighted), source: Google / Google Earth.

Test cases included in the test scenario(s) for the Innsbruck pilot site

The test scenario assumes heavy rain in the city of Innsbruck. These environmental conditions cause pluvial flooding affecting, amongst other objects, a shopping mall during opening hours. Thus, an unknown number and grouping of people gets trapped in that shopping mall. Responding to that situation, all aforementioned emergency response organisations are alarmed and dispatch resources to the scene. Due to early forecasts, civil protection agencies already expected the event in advance. They rated the criticality of the event by assuming that its probability corresponds to 50y rainfall. As a consequence, they activated corresponding response plans. At start of the test scenario, the rain hit western edges of Innsbruck, then approached the city centre, and led to an increased gauge of the river Inn west of Innsbruck. An increase of emergency calls due to heavy rain at CC Tyrol was obvious, situation management was transferred to FD IBK. The situation is escalated by assuming emergent influences like heavier rain than expected, unexpected blockages in the sewer system, unexpected flooding of buildings, vulnerable and immobile people in these buildings. Table 5 correlates this scenario with test cases.

Table 5: Test Scenario TS_001: Innsbruck

Test_Scenario_Name	Test Scenario Innsbruck
Test_Scenario_ID	TS_001
Test_Scenario_Owner	DCNA
Test_Case_IDs	TC_001 ARGOS demonstration TC_003 CEF sewer network (TC_004 CEF flooding)* (TC_005 Parameter exploration in urban flooding simulation model)* TC_006 3D mapping using UAV imagery TC_007 Text mining TC_011 AR routing TC_012 AR visualizing points of interest
Test_Item(s)	Several CREXDATA sub-systems: ARGOS-system, text mining, complex event forecasting, Augmented Reality (hardware/software-application), MIKE+, UAV

Test_Scenario_ Activities	• View situational map and available forecasts for an expected emergency → TS_001 → set emergency category and level
	• View data analysis of incoming emergency calls correlated with rain gauges and precipitation forecasts → TC_004 → enumerate potential use cases resp. demands
	• View text mining results → TC_007 → evaluate contribution to situational picture
	• Categorisation of a situation in terms of a) drainage relevance and/or b) emergency escalation → TS_003 → change setting of outlets, pumps etc.
	• View situational map of flooding situation (TS_001), view options in urban flooding models → TS_005 → enumerate potential use cases resp. demands esp. with respect to set of envisaged use cases (cf. D4.1)
	• Send/assume command to on-site commanders to explore the situation, augmenting the view by predictions from the system/simulation → TC_012 → providing feedback with recommended decision for action
	• Assume necessity to rescue people from a shopping mall, assuming a certain duration of necessary rescue operations in correlation with trapped vulnerable people
	  ○ exploring accessible entries/exits around the building → TC_006 → evaluate 3D representation, decide (best case) in favour of visible entry/exit
	  ○ exploring entry routes and, assuming delay for rescue operations, safe exit routes → TC_011 → evaluate system recommendations, decide (best case) in favour of recommended route

Test_Scenario_ Environmental_Parameter	Derived from precipitation data of heavy rain event in July 2016 in the Amras district in Innsbruck

Detailed planning

The workflow / planning of the field trial in Innsbruck was as follows (Figure 6):

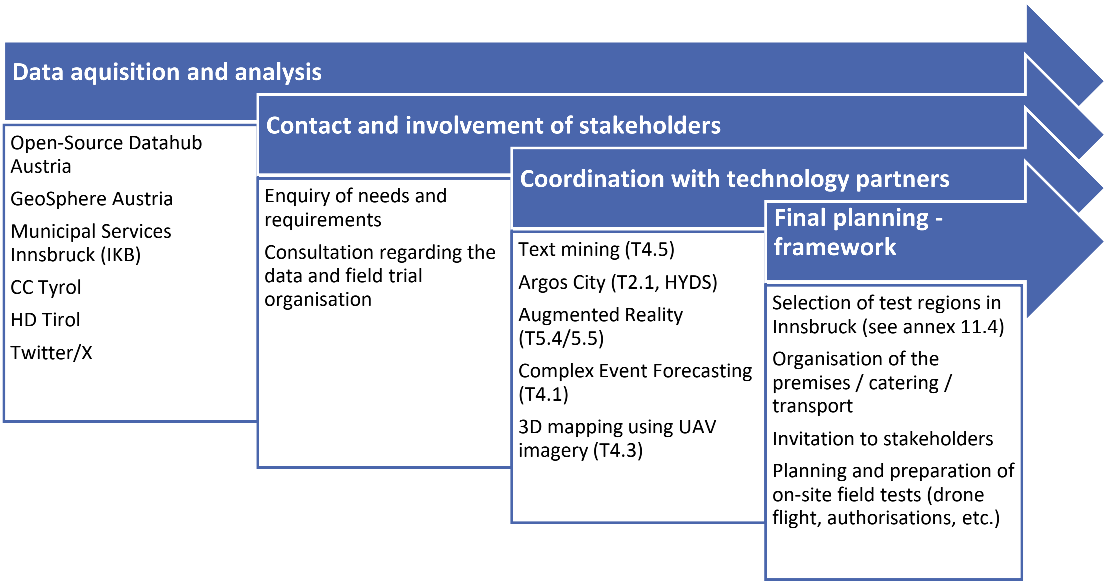


Figure 6: Workflow of trial preparations in Innsbruck

3.3.3 Description of pilot site Dortmund

During the field trial in Dortmund a selection of technologies have been tested, which implied different requirements to the environment as well as the methodology of the trial. Due to the nature of urban floodings, the EmCase team identified the DRZ as a suitable pilot site, where an outdoor area with debris could represent damaged infrastructure comparable to a large-scale flooding scenario. Moreover, its proximity to the densely populated area of Dortmund's city centre provided opportunities for the application of AR components with various points of interest, which are critical rescue targets in case of a flooding scenario. Also, facilities of FDDO have been used to present the various digital/online-tools of CREXDATA. All applications of subsystems have been observed or interactively attended by relevant stakeholders from the field of emergency response operations. Their feedback will further improve on the effectiveness and usability of all technologies.

Organisational settings

The hierarchical structures within emergency response organizations on-site are described in [D2.1]. They put up a structure in which technological systems as well as information can be more or less usable depending on both the actual user and her/her role. Therefore, the chain of command was considered in stakeholder selection for each Test Case (TC, see Section 3.2.2). In this process, an essential focus was the expected demand for information which flows not only upwards but also downwards along the chain. The targeted selection procedure was conducted by TC owners to ensure that stakeholder qualification matched a clear technology-to-requirement-fit from the start. That should imply the ability to collect higher quality feedback for upcoming project phases with corresponding further evolution of the technologies in WP4/WP5.

Stakeholders

During all test cases potential end-users from FDDO were present and observed or interacted with the subsystems. Depending on their role within the chain of command as either decision makers or action-planners, different parts of the trial were focused on them respectively. Therefore TC_002, TC_003, TC_004 and TC_005 were addressed mainly towards the B-Level fire officers being decision makers. On the other hand, TC_006, TC_013 and TC_014 were considered more relevant for C-Level officers due to their proximity to the scene and the responders during an emergency scenario. TC_008 was identified as an example for a technology, which provides information to be used on different levels of command depending on the goal of interpretation or the level of detail.

Geographical settings

The lab setting trials are held on the DRZ's premises at Rohdesdiek 32-34, 44357 Dortmund, Germany. The test cases are evaluated inside the DRZ's spacious "LivingLab" Hall (1300 m^2) and the adjacent outdoor testing area (1500 m^2).

Figure 7: Laboratory setting trials on the DRZ's premises

In order to identify suitable locations for exercises within the city of Dortmund, FDDO carried out a location analysis. The parameters for determining suitable locations for this use case were as follows: A buffer of 200 metres around schools, as these are locations with the most points of critical infrastructures and various heights of water during heavy rainfall. The following critical infrastructures were considered: nursing homes, schools, kindergartens, and hospitals. One location was prepared to be visited as part of the test with AR Head Mounted Devices (HMDs). Detailed descriptions are provided in appendix Section 11.3.

Test cases included in the test scenario(s) for the Dortmund pilot site

The test scenario was based on the same fundamental assumptions like in Innsbruck (see Section 3.3.2). Pre-conditions stated that heavy rain events could be observed in neighbouring cities, approaching Dortmund from Cologne via Bochum and hitting western edges of Dortmund (Lütgendortmund/Hombruch, passing the small river Emscher) and approaching the city centre. An increase of emergency calls was meant to be observable at the control centre of FDDO. The situation escalated by assuming emergent influences like heavier rain than expected, unexpected blockages in the sewer system, unexpected flooding of buildings, vulnerable and immobile people in these buildings. Table 6 correlates this scenario with test cases.

Table 6: Test Scenario TS_002: Dortmund

Test_Scenario_Name	Test Scenario Dortmund
Test_Scenario_ID	TS_002
Test_Scenario_Owner	FDDO
Test_Case_IDs	TC_002 ARGOS demonstration (TC_003 CEF sewer network)* (TC_004 CEF flooding)* (TC_005 Parameter exploration in urban flooding simulation model)* TC_006 3D mapping using UAV imagery TC_008 Text mining (TC_009 XAI based on biometrical data)* (TC_010 Visual Analytics of emergency protocols and precipitation data)* TC_013 AR routing TC_014 AR visualizing points of interest
Test_Item(s)	Several CREXDATA sub-systems: ARGOS-system, text mining, complex event forecasting, Augmented Reality (hardware/software-application), MIKE+, UAV, Visual Analytics

Test_Scenario_ Activities	• View situational map and available forecasts for an expected emergency → TS_002 → set emergency category and level
	• View text mining results → TC_008 → evaluate contribution to situational picture
	• View data analysis of incoming emergency calls correlated with rain gauges and precipitation forecasts → TC_009 → enumerate potential use cases resp. demands
	• Side-discussion based on Innsbruck data: Categorisation of a situation in terms of a) drainage relevance and/or b) emergency escalation → TS_003 → change setting of outlets, pumps etc.
	• View situational map of flooding situation (TS_001), view options in urban flooding models → TS_005 → enumerate potential use cases resp. demands esp. with respect to set of envisaged use cases (cf. D4.1)
	• Send/assume command to on-site commanders to explore the situation, augmenting the view by predictions from the system/simulation → TC_014 → providing feedback with recommended decision for action
	• Assume necessity to rescue people from a shopping mall, assuming a certain duration of necessary rescue operations in correlation with trapped vulnerable people
	  ○ exploring accessible entries/exits around the building → TC_006 → evaluate 3D representation, decide (best case) in favour of visible entry/exit
	  ○ exploring entry routes and, assuming delay for rescue operations, safe exit routes → TC_013 → evaluate system recommendations, decide (best case) in favour of recommended route
	• Assume firefighters having started the rescue operation, assume task to monitor their health status → TC_009 → enumerate potential use cases resp. demands to take actions based on detected events in biometric data instead of standard operating procedures, elaborate on requirements for explanations

Test_Scenario_ Environmental_Parameter	Derived from data used for Visual Analytics explorations

Detailed planning

The workflow / planning of the field trial in Dortmund was as follows (Figure 8).

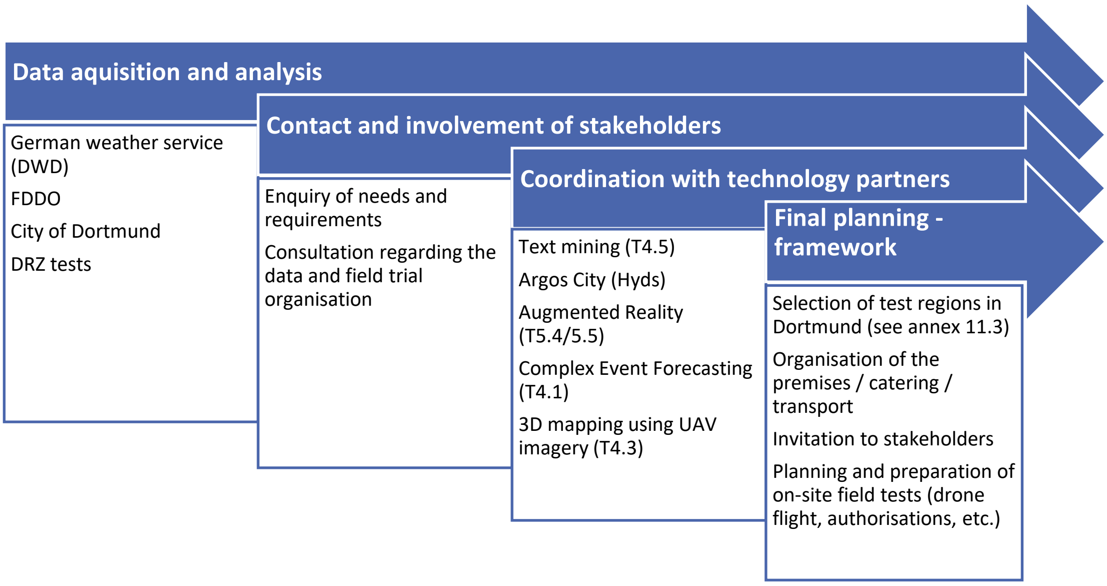


Figure 8: Workflow of trial preparations in Dortmund

3.3.4 Involvement of Finnish experts

The involvement of Finnish experts includes the Finnish Meteorological Institute (partner, FMI), the Ministry of Interior Finland (partner, MoIFI) and the Rescue Department of Helsinki (external stakeholder). In this first evaluation phase, the involvement is scheduled with two objectives: Firstly, experts contribute by collaboration in development and evaluation at both the German and the Austrian pilot site. Secondly, pre-studies are built around showcasing the use of machine learning in weather-related impact forecast.

Organisational settings

The Rescue Department of Helsinki is an external stakeholder of CREXDATA and represents the local, 'grass root level' actor, operating in Helsinki urban area, and in Uusimaa county. The Rescue Department of Helsinki employs around 700 experts distributed on nine fire brigades around the city of Helsinki ensuring the safety of the capital. The tasks of the national level and local level emergency management are very different. On the national level decisions are made on strategic planning, policy development and major national and international catastrophes. The local level rescue departments are responsible for instance on risk assessment, prevention of accidents and preparing for weather hazards on their own area. In weather-related hazards, both national and local level actors are dependent and base their decisions on accurate and modern weather forecasts, weather warnings and next-generation impact-based forecasts of FMI. The role of the Rescue Department of Helsinki in CREXDATA is to provide expertise, test the machine learning-based tools developed by FMI and possibly also other technologies in WP3-5, and bring end-user perspectives to the tools development.

Specific demonstrator components

FMI provides meteorological expertise in weather related hazards to ensure best possible interpretation and understanding on the weather information across the various showcases (cf. Deliverable D2.1). In addition to collaboration on the task in the different use cases, FMI is committed to develop machine learning models which are derived from the available impact data and real-time weather forecasting models used in FMI operational weather service in Finland. The model uses a Gradient Boosting machine learning technique originally developed in the national Finnish SILVA research project (2020-2023). The impact-based tools are developed in close cooperation between FMI, MoIFI and the Rescue Department of Helsinki, and user experiences are collected to develop and elaborate the machine learning model during the project time period. To date, during the CREXDATA project hourly municipality level impact data consisting of alert tasks, fire-fighting damage control, number of vehicle accidents, traffic accident number and fire-fighting missions in the field have been included into the machine-learning model.

Furthermore, regional forecasts have been extended during the project to a new municipality level product with new risk level predictions presented in **Error! Reference source not found.**. Also, previously the modelling used only the ECMWF HRES weather model but during the project a HARMONIE-AROME-based weather model system MEPS (Mesoscale Ensemble Prediction System) model has been added as predictor data. This high-resolution MEPS operational forecasts are produced every 3 hours and extend to 66 hours. Also, new parameters derived from the numerical weather prediction models have been added, which will better predict for instance the impacts of thunderstorms.

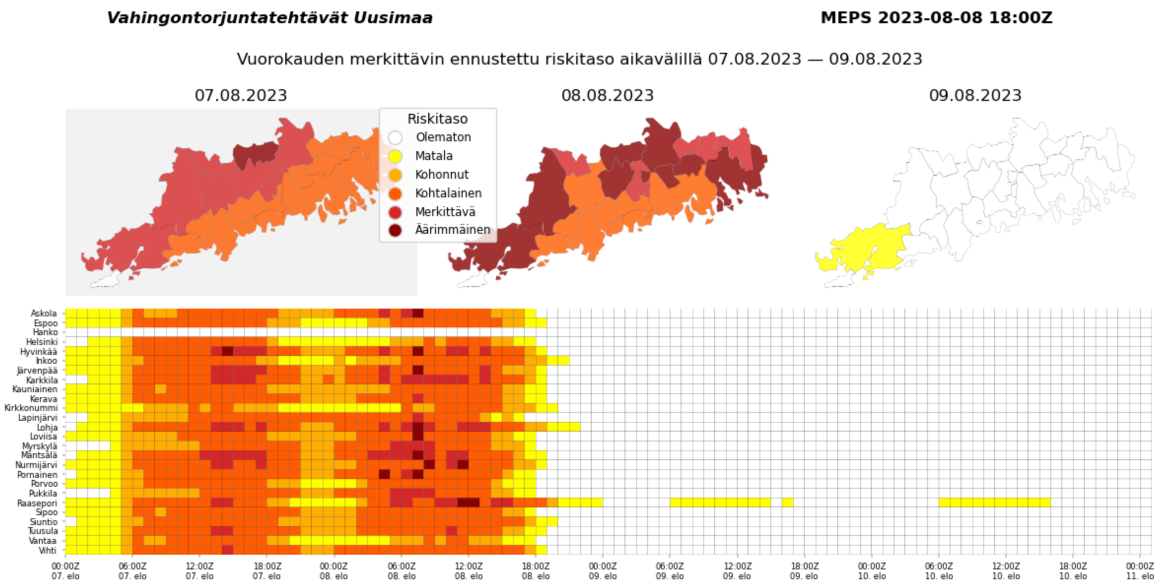


Figure 9: Gradient Boosting Impact-based model forecasts for damage clearance tasks in Helsinki and the whole Uusimaa region with different risk levels

Test cases included in the test scenario(s) for the Finnish involvement

Compared to scenarios TS_001 and TS_002, which are focused on the impacts of one specific weather hazards to eliminate too many influencing variables in the tests, the scenarios of the Finnish showcase were already built on the impacts of different weather hazards. The scenarios of the Finnish showcase are built rather on the impacts of different weather hazards than one specific weather hazard. In Finland a variety of weather-related hazards are experienced throughout the year. In winter months, for instance, strong windstorms and difficult road conditions are usual. In contrast to that, during the summer months, Finland experiences in increasing frequency extreme heatwaves and dry weather conditions, which increase the risk of wildfires. Severe thunderstorms with strong wind gusts or heavy rainfall occur with varying frequency from year to year. The Rescue Department of Helsinki assists in clearing damage caused by storms. The most common damage caused by storms includes fallen trees, torn roofs and water damage caused by heavy rainfall and broken pipes.

The verification of the Finnish machine-learning model will be mostly based on careful cross-validation and has been previously cross-validated with data from 2002-2021 and separately from 2022. The model performance will be monitored actively, emphasizing on extreme events. After a testing phase, the performance results will be collected for the analysis of the overall skill using appropriate metrics. The Gradient Boosting method has been found reliable and accurate in several different applications in previous research projects, thus it was chosen also for the main method in the Finnish showcase.

Table 7: Test Scenario TS_002: Finland (esp. Helsinki)

Test_Scenario_Name	**Test Scenario Finland (esp. Helsinki)**
Test_Scenario_ID	TS_003
Test_Scenario_Owner	FMI
Test_Case_IDs	TC_015
Test_Item(s)	Gradient Boosting machine learning technique
Test_Scenario_ Activities	• Present output of new risk level predictions to stakeholders related to Fire Departments → TC_015 → enumerate potential use cases and demands on practises how to use the tools • Present output concerning pre-hospital service operations to stakeholders related to rescue services → TC_015 → enumerate potential use cases and demands on practises how to use the tools
Test_Scenario_ Environmental_Parameter	• different weather hazards (not one specific weather hazard) • dependencies of various types of pre-hospital medical emergency service operations (ambulance calls) in Helsinki → The study is integrated with the Health Use Case and employs a statistical regression model DLNM (distributed lag nonlinear model).

3.4 Simulation tools

In the EmCase Use Case "Simulation", there is a variety of use cases for simulations[1] spanning from reproducing data from archives to identifying optimal parameters targeting a desired situation (see Figure 10, based on Deliverable D2.1, pp. 28-29 and EC_UC_40]).

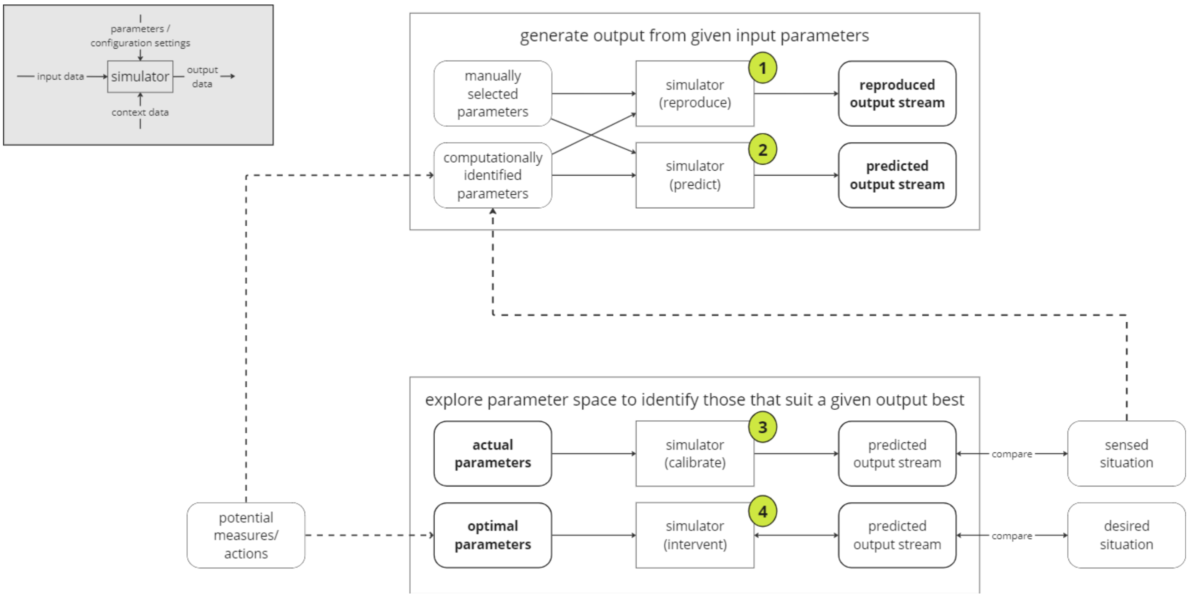


Figure 10: Classification of use cases for simulators in the EmCase

3.4.1 Urban flooding simulator

In the first half of the CREXDATA project, a focus was set to "weather related simulation" (cf. Deliverable D2.1, p.28) providing research context for both T4.2 (Interactive Learning for Simulation Exploration) and T5.4/T5.5 (Augmented Reality). Anchored within the concept of a Simulator HyperSuite for the EmCase (see **Error! Reference source not found.**), MIKE+[2] was selected as a simulator engine. MIKE+ is a representative sample of similar simulators, also preparing for different natural phenomena like forest fires. Even though it is not open source, it covers the requirements of WP4. It was selected due to priorities of end users. The selection process involved further alternatives like visdom, Hydrologic Modeling System (HMS), Storm Water Management Model (SWMM, open source). MIKE includes functionality also for river modelling, available also in products like River Analysis System (RAS) and XPSWMM (open source) . It can be coupled with further models like the Life Safety Model (LSM).

---

[1] "A simulation imitates the operation of real world processes or systems with the use of models. The model represents the key behaviours and characteristics of the selected process or system while the simulation represents how the model evolves under different conditions over time." [https://www.twi-global.com/technical-knowledge/faqs/faq-what-is-simulation]

[2] MIKE+ documentation: URL https://manuals.mikepoweredbydhi.help//latest/MIKEPlus.htm, last access 20.05.2024

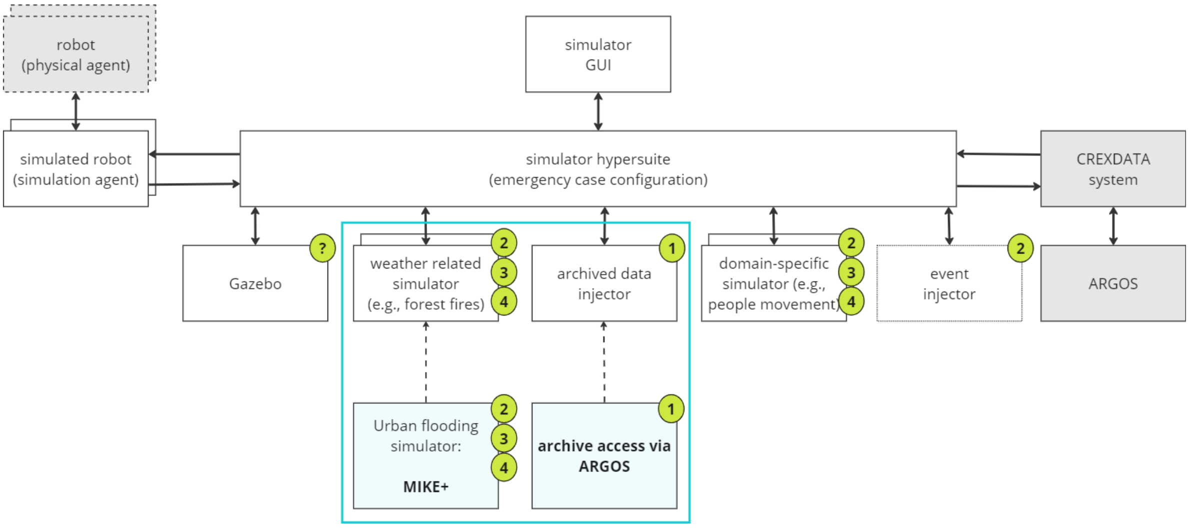


Figure 11: EmCase Simulator HyperSuite: focus on urban floodings in M1-M18

Typically, MIKE+ is not used in online but in offline use cases: Public Protection and Disaster Relief Agencies (PPDR) use it in preparation phase to simulate potential "what-if" scenarios for their area of responsibility (cf. flood risk map for the city of Dortmund in Deliverable D2.1). Such risk maps are simulated for scenarios based on probability of certain events, such as a 50 years' flooding up to a 200 years' flooding event. Based on simulation outputs, response plans are prepared. In case of an expected extreme weather event, such plans are activated prospectively. In the event itself, it is of major importance for decision makers to understand whether the event confirms the expected simulation or, more pragmatically, whether the chosen response plan is the appropriate one or needs to be escalated.

The software is available with a research license, offering an API to parametrize and invoke simulation runs[3] and access output results[4]. Specific use cases are detailed in Deliverable [D4.1]. While ARGOS allows reproduction of data (use case type 1 presented before in Figure 10), urban flooding simulation covers the other three use case types:

- Prediction (2) applies the urban flooding model with a given parameter set to create an output in the sense of predicting a potential future situation (here: flooding of streets, buildings etc. over time) which is typically visualized for users
- Calibration (3) seeks to estimate the model parameters based on an observed situation to make the error between these observed values and those predicted by the model as small as possible (e.g., to verify that premises for a decision which were derived from prediction at an earlier point in time were actually valid)

[3] MIKE+ Py enables esp. the modification of the MIKE+ database stored in .sqlite from python (geometry data in that database stored with the SpatiaLite format). See URL https://github.com/DHI/mikepluspy, last access 17.05.2024

[4] MIKE IO enables common data processing workflows for MIKE files that are compliant with DFS and res1d in Python. URL https://github.com/DHI/mikeio, last access 17.05.2024

- Intervention (4) means that parameters are explored to identify those that are most likely leading to a desired resp. expected situation like “minimum average water level in a city” for deciding on corresponding response measures

As an example, a type 4 use case (intervention) is that decision makers try to understand under which conditions (i. e., parameter values) a certain situation occurs that was prepared in advance. For instance, a city like Dortmund or Innsbruck sets up response plans in correlation with the likelihood of events. Such an event might be, e. g., a “100 years flooding” (flooding that is likely to appear every 100 years, cf. Deliverable D2.1). To activate such a response plan, decision makers need to understand whether the current emergencies is actually a “100 years flooding” – or even worse.

3.4.2 Interfaces of the urban flooding simulator

Context data for simulations in MIKE+ (see Figure 12) are given in terms of

- grid data: Basic data is required in terms of bathymetry or, respectively, DTM plus DSM (publicly available at https://www.data.gv.at/, cf. Section 3.3.2**Error! Reference source not found.**).
- network data: The sewer network of a city can be modelled by nodes like manholes or pumps, and links like pipes. Nodes are modelled by coordinates, while links are modelled by references to nodes.

MIKE Powered by DHI software utilizes sqlite for specifying input data. For output, res1d and the DFS (Data File System)[5] are used, including data formats for scalars (dfs0), vectors (dfs1), matrices (dfs2) and unions of these three (dfsu).

MIKE+ provides a huge amount of parameters to be set for a scene. Examples are

- attributes of nodes (e.g., manhole) and links (e.g., pipe) like “infiltration_to_node”
- precipitation over time
- wind speed/direction
- possible discharge/evaporation of spatial areas

Precipitation and wind conditions are examples of parameters that can be set by actual measurements from weather stations or based on forecasts from weather models. Discharge and evaporation are examples of parameters that need to be derived from actual or forecasted data like temperatures and drought indices. Basic data for all of these examples is provided by the ARGOS system, covering measurements, forecasts and archived data from the past (see “archived data injector” in Figure 11).

[5] URL https://docs.mikepoweredbydhi.com/core_libraries/dfs/dfs-file-system/, last access 16.05.2024

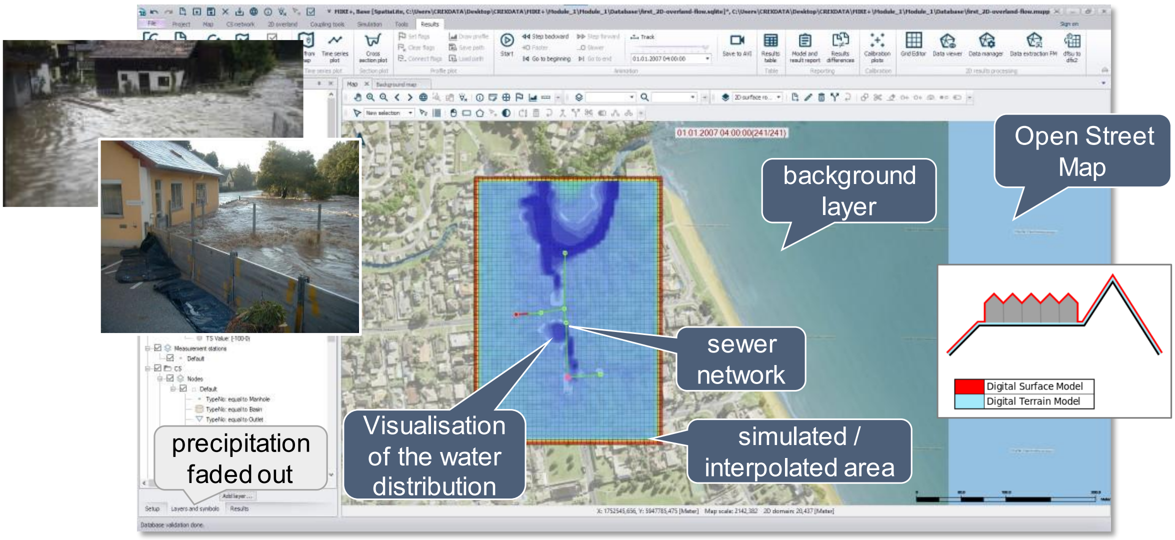


Figure 12: Sample MIKE+ project (photos: Innsbruck; screenshot: DHI)

To efficiently invest efforts in CREXDATA, the focus was set on value-adding activities integrating the simulator instead of investing into modelling a real city like Innsbruck or Dortmund with bathymetry, sewer network etc. The conceptual software architecture for the integration of MIKE+ (Figure 13) covers the interface to the stream-based CREXDATA system, the stream-enabled ARGOS system invoking simulations and pushing simulation results, as well as a sample deployment of MIKE+ extended by a batch process. This batch process starts a simulation run, uniquely identified by a simID, converts output results from DFS formats into those formats required by T4.2 and T5.4/T5.5, and merges all relevant result packages (esp. grid-related and network-related simulation results).

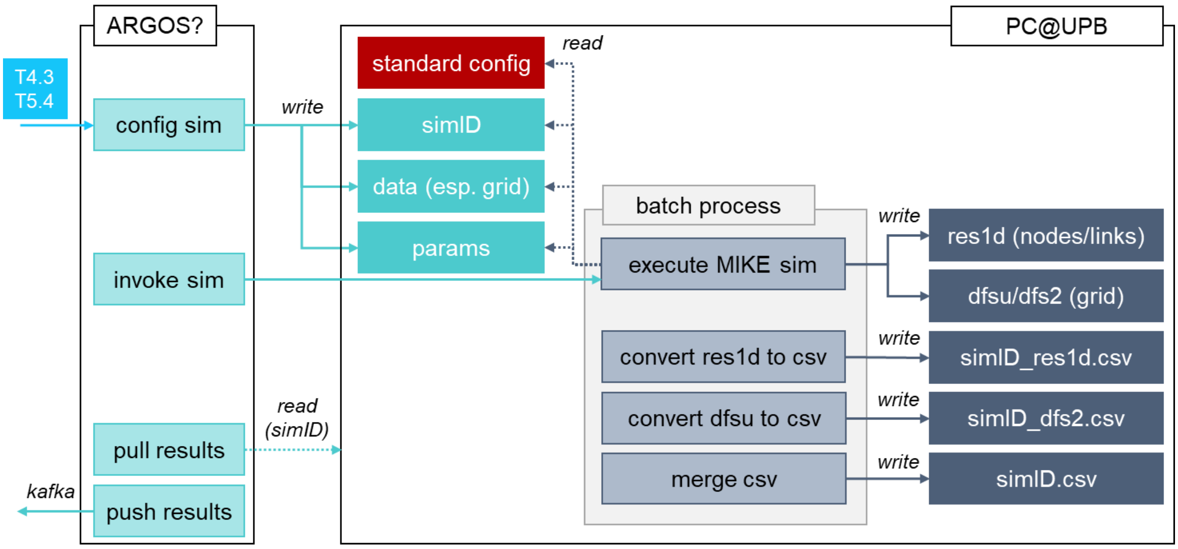


Figure 13: Conceptual software architecture for the integration of MIKE+

3.5 Initial Use Case Evaluation

In the weather emergency case, the impact evaluation is enabled by a correlation of injected incident data and observed behaviour of test personnel, captured by mobile observatory labs. KPIs are (see DoA, p.8 and Deliverable D2.1, p.89):

a) 80% accuracy in critical event prediction in test bed scenarios,
b) System Usability Score (SUS) of interactive exploration tools for XAI above average,
c) SUS score of uncertainty visualization (in AR and control centre) above average,
d) observable impact of system’s output to action-planning in 90% of injections,
e) even under perceived risk, users follow system’s advice in 80% of situations.

The first evaluation phase is characterized by trials of early prototypes of a sub-set of the various technologies in the project. The aforementioned KPIs are not yet achievable and even realistic to be validated, but the trials seeked to find early indications especially towards system usability (KPIs b/c) and observable impact on action-planning (KPI d). For the latter, subjective feedback on utility and future use was collected. Section 3.5.1 summarizes the overall settings, while Section 3.5.2 documents results. The main intention is to formatively provide the basis for planning the second half of the CREXDATA project. Careful overall interpretations are provided in Section 3.5.3.

3.5.1 Description of field trials

As part of the first half of the CREXDATA project, the initial project results have been evaluated through field tests in the EmCase at the pilot sites in Innsbruck and Dortmund. These field trials were specifically organized to test the effectiveness and practicability of the developed technological solutions in the context of pluvial floods in urban areas. Representatives of the fire departments and local authorities were invited as experts for these field tests. Their task was to critically evaluate the current state of development of the technologies and test their applicability in real flood situations, providing feedback for further development. The selection of these experts was based on their extensive experience in disaster management, a certain understanding of security research and their in-depth knowledge of local conditions and challenges.

The EmCase Field Trial comprises two test scenarios for Innsbruck resp. Dortmund (see Section 3.2). Figure 14 below is the detailed agenda of the field trials conducted at the pilot sites in Innsbruck and Dortmund.

	Innsbruck		**Dortmund**	
time (CET)	**Monday 2024-06-10**	**Tuesday 2024-06-11**	**Thursday 2024-06-13**	**Friday 2024-06-14**
9:00 AM	**T4.5** Text Mining	**T5.4/T5.5** Augmented Reality in the field **T2.1/T4.3** UAV flight and NeRF [including travelling to the test site and back]	**T2.1/T4.3** UAV flight and NeRF	**T4.5** Text Mining
	T2.1 ARGOS demonstration		**T5.4/T5.5** Augmented Reality in the field + Evaluation of Augmented Reality	**T2.1** ARGOS demonstration
	T5.4/T5.5 Introduction into Augmented Reality	**T2.1/T4.3** NeRF – background information about the technology and evaluation/workshop	**T2.1/T2.4** MIKE+ Simulation **T4.2** Parameter prioritisation	**T5.3** Visual Analytics (results of emergency call protocols)
	T4.1 Complex Event Forecasting of sewer network data	**T2.1/T2.4** MIKE+ Simulation **T4.2** Parameter prioritisation	**T2.1/T4.3** NeRF – background information about the technology and evaluation/workshop	
		T5.4/T5.5 Evaluation of Augmented Reality		
5:30 PM		*Presentation of activities at Fire Brigade TU Graz*		

Figure 14: Trial agenda of the EmCase

3.5.2 Results of trials

The individual test cases each consist of the test, the evaluation by the experts and an in-depth discussion between the end users and the relevant technology developers of the CREXDATA consortium. In the technology-related discussion, the focus is on identifying potential for improvement and drawing up recommendations for future development work. These evaluations are crucial to ensure that the technologies developed are not only theoretically sound, but also practically applicable and effective in supporting emergency services and authorities during pluvial floods.

3.5.2.1 Text Mining

During the Field Trial of the EmCase in Innsbruck and Dortmund, an application was demonstrated to assess the language model's effectiveness in detecting relevant social media posts and the question answering model's ability to answer queries about an event, aligning with T4.5 Text Mining objectives. To assess the performance of the system in terms of attributed relevance, the practitioners red manually through the detected posts to determine if the posts were relevant to a flooding event or not. The demonstration featured two sections: one where experts identified misclassified posts to calculate the language model's accuracy, and another where they queried relevant posts to evaluate the coherence of the QA model's responses. Practitioners working in the emergency control centre and dispatchment as well as firefighting took part in Innsbruck. In Dortmund the tool was presented to experts of different hierarchy level of the fire department.

In Innsbruck two data samples were evaluated: an English sample with 30 tweets from a 2013 flood in Alberta, Canada, and a German sample with 31 tweets from a 2021 wildfire in Austria. The experts evaluation were as follows: The English test data set has 8 incorrect and 22 correct classifications, which corresponds to an accuracy of around 73%. The German test data set has 11 incorrect and 20 correct classifications, which corresponds to an accuracy of around 64%. For the QA model, response was coherent and in accordance with the relevant posts. Two LLMs were evaluated for question answering, Mistral 7B and Phi-3 mini. Both models performed well in answering and generating English texts, but Phi-3 mini performed better in German.

In Dortmund two data samples were evaluated: an English sample with 30 tweets from a 2013 flood in Alberta, Canada, and a German sample with 30 tweets from a 2024 flood in Passau, Germany. The experts evaluation were as follows: In the English data set, there were 5 incorrect and 25 correct classifications, which corresponds to an accuracy of around 83 %. In the German data set, there were 15 incorrect and 15 correct classifications, which corresponds to an accuracy of around 50 %. For the QA model, response was coherent and in accordance with the relevant posts. Due to time constraints, only the Phi-3 mini model was evaluated.

The experts were concerned about properly defining the criteria for tweets to be selected as relevant. They were especially interested in tweets that contain information that will lead to early detection of an emergency event to aid rescue operations. Experts were also interested in the possibility of improving the classification model's accuracy in real time with input from a human (e.g., online learning). They were also interested in displaying the relevant posts and their attached images in real time as they are detected, which will ultimately be done in a user interface like the ARGOS platform. The experts also mentioned the possibility of having a warning level alert to inform authorities of the need to respond to panic on social media platforms (for example if there's a lot of tweets in a short time regarding a certain incident), and the possibility of detecting false information in relevant posts. They further advised on a future application of analyzing images to extract more information. Concerns were raised in ensuring that the QA model is explicit of its uncertainty when providing information that it is unsure of. Finally, when asked if the German answers provided by Phi-3 mini were grammatically correct, they answered affirmatively.

3.5.2.2 ARGOS

To evaluate the ARGOS system, the experts were shown a demonstration and had the opportunity to test the platform by themselves. The evaluating practitioners in Innsbruck were again the emergency control centre and dispatchment as well as firefighting. In Dortmund the expert were firefighters. The ARGOS presentation included its main ideas, the data that is included and how it is used at the moment. The use case of Innsbruck was showcased afterwards with different functionalities, e.g. vulnerable elements, thresholds, logical rules for warnings and how the various data are or can be displayed. The experts had various experiences with map-based incident management systems and appreciated the visualization of the tool as well as the possibility to add different information in different layers. The end users were given an introduction to the ARGOS system and were impressed by its ability to provide accurate and up-to-date data on extreme weather situations. In particular, the fact that the platform bundles all relevant information was found to be extremely useful. It was found that the system can only be fully utilised if additional sensors and weather stations are installed. The installation of various sensors at key nodes and critical infrastructure needs to happen first in order to realise the full potential of the system. Furthermore, an expert would need to check the relevant sensors linked in ARGOS in order to specify the threshold values for the risk assessment. Despite these preparatory measures, the ARGOS system was received very positively by the emergency personnel. The system is now to be presented to the higher-ups as great potential has been recognised. They could foresee the ARGOS system used in a wide range of scenarios, especially for forecasting of weather phenomena and rescue planning for rescue missions. Further improvements could be made in the possibility for reporting a new incident and using the user role definition for different access levels in terms of information showcased in the system.

Complex Event The sewer network data of Innsbruck was analyzed by the technology partner for task T4.1. The demonstration and presentation of the results took place in a hybrid setting with the main expert who is representative of the communal services of Innsbruck. The goal of the event forecasting is to detect once the water in the sewer system rises above a certain threshold to inform the decision-maker and trigger a decision process whether to open valves to discharge water out of the system or other related decisions. This session was discussion-based and allowed for an open exchange between expert and technology partner. Questions arose concerning specific sensors of the sewer network that should be included in the forecasting system and which events should be detected as it is difficult to find incidents in the existing data which makes it hard to detect relevant events. The sewer network is a complex system and critical measuring points for a forecast of overflow are missing which makes the forecasting particularly challenging. A solution of this problem was proposed by measuring the water level in the sewer network at its lowest point. Other potential use cases were discussed, such as predicting the water level of the river Inn because it is possible that water from the river enters the sewer system and causes an overflow.

3.5.2.3 Augmented Reality

The session for Augmented Reality was split into different parts. First, an introduction into the technology and its workings was given by the technology partners to the participating experts, followed by a field test. The field test in Innsbruck was carried out in downtown Innsbruck (left photo in Figure 15). The field test in Dortmund was executed at the DRZ facilities (right photo in Figure 15). The calibration process of the AR system was explained as well as how points of interest such as schools, hospitals, or train stations can be shown and used in the AR. The system furthermore includes different colours that can indicate danger levels for the points of interest and facilitate priorities for rescue, as well as water visualization for flooding events on different time scales. The distance to the POIs is shown by aerial lines under the icons. At the same time, the AR glasses allow access to the city map with the option of adjusting the scale as required. The POIs are also displayed on the overall map. In addition, the AR glasses can display the potential water level on site and visualise the course of the water level over the day to examine the effects of changing conditions over time. The experts testing out the system in an urban environment in Innsbruck were fire brigade officers, fire fighting volunteers and representatives of the State Warning Centre who had never used an AR application for mission-related purposes before.

Figure 15: test case of Augmented Reality in the field with experts in Innsbruck (left) and Dortmund (right)

The participants particularly liked the additional, easy to use information provided by AR, the water level visualization. The downsides of it were marked by the state of the art and use of AR in general, not the CREXDATA application in particular. Doubts were issued in terms of the additional mental demand because of the detachment of the “real world” and the limited contact to the other forces in the field (e.g. colleagues) and the current status of technical equipment (not waterproof, expensive, not yet compatible with protective equipment). Figure 23 presents qualitative feedback regarding AR functionalities. Test users rated those functions very positively that concern geo-references information in an operation, displayed in terms of POIs (left diagram). Similarly, but a bit more reserved, functions regarding routing and orientation were rated. In discussions became clear that there is a strong bias from a) operational experiences that users had in mind, and the ability to consider AR in future operations out of a real operational environment.

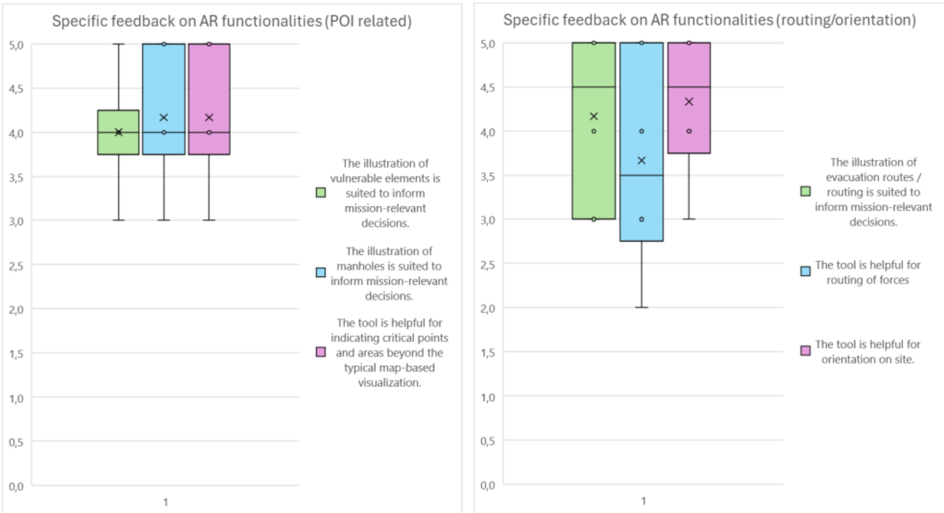


Figure 16: Qualitative feedback on AR functionalities regarding the applicability in emergency response

3.5.2.4 NeRF generation using images from UAVs in combinaiton with Federated Learning

In this section, the outcomes of the structured interviews carried out with the experts in Innsbruck and in Dortmund are summarized, addressing current use of UAV imagery and 3D modeling with traditional tools, and assessing the potential benefits of novel features enabled by using a machine learning-based approach, in particular Neural Radiance Fields (NeRF) combined with Federated Machine Learning, developed in T4.3. The experts in Innsbruck and Dortmund are all using 2D drone imagery (i.e., (ortho)photos and videos) regularly or frequently for a variety of purposes, including situation overview, aerial reconnaissance, search for persons, detection of embers in forest fires, monitoring bridges in floods, mission planning, risk and damage assessment, documentation. Images are typically analyzed by drone operators and the incident commander on site. Images are not uploaded to a situation awareness system. The experts in Innsbruck stressed that decisions regarding passing images/information to higher command levels should be made on-site based on operational needs. Currently images are not stored at all or stored locally by the individual fire departments. The issue of centralized storage for nationwide access, which would be desirable for research and analysis, was discussed by the experts in Innsbruck. Privacy protection would be an important factor to be taken into consideration, especially when images contain identifiable people. 3D models have also already been used by the experts. The Dortmund experts have expressed a rather reserved attitude, based on concerns about reliability due to missing objects or potential artefacts, as well as long computation times. The Innsbruck experts use 3D models more extensively and have experience with various tools. Individual opinions include: “It works well, many details, cloud density is often too big.” and “It is complicated but powerful.” Added value is seen for visualizing slope gradients for water flow direction in flash floods or thermal lift in forest fires; assessing terrain accessibility in mountainous regions during forest fires; monitoring the break-off edge in flooded areas; inspection of building collapses; interior and exterior structure surveying; detailed planning and simulation (what-if scenarios).

Generally, both model computation time and quality are important: information is needed quickly and must be reliable. Some scenarios, such as flash floods, are developing very fast and need immediate situation awareness. Other scenarios, for example river flooding or forest fires, allow more time for decision-making. Slow-developing scenarios like forest fires need evolving models over time. Time is seen as absolutely crucial for saving people, model quality is more relevant for saving property, cleaning operations and coordination. The experts have varied opinions regarding the question whether an iterative approach would be useful, i.e., whether it is useful to first get a 3D model of lower quality fast(er) and then gradual improvements over time. Some said that it would be useful to start with a quick, rough model for initial decisions and refine it over time for further decision-making. Others stressed the danger of errors based on an initial imprecise model and prefer waiting for a better model. The dilemma is “Do I prefer to make a wrong decision quickly or a right decision slowly?”. The current low overall quality and also the slow rendering of the NeRF-based 3D model are a big concern for the experts and need to be improved in the future. All experts agreed that it is important to have a clear indication which areas of the model are reliable (generated from sufficient data, high confidence) and which are not. Some experts would prefer gaps in the model (as is the case, e.g., in the WebODM tool), others can see the usefulness of guesses in the model (such as produced by the NeRF approach) as long as (un)certainty is properly indicated. It was stated that such “completed” models provide more visual clarity, where the notion of “completed” should be conceptualized in T5.4-T5.6.

The possibility to integrate multiple small(er) models was discussed for two use cases: One, a large incident area is divided into segments, each segment is mapped separately by a drone flying in the segment. Combining the segment-models into one global model is considered of limited use for decision-making on the spot, because large incident sites are typically handled segment-wise. However, such an integrated model is considered useful for higher levels of command and for documentation and/or training purposes. Two, a large area is mapped by several drones flying simultaneously in the same area, each collecting only some of the pictures. Since this reduces flight time and enables faster data collection, the experts agreed that it would be useful. They were somewhat concerned whether the computation time benefit would outweigh the technical effort. They also pointed out organization challenges due to having to coordinate the flights, although it was agreed that this could be alleviated by drones flying at different heights.

Figure 17: UAV flight for NeRF generation in Innsbruck

3.5.2.5 Visual Analytics

The results of Visual Analytics was presented as part of the field trial. The analysis was made based on a data set of the city of Dortmund. The aim of the tool is to find out which precipitation conditions lead to an increased personnel and operational load. This should ensure that incident commanders can find out how many personnel they would need in which weather conditions in order to be able to plan accordingly in advance. On the one hand, the deployment data from 2021 and the weather data from the same year were analysed. The deployment data from 2021 contained information on whether a deployment was associated with heavy rain, the required number and qualifications of personnel and the date and time of each deployment. The precipitation data was provided by two different weather stations. On the one hand, the weather station in Dortmund city centre (Brückstraße), which provides precipitation data with a ten-minute time resolution. Another weather station in Waltrop, on the other hand, provides additional information on the daily precipitation total for the same period. Initial interim results have shown that some periods of increased precipitation were accompanied by an increased number of deployments, but not in all cases. It was also hypothesised that successive periods of moderate rainfall also led to an increase in deployments. The end users assessed the information very interesting, especially as the results could provide information on preventive deployment planning. In this way, emergency personnel can be bundled and coordinated in advance. However, the problem here was that, that it was not yet possible to provide precise results on how much personnel is needed for which precipitation totals. There is currently a lack of comparable data to exclude outliers and to ensure that the results are representative and can therefore be used for deployment.

To solve this problem, one expert suggested to use a different and bigger data set to be able to combine the data for further analysis.

3.5.3 Evaluation of trial results

In general, the trials were conducted with technologies in very different maturity levels. While some technologies were tested hands-on by end users (AR, NeRF, ARGOS), others were explored by presenting their capabilities in the context of described test scenarios (Text Mining, CEF, simulation exploration, Visual Analytics). Therefore, all evaluation sessions were finalized by a prospective estimation of end users inhowfar the individual technologies and especially the integrative CREXDATA system would useful and usable in emergency response, considering use cases of action-planning and decision making. The following sections provide insights into end user's feedback, which will be very valuable for setting clear a scope for M19 to M36. Such planning sessions will take place in June and July 2024.

3.5.3.1 Usability evaluation incl. System Usability Score (SUS)

Even though the CREXDATA system was not yet available as an integrated system, applications like ARGOS, AR and NeRF in combination with demonstrations and initial tests made it possible to engage with end users very constructively. Figure 18 (a) shows the perceived usability of the system, represented by a SUS of 57.8 points. While the objective for the end of the project duration is much higher, this score is acceptable considering the current state of implementation. Specific answers to questions show that end users clearly distinguish between currently implemented features and the high potentials of the envisaged system, which was expressed several times by several experts. Figure 18 (b) is one of these confirmations, strengthening current directions of CREXDATA research : Even though we assume that AR hardware means restrictions on usability, thinking of integrated AR helmets, AR functionalities were rated very positively. With regard to NeRF, Figure 18 (c) shows that there is a clear correlation with the quality of outputs. Following the rendering of 3D objects from UAV photos, it is obvious that end users can distinguish the quality provied by an enlarged number of images. In its largest test setting in NeRF studio, the approach received very positive feedback at least from many of the experts.

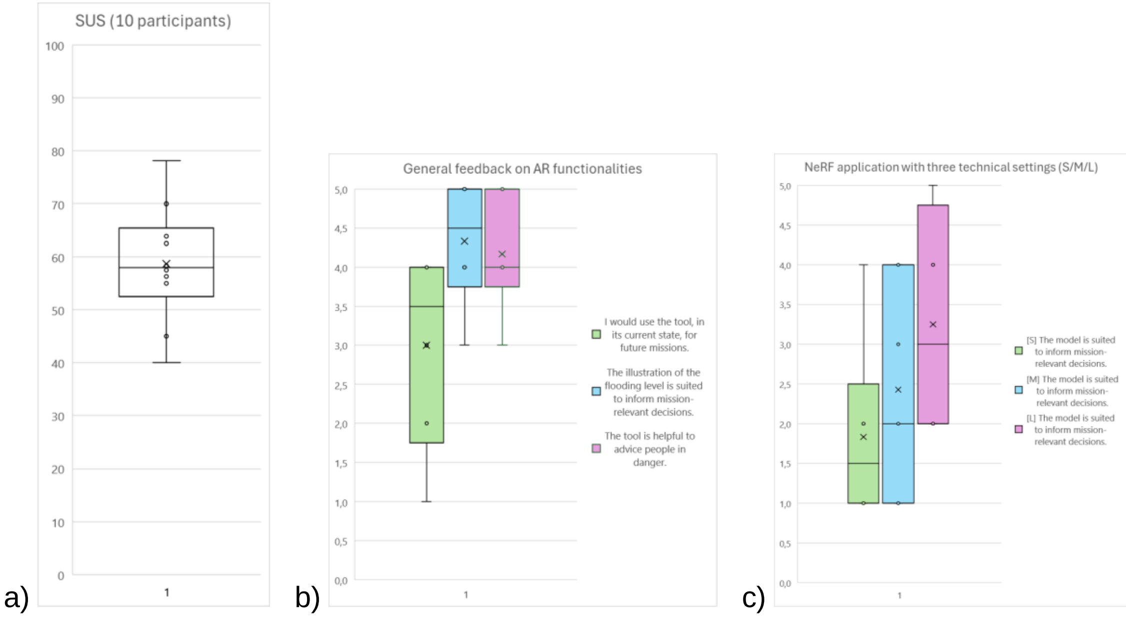


Figure 18: SUS of the CREXDATA EmCase demonstrator system (box plot (a)), general feedback on AR functionalities (box plot (b)), and feedback on NeRF applications showing an increase of perceived utility depending on the richness of NeRF outputs (box plot (c)).

3.5.3.2 Feature requests for core demonstrator system components

ARGOS is used both as data source, mediation component and Graphical User Interface. Utilizing the system as an established platform in connection with new CREXDATA technologies, fundamental functionalities were rated very positively by end users. Figure 19 presents three additional features that end users considered beneficial. Here, the command levels influence results as only higher levels are using geo-information system (GIS) in an extensive way. Communication is often conducted in different systems, not directly coupled with GIS.

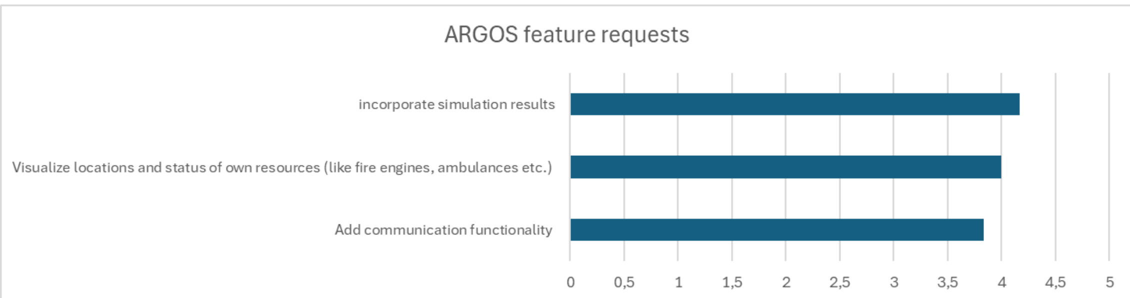


Figure 19: ARGOS feature requests

In terms of potential scenarios where the AR application could be useful, especially flooding was mentioned where the AR could support decisions in object protection as well as the protection of personnel if the location of manholes as points of interest could be included in a flooding situation. Additionally, it was pointed out that the AR in combination with drone footage could be useful in the visualization of hotspots and embers during wildfire that are difficult to spot from the ground. It was also suggested that it would be useful to be able to manually rank the POIs according to the risk assessment in order to visualise the importance of the more vulnerable POIs. It was also suggested that pop-up windows for further information from the POIs should contain a link to the property plans of the respective facilities. Despite these challenges, end users in Dortmund and Innsbruck felt that with further improvements and customisation, the technology could become a valuable tool for use and see strong potential for future use. The constructive and encouraging perspective is visible in feedback on potential features which are considered in the EmCase (Figure 20).

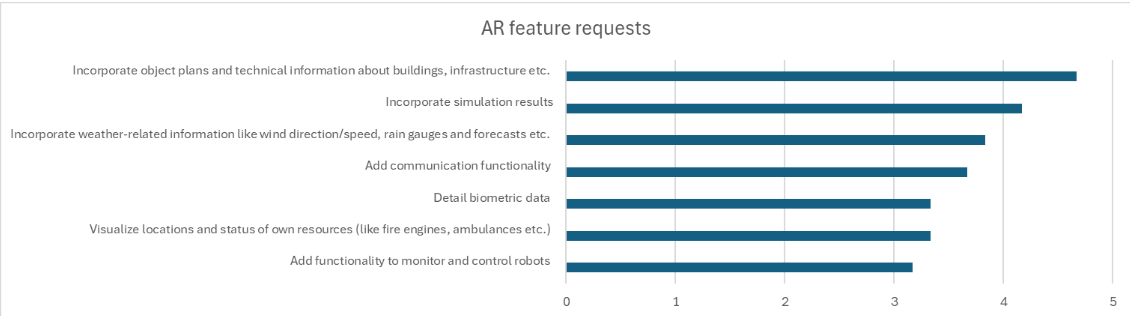


Figure 20: AR feature requests

Regarding use of Federated Machine Learning, implemented by the NeRF approach, the experts in Innsbruck discussed the usefulness of model modification, which the NeRF-based approach makes possible, although only modification of objects, e.g., human faces, and not scenery has been shown in research demonstrators so far. This potential future feature has been found very useful, especially in combination with simulation/prediction, that is, visualizing possible future developments of the incident, for example, a flooding of an area. The discussion highlights the potential of 3D modeling in enhancing incident response missions, particularly in planning and documentation. Real-time operational value is limited by technical complexity and the need for accurate, timely data. Centralized image storage and cautious incremental model development are recommended, ensuring operator oversight on model reliability. Operational tasks require a balance between training time and the quality of the 3D reconstruction. Clearly, the optimal result would be achieving both objectives. The technology developed in T4.3 accomplishes this by distributing the image collection and model training processes across multiple computational nodes (drones or servers) and parallelizing them. This approach also reduces communication costs, as only NeRF model's variables, rather than raw image data, traverse the network. Our current goals are to achieve optimal quality in our distributed experiments, create a deployable version of the solutions developed in T4.3 and make use of communication-efficient algorithms that we have developed (Federated Dynamic Averaging), to further reduce communication between nodes. Feedback by end users (see Figure 21) needs to be considered very carefully as utility and usability depend highly on the objects and environments that are observed.

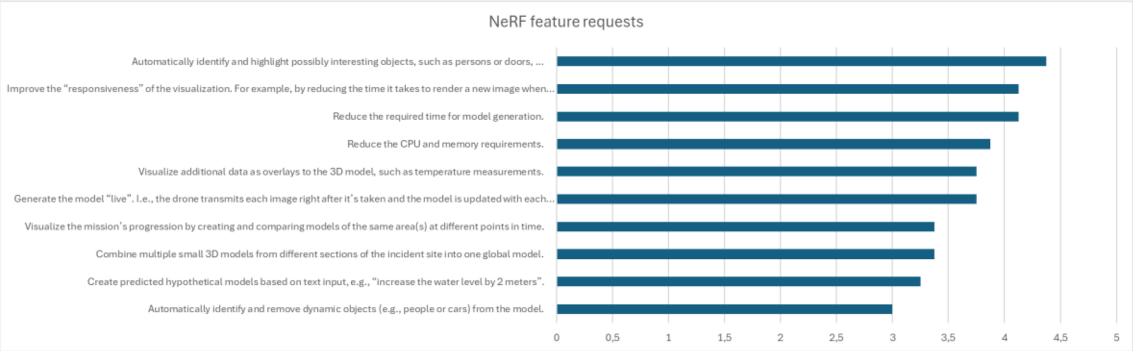


Figure 21: NeRF feature requests

3.5.3.3 CREXDATA objectives

Regarding overarching CREXDATA objectives, only indications could be gathered to understand the current state and the end user's view on potentials of the technologies. Rating these objectives on a scale from 1 to 5 (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful), outcomes of the trials are:

- End users confirmed the potentials of being able to have extreme-scale data ingestion/generation, fusion and exploitation as "quite useful". With regard to different types of media, ARGOS, especially AR based on simulation feeds and NeRFs rendered from UAV imagery were considered beneficial.
- Ingesting multimodal data (images, simulations, social media publications, etc) was mainly demonstrated by text mining, considering integration with, e.g., ARGOS. Users were very eager to discuss use cases, and confirmed the technology clearly to be "very useful". Handling multilingual social media data in real-time as such was seen "quite useful".
- Technologies like CEF could not yet be demonstrated in a hand-on fashion. Thus, the ability of using dynamic modelling to predict the systems' behaviour, having real-time predictive knowledge and forecasts as well as having multiresolution complex event forecasting under uncertainty could not yet be evaluated in a reasonable manner.
- Using online federated learning is kind of hidden being embedded within the NeRF technology. Thus, feedback is indirect, as users are not able to assess the impact of federated learning itself, only the NeRF outputs. The added value of federated learning is that it enables more efficient computation of NeRFs. This is an essential requirement, since users need models of highest possible quality in shortest possible time.
- Interpreting the quality of discussions with end users on use cases, the potential of reducing the perceived complexity is clearly seen and rated to be "quite useful". The extend depends highly on command levels and roles.

Similarly, focusing on roles close to an incident site, using augmented reality under uncertainty on-site & remotely was discussed in a highly constructive manner, indicating that functionality is rated as "quite useful".

3.6 Perspectives

Perspectives result from both evaluation results and plans which were already determined in the requirements elicitation phase. With regard to such requirements and plans, from M19 on there will be an expansion of activities towards two domains (cf. [D2.1]):

- Extending relevant scenarios beyond floodings, the [DoA] subsumes wild fires as a second type of scenarios. In contrast to floodings, simulator use cases are much more prioritized by end users as a) the beginning of fire cannot be forecasted and b) there are much more factors causing dynamic evolutions in a situation. In the CREXDATA project, FMI has already presented an open-source, WISE wildfire spread model to be explored in Phase II.
- Experimentation with simulators was driven by demands from the interaction of ARGOS, prioritized AR use cases and reasonable test scenarios around NeRF applications. The integration of simulators and robotics using Gazebo is planned, leading also to the envisaged EmCase simulator HyperSuite concept.

In addition to these plans which were already documented, there are considerations interfacing demand side and reasonable test cases for WP4/WP5 technologies.

3.6.1 EmCase machine learning models

Especially with regard to tasks in WP5, further machine learning models are considered as relevant basis for T5.1 eXplainable AI, T5.2/T5.3 Visual Analytics and T5.5 Uncertainty Visualisation. The opportunities and priorities are derived from results of the first project period and especially feedback by stakeholders in the first evaluation phase.

3.6.1.1 Fatigue assessment based on biometric data

The aim of T5.1 is to use machine learning to enable a precise assessment of the state of fatigue of emergency services personnel. This is to be achieved by generating a result that provides information on whether or not the emergency services are exhausted during their activities.

Data generation is an essential step in the development of a successful machine learning model. The project partners have therefore decided to use an existing data set and expand it with additional data from emergency services. On the one hand, this ensures that the machine learning model has sufficient data to make robust predictions. It should be emphasised here that the existing data set primarily comprises data from nurses. This data will of course be adapted accordingly to meet the specific requirements of the project. The available data includes various parameters, including heart beats per minute, body temperature and electrodermal activity. In order to add further data to this dataset, the same parameters will be collected. A device like the Fitbit Sense 2 is to be attached to trainees at the fire service institute in exercises designed to simulate a realistic operation. During these exercises, the vital parameters are to be recorded continuously. The data collected will serve as the basis for training the machine learning model. By analysing this data, the model will learn to make objective assessments of the emergency services' state of fatigue. This should enable decision-makers to make informed decisions and optimise the safety and efficiency of the emergency services. For this purpose, a device like the Fitbit Sense 2 is to be procured as part of the project. This wearable device enables the continuous recording of the above mentioned parameters as well as other potentially relevant data.

To evaluate the functionality of the model, TC_009 is created to provide the decision-makers information on the vital condition of the first responders in the field. Expected results of TC_009 automatically generated information/score on mental status and fatigue based on physical biometrics, as well as an officer dashboard with unit overview and automated alerts when individuals currently on duty are approaching their fitness limits and need to be rested.

In order to communicate these results effectively, it is planned to visualise them in augmented reality so that decision-makers can view the data intuitively and quickly.

3.6.1.2 FMI machine learning models

In addition to generation of new scientific results with statistical modelling methods, the next step in Phase II is to explore the data with machine learning techniques to study weather dependencies seasonality and prepare for future impact-based forecasting tools to the health sector. The Facebook Prophet and partly Gradient Boosting machine learning model is utilized as a base to study the number of ambulance units needed in Helsinki in a weather emergency and to be used for developments of creation of impact-based forecasts for ambulance rescue management tasks. The next steps involve exploring if the FMI impact-based model could be run with impact data from Innsbruck and Dortmund and preliminary data discussions between DCNA and FDDO was held in the 4th CREXADATA meeting in BSC. We still have to investigate key data requirements during the next months and then continue the modelling tests.

Additionally, the Finnish showcase will utilize the Rapidminer Studio platform in data distribution in the upcoming months for selected impact-based forecast output data and thus utilize the same platform as the Dortmund and Innsbruck pilots will be using. In the ARGOS platform, open-source meteorological datasets for Helsinki are already available.

Furthermore, based on results from the field trials in Innsbruck and Dortmund, the possibility to uptake methods and technologies from WPs 3-5, especially Augmented Reality, will be explored in the second half of the project.

3.6.2 Parameter Prioritisation for Interactive Learning

As part of the field trials carried out with stakeholders in Innsbruck and Dortmund, the possible use of a simulation model in extreme weather events was also discussed. The focus here was on the MIKE+ simulation model for pluvial floods. The end users received a detailed introduction to the Urban Flood Simulator and rated it positively overall. The flood simulation model can be used in various situations in a critical scenario. They found the ability to simulate different flood scenarios and observe the impact on evacuation measures particularly useful. Furthermore, insights can be gained into where the placement of flood barriers would have the greatest effect. Firstly, given certain sensor data as input, one can determine the expected situation as output, a process known as reproduction. Secondly, the model can be used for prediction, allowing the assessment of scenarios where an assumption becomes invalid, such as the breaking of a dike or barrier. Thirdly, calibration involves comparing the sensed situation to the pre-simulated situation to ensure accuracy. Lastly, exploration entails identifying the measures that would lead to the desired situation. The simulation models used in the context of the EmCase are to be used in conjunction with Task T4.2 Interactive Learning for Simulation Exploration, among others. Firstly, certain use cases were defined for the pluvial flood, which in turn have correlations with certain parameters. The flood simulation model and the possible use cases were presented to the stakeholders in Innsbruck and Dortmund. During the field trials, the experts prioritised the use cases in order to be able to further develop the most relevant use cases from the user's perspective in an extreme weather event. The prioritisation showed that use cases 6.1 and 6.2 are the most relevant from a user perspective, as the safety of the emergency forces is the top priority for the rescue operation. Use cases 1 and 2 were also categorised as very relevant by the users. Another request for an as yet undefined use case was the capacity of a creek flow, which can be blocked by debris in heavy rainfall events and lead to unpredictable flooding.

4 Life Science Use Case

This Use Case has two clearly defined scenarios: the epidemiological and the multi-scale infection one. Thus, all Sectionsections of this Chapter will be divided into these two scenarios.

4.1 Scenario description

The scenario was drafted in Deliverable 2.1 in M6 and we here described the final status of how the two different scenarios of this iusse will be used in this project.

4.1.1 Epidemiology scenario

In this pilot scenario, our objective is to simulate the spatiotemporal dynamics of epidemic processes using a mathematical model for disease progression, coupled with anonymized phone-based mobility data. We focus on the COVID-19 pandemic in Spain. The goals encompass calibrating epidemiological parameters, understanding the epidemic's spread, and simulating various counterfactual scenarios for controlling the epidemic process. These scenarios involve evaluating different interventions to reduce the epidemic's impact on total fatalities and the health system (i.e., flattening the curve). Strategies for controlling the epidemic include designing confinement and mobility reduction policies (ranging from national-level lockdowns to region-specific restrictions) and finding effective vaccination strategies under a limited number of vaccine rollouts.

In the first stage, we have collected COVID-19 reports that include new cases, fatalities and hospitalizations reported daily and weekly at different levels of spatial resolution (e.g. country-level, provinces, municipalities) from different sources [4][6]. Furthermore, we have also gathered phone-based anonymized daily mobility data from the Ministry of Transportation[7] and reported values for the different epidemiological parameters. We used these comprehensive datasets to define a **reference model** for simulating the spread of COVID-19 in Spain using population mobility data that will be used by the demonstrators.

In the second stage, we have extended the MMCACovid19 simulator with new features. First, we have added three functionalities: i) modelling the effect of introducing vaccination; ii) considering herd immunity, and iii) a parameter that accounts for the reinfection rate. These extensions are critical for achieving the scenarios defined for this project. Additionally, we have defined a new and simple file in JSON format to store all the model's parameters.

In the third stage, we developed an extreme-scale model exploration workflow based on the EMEWS[8] framework. Our workflow facilitates the execution of model exploration experiments on high-performance computing infrastructures, scaling dynamic computational experiments to millions of model instances. It also integrates various metaheuristic optimization algorithms, including Genetic Algorithms (GA) and Covariance Matrix Adaptation (CMA-ES). These optimization algorithms are used to calibrate model parameters by fitting simulations to real data and conducting in-silico experiments to design effective intervention strategies. The reference dataset, extended simulator, and workflow

---

[6] https://cnecovid.isciii.es/covid19/

[7] https://www.transportes.gob.es/ministerio/proyectos-singulares/estudio-de-movilidad-con-big-data

8 https://emews.github.io/

are the primary components necessary to achieve the objectives of the epidemiological scenario. We have focused on two key problems:

- Calibration of unknown parameters (fitting) and estimating uncertainty.
- Design, exploration and analysis of effective interventions to control the epidemic.

Integrating CREXDATA's AI technologies significantly enhances the effectiveness and usability of the simulator and the model exploration workflow using Altair's RapidMiner Studio. Specifically, these enhancements include interactive learning to explore simulations, efficient analysis of large datasets, and visual analytics for decision-making under uncertainty. In sum, this pilot project highlights the potential of **using epidemic simulations and extreme-scale data analytics to comprehend the intricate spatiotemporal patterns inherent in epidemic processes**. It sheds light on the most effective strategies for mitigating such processes, whether through population mobility restrictions, reduced social contact, or optimized vaccination programs to manage the severity of COVID-19 in patients. The insights gained from this endeavour can inform and refine the design of containment measures should a future pandemic arise.

4.1.2 Multi-scale infection scenario

This pilot scenario aims to simulate the dynamics of COVID-19 infection in the lungs and the impact of airflow in the conductive and respiratory zone of the lungs using two computational tools, Alya and PhysiBoSS (see Section 4.3.2). The goal is to understand the different severity of COVID-19 in patients and optimise both non-drug interventions and drug treatments using CREXDATA technologies.

The project begins with the collection of anatomical data and the creation of a 3D model of the conductive and respiratory zone of the lungs. Alya, a high-performance computational mechanistic simulations framework, is then used to simulate the airflow and the distribution of SARS-CoV-2 virus particles in the upper airways. This step is crucial as it provides insights into how the virus may spread within the airways. Next, PhysiBoSS, a hybrid agent-based modelling framework, is used to simulate the infection dynamics of COVID-19 in the lung epithelia. This step is key to understanding how the virus interacts with the cells and triggers an immune response.

To explain the different severity of COVID-19 in patients, omics data (genomics, transcriptomics, proteomics, and metabolomics) is integrated into the PhysiBoSS simulation. This allows the model to take into account the individual genetic and molecular profiles of patients, which can significantly influence the course of the disease. The simulations are initially performed on a simplified model with a few bronchioles, and then scaled up to a full lung model. This approach ensures that the model is robust and can accurately represent the complex dynamics of airflow and infection in the lungs.

The results of the simulations are then used to optimise non-drug interventions, such as ventilation and air filtration systems, and drug treatments by using the different technologies developed at CREXDATA. For example, using Altair's RapidMiner Studio, complex event forecasting can determine when the best time is to administer mechanical ventilation to a patient, or the federated machine learning can give insights into having optimised patient-specific drug treatments.

The integration of CREXDATA's AI technologies further enhances the effectiveness and usability of the multi-scale modelling approach. These technologies provide additional capabilities such as model interpretability, prediction of complex events, interactive exploration of simulations, efficient analysis of large amounts of data, and decision-making under uncertainty.

In conclusion, this pilot project demonstrates the potential of using multi-scale modelling and extreme-scale data analytics to understand the different severity of COVID-19 in patients and optimise interventions. The insights gained from this project could potentially inform the development of new treatment strategies or preventive measures, ultimately contributing to the fight against COVID-19.

Additionally, here is a summary on how this scenario aligns with CREXDATA objectives:

Objective O1: Extreme-scale Data Ingestion/Generation, Fusion and Exploitation:

- **MO1.1:** The use case involves the ingestion and fusion of multimodal data. For the simulations, we are using anatomical data from medical imaging and omics data from patients. For the analysis of the simulation results by the technologies developed in CREXDATA, we are using online data from Alya and PhysiBoSS as well as the information from the medical images and omics data.
- **MO1.2:** The use case employs dynamic simulation models, Alya and PhysiBoSS, to simulate the airflow in the conductive and respiratory zone of the lungs and the infection dynamics of COVID-19 in the lung epithelia.
- **MO1.3:** The use case doesn't directly involve handling multilingual social data in real-time.

Objective O2: Real-time Predictive Knowledge and Forecasts:

- **MO2.1:** The use of Federated Machine Learning in the use case allows for real-time learning from decentralized patient-specific omics data, contributing to predictive knowledge about the different severity of COVID-19 in patients.
- **MO2.2:** The use of Complex Event Forecasting in the use case allows for multiresolution forecasting of the progression of COVID-19 infection in the lung epithelia under uncertainty.
- **MO2.3:** The use case involves the optimization of non-drug interventions and drug treatments based on the results of the simulations, which can be seen as a form of Prediction-as-a-Service (PaaS).

Objective O3: Reduced Perceived Complexity:

- **MO3.1:** The use of Graphical Workflow Specification in the use case allows for the visual management of the complex workflow involved in the simulations, reducing the perceived complexity.
- **MO3.2:** The use of Visual Analytics supporting XAI in the use case allows for the visualization and explanation of the results of the simulations and the predictions made by the models, aiding in understanding complexity and reasoning under uncertainty.
- **MO3.3:** While the use case doesn't directly involve augmented reality, the insights gained from the simulations could potentially be visualized using augmented reality tools, allowing for on-site and remote understanding of the results under uncertainty.

4.2 Pilot definition and demonstrator

In the following section and for the two scenarios of this use case, we describe the definition of the pilot and an explanation on how we are going to demostrate its use in a real-world scenario to assess its practical application.

4.2.1 Epidemiology scenario

The epidemiology scenario will use a pilot defined in the following section and will be proven in the following demonstration.

4.2.1.1 Pilot definition:

The pilot is the simulation and analysis of the COVID-19 pandemic in Spain. The goals of this pilot are to develop tools that help understand the complex interplay between population mobility, social contacts, vaccination and the spatio-temporal patterns of the epidemic process; and to use the tools to guide the design of effective interventions to control an epidemic using a combination of simulations, AI and visual analytics technologies.

To achieve these goals, we will use a model exploration workflow that will allow the calibration of epidemiological parameters. The parameter calibration will be accomplished by fitting the simulations to a multidimensional dataset that includes a time series of cases, hospitalizations, and deaths for different regions and age groups. Various strategies, including optimization algorithms (e.g., GA, CMA-ES) and interactive learning, guide the parameter calibration process.

This will enable us to achieve the desired KPI for **calibrating epidemiological parameters to fit incidence time series**. This, in turn, will allow us to identify critical epidemiological parameters such as infection, hospitalization, death, and recovery rates for different age strata. Furthermore, the combination of optimization algorithms with active learning will help us achieve the KPI of **characterizing the parameter space with 50% fewer simulations**, thereby reducing the time and computational resources required.

The calibrated model will be used for the model-based design and analysis of effective interventions, considering: i) the exploration of mobility reduction and social distancing policies that consider health and socioeconomic impacts, and ii) the exploration of vaccination strategies for COVID-19. We will use the model exploration workflow, combining optimization, interactive learning, and visual analytics to achieve the third KPI of **forecasting seven parameter sets that reduce COVID-19 infections.**

This pilot has several elements that are needed:

MMCACovi19-vac simulator: this component of the workflow is used to simulate the evolution of the epidemic in the reference model under different scenarios

Model Exploration Workflow: this component allows the exploration of large parameter spaces by evaluating a large number of simulations in parallel and by adaptively learning properties of the parameter space.

Interactive Learning for Simulation Exploration (T4.2) will allow users to interactively explore and learn from simulations and the effects of different policies. In this context, we are working in implementing a reinforcement learning approach to identify effective intervention strategies in different contexts **(T4.2).** We will consider the effect that the uncertainty of the parameters has on the different strategies and policies. For this, we will evaluate ensembles of simulations with parameters sampled from learned distributions. We will use tools from **Visual Analytics for Decision Making under Uncertainty** to analyze the results of the enable trajectories and support decision-making under uncertainty. All these elements are being included in **Graphical Workflow Specifications** to visually manage the complex workflow involved in this step using RapidMiner. Furthermore, some additional CREXDATA tools will be used if we consider them useful to have meaningful results:

- **Federated Machine Learning** could be used for parameter calibration when data exchange between different stakeholders is restricted due to privacy concerns.
- **Distributed "Analytics as a Service"** could provide analytics capabilities as a service, optimized for distributed computing environments, allowing for efficient analysis of large amounts of data generated by the simulations.

4.2.1.2 Demonstrator:

Step 1: Data Collection and Preprocessing

- Collect and preprocess COVID-19 reports for Spain at different levels of spatial resolution. The reports include the time series of daily reports of cases, hospitalisation and fatalities grouped by age and for the different regions.
- Phone-based, anonymized population mobility data. The mobility indicator includes daily reported origin-destination matrices that account for the daily trips between the different mobility areas as well as the percentage of moving population in this area.
- Geographic layer for the territorial units where the time series of the different indicators are geo-referenced (e.g. mobility areas, municipalities, provinces).
- Bibliographic review of reported epidemiological parameters (e.g. infective period, age-dependent mortality rate, protective effect of vaccines).
- Preprocess the data to filter inconsistencies, apply smoothing when needed, input missing data, and store the data in the NetCDF format enabling the fusion of the different data sources. This step ensures that the data can be directly compared to simulations and helps to reduce computational complexity.
- This step has been achieved.

Step 2: Definition of a reference model

- We first defined a standard configuration format based on JSON to formulate an instance of the metapopulation model. The
- In this step, we process and integrate population demographic data with mobility to create a metapopulation for Spain. This includes population by age for the different regions of the metapopulation, the surface of each area, the average number of contacts between age groups, etc.
- We also collect the parameter values from the original model reported by Arenas and collaborators [5].

- The main difference with the original model is the metapopulation structure that is now defined over mobility areas, enabling the direct integration of the mobility data.
- This step has been achieved.

Step 3: Simulation of the first and second waves

- Adjust the initial conditions for the pandemic by identifying the first reported cases and mapping them to the corresponding regions of the metapopulation model.
- Run the MMCACovid19-vac simulation to model the spread of COVID-19 in Spain during 2020.
- Post-process the simulation results to obtain the set of observables reported at the same level of spatial resolution as the time series of COVID-19 reports.
- Analyse and visualise the results by comparing the simulated data to the real-time series using various visual analytic techniques and error metrics. This step has been completed.
- This step has been achieved.

Step 4: Evaluate vaccination strategies

- Define different vaccination strategies.
- Run the MMCACovid1-vac simulation to model the epidemic dynamics under the different vaccination campaigns.
- Post-process the simulation results to analyse and visualise the effect that the different vaccination strategies have on critical epidemic indicators such as the number of fatalities and the peaks of the waves of cases and hospitalizations.
- This step has been achieved.

Step 5: Extending the MMCACovid19-vac simulator

- **Temporal Mobility Network**: Currently, the simulator accommodates a static mobility network. We are developing an approach to incorporate time-dependent mobility networks, enabling us to account for seasonal changes in the mobility structure and to implement mobility reduction policies that affect this structure. This is ongoing work.
- This step is currently under development.

Step 7: Application of AI and Machine Learning Technologies

- Use **Optimization-via-Simulatio**n for calibrating epidemiological parameters.
  - This has already been done
- Use **Reinforcement Learning for Simulation Exploration** to find effective interventions (confinements and vaccinations to control the epidemic.
  - This is currently a work in progress.
- Use **Active Learning for Simulation Exploration** to fully characterize the complete parameter space running a reduced number of simulations
  - This is currently a work in progress.
- **Visual Analytics for Decision Making under Uncertainty** to analyze the results of the enable trajectories and support decision-making under uncertainty.

    - There has been some work on using visual Analytics techniques to detect causal associations between population mobility between regions and the dynamics of new infections at the target region.
- Other technologies developed in CREXDATA will be used if considered necessary:
    - Use **Graphical Workflow Specifications** to visually manage the complex workflow involved in this step using RapidMiner.
    - **Federated Machine Learning** could be used for parameter calibration when data exchange between different stakeholders is restricted due to privacy concerns.

Step 8: Visualization

- We are using **Visual Analytics for decision-making under Uncertainty** to analyze the uncertainty in the simulation results and the model predictions and to support decision-making under this uncertainty. Specifically, we are interested in the early detection of outbreaks and hotspots of spreading
- We will evaluate ensembles of simulations with parameters sampled from learned distributions. We will use visual analytics to analyze and enable trajectories of simulations and support decision-making under uncertainty.

Step 9: Validation

- Validate the results of the simulation against real data to ensure that the simulations accurately represent the real-world dynamics of the COVID-19 pandemic. This step is currently in progress.
- We will also consider applying the entire framework in different contexts, such as other countries for which COVID-19 reports and mobility data are available in the public domain. This will provide another method of validating the framework.

Each of these steps plays a critical role in the overall process and contributes to the success of the demonstrator. They ensure that the simulations are based on accurate data, are computationally efficient, and provide meaningful and actionable insights. The integration of AI and machine learning technologies further enhances the effectiveness and usability of the multi-scale modelling approach.

4.2.2 Multi-scale infection scenario

The multi-scale infection scenario will use a pilot defined in the following section and will be proven in the following demonstration.

4.2.2.1 Pilot definition :

The pilot is the simulation of upper airways flow and COVID-19 infection of lung epithelia with advanced ai technologies. The goal of this pilot is to understand the dynamics of COVID-19 infection in the lung epithelia and the impact of airflow in the upper airways using a combination of computational tools and AI technologies. This pilot has several elements that are needed:

Airflow Simulation with Alya: The first part of the simulation involves using Alya to model the airflow in the upper airways. This would involve creating a 3D model of the conductive and respiratory zone of the lungs. The model considers the physical and mechanical properties of the airway tissues, as well as the airflow dynamics such as velocity, pressure, and turbulence. The output of this simulation is the airflow patterns and the distribution of airborne particles (like the SARS-CoV-2 virus).

COVID-19 Infection Simulation with PhysiBoSS: The second part of the simulation involves using PhysiBoSS to model the infection dynamics of COVID-19 in the lung epithelia. This involves creating an agent-based model of the lung epithelial cells, and simulating the intracellular signalling pathways that are activated upon SARS-CoV-2 infection. The model considers the biological and physiological properties of the cells, as well as the dynamics of the viral infection and the immune response. The output of this simulation is the progression of the infection in the lung epithelia, and the response of the immune system.

Integration and Analysis: These two simulations need to be integrated and the combined effect of airflow and infection dynamics on the progression of COVID-19 analysed. This provides valuable insights into how the disease spreads and progresses in the lungs and can inform the development of new treatment strategies or preventive measures. **Complex Event Forecasting** is being used here to predict the occurrence of complex events based on historical data, while **Interactive Learning for Simulation Exploration** allows users to interactively explore and learn from simulations.

All of these elements are being included in **Graphical Workflow Specifications** using Altair's Rapidminer Studio to visually manage the complex workflow involved in this step.

Moreover, some additional CREXDATA tools will be used if we consider them useful to have meaningful results:

- **Federated Machine Learning** could be used to train models on decentralised data, helping to understand the different severity of COVID-19 in patients.
- **Optimised Distributed "Analytics as a Service"** could provide analytics capabilities, optimised for distributed computing environments, allowing for efficient analysis of large amounts of data generated by the simulations.
- Throughout the process, **Explainable AI (XAI)** and **Visual Analytics supporting XAI** could be used to explain the results of the simulations and the predictions made by the models, increasing trust in the results and facilitating decision-making.
- **Visual Analytics for Decision Making under Uncertainty** provides visual analytics tools for decision-making under uncertainty, supporting decision-making under the uncertainty in the simulation results and the model predictions.

4.2.2.2 Demonstrator :

With the pilot defined, we have linked the different technologies and infrastructure parts to complete the use case successfully. We have divided the demonstrator in different parts that have different expected timelines (Figure 22).

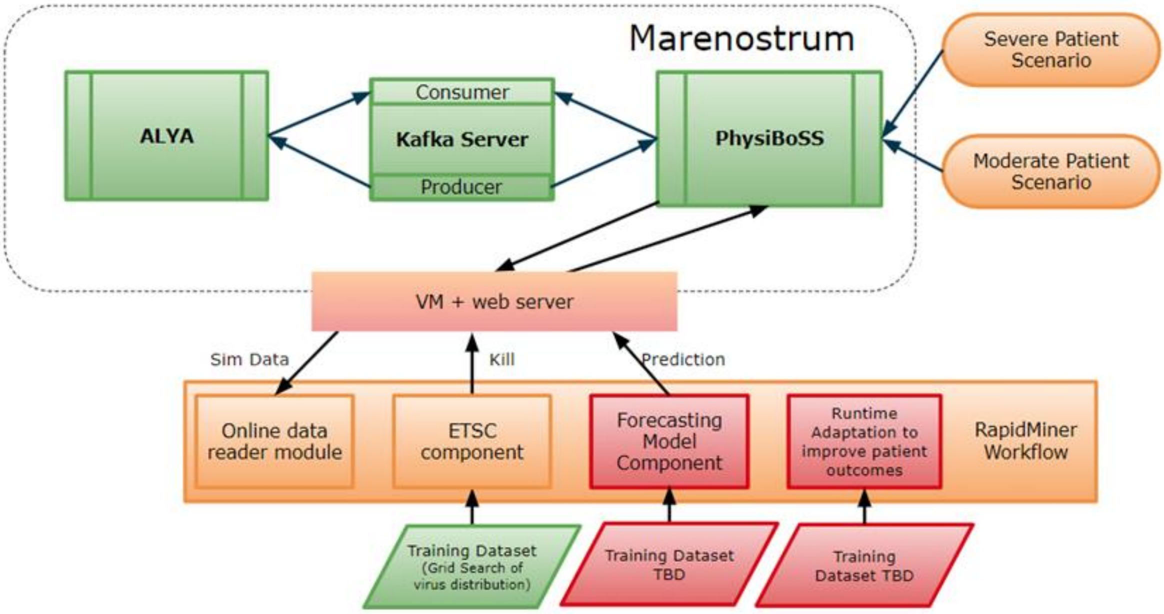


Figure 22: Flowchart of the pilot workflow. In green, parts that are already available. In orange, parts that are being tested and that will be ready for the mid-term review in September 26. In red, parts that will be incorporated in the second term of the project. VM stands for virtual machine; ETSC stands for Early Time Series Classification.

Step 1: Data Collection and Preprocessing

- Collect anatomical data of the upper airways. This was done through medical imaging techniques like CT or MRI scans to provide the necessary data to create a realistic and accurate model of the upper airways.
- Preprocess the data to create a 3D model of the upper airways to be used in the Alya simulation. This step ensures that the data is in a suitable format for the simulation and helps to reduce computational complexity.
- This step has been achieved.

Step 2: Integration of Omics Data in Boolean models

- Integrate patient-specific omics data (genomics, transcriptomics, proteomics, and metabolomics) into the PhysiBoSS simulation, enabling us to understand the different severity of COVID-19 in patients.
- This step has been achieved.

Step 3: Airflow Simulation with Alya

- Define the boundary conditions, such as the airflow rate and the particle concentration representing the SARS-CoV-2 virus to set the initial conditions for the simulation based on real-world parameters.
- Run the Alya simulation to model the airflow and particle distribution to understand the virus spread within the airways.

- Post-process the simulation results to obtain the airflow patterns and particle distribution to analyse and visualise the results of the simulation in a meaningful way.
- This step has been achieved.

Step 4: Infection Simulation with PhysiBoSS

- Define the initial conditions, such as the number of epithelial cells and the initial viral load to set the stage for the infection simulation, based on realistic biological conditions.
- Run the PhysiBoSS simulation to model the infection dynamics in the lung epithelia to understand the virus interactions with the cells and its triggering of the immune response.
- Post-process the simulation results to obtain the progression of the infection and the immune response to analyse and visualise the results of the infection simulation.
- This step has been achieved.

Step 5: Scaling Up Simulations

- **Version 1 - Few Bronchioles:** Start with a simplified model that includes only a few bronchioles to allow for quicker simulations and enable the team to test and refine the model and simulation parameters.
    - This is currently work in progress.
- **Version 2 - Full Lung:** Once the team is confident in the accuracy and reliability of the simulations, scale up to a full lung model to provide a more comprehensive understanding of the disease progression in the whole lung.
    - This is in scope.

Step 6: Integration and Analysis

- Integrate the results from the Alya and PhysiBoSS simulations to understand the combined effect of airflow and infection dynamics.
- Analyse the combined effect of airflow and infection dynamics on the progression of COVID-19 to provide insights into how these two factors interact and influence the disease progression.
- This step is currently work in progress.

Step 7: Application of AI and Machine Learning Technologies

- Use **Complex Event Forecasting** to predict the progression of COVID-19 infection in the lung epithelia based on the initial conditions and the results of previous simulations.
    - This is currently work in progress.
- Use **Interactive Learning for Simulation Exploration** to understand the dynamics of airflow and infection in the lungs, and to identify patterns and trends that could inform the development of interventions and treatments.
    - This is currently work in progress.

- Use **Graphical Workflow Specifications** to visually manage the complex workflow involved in this step using RapidMiner.
    - This is currently work in progress.
- Other technologies developed in CREXDATA will be used if considered necessary:
    - Use **Federated Machine Learning** to train models on patient-specific omics data to help in understanding the different severity of COVID-19 in patients.
    - Use **Explainable AI (XAI)** and **Visual Analytics supporting XAI** to explain the results of the simulations and the predictions made by the models to increase the trust in the results and facilitate decision-making.
    - Use **Optimised Distributed "Analytics as a Service"** to analyse the large amounts of data generated by the Alya and PhysiBoSS simulations, and to derive insights from this data in an efficient and scalable manner.

Step 8: Visualisation

- Visualise the results using appropriate visualisation tools to communicate the results of the simulations in a clear and understandable way.
- Potentially, we will use **Visual Analytics for Decision Making under Uncertainty** to visualise the uncertainty in the simulation results and the model predictions, and to support decision-making under this uncertainty.
- This step is currently in scope.

Step 9: Validation

- Validate the results of the simulation against experimental or clinical data to ensure that the simulations accurately represent the real-world dynamics of COVID-19 infection.
- This step is currently in scope.

Each of these steps plays a vital role in the overall process and contributes to the success of the demonstrator. They ensure that the simulations are based on accurate data, are computationally efficient, and provide meaningful and actionable insights. The integration of AI and machine learning technologies further enhances the effectiveness and usability of the multi-scale modelling approach.

4.3 Simulation tools

In this section we describe the different tools used in each one of the scenarios.

4.3.1 Epidemiology scenario

There are different approaches to model and simulate the dynamics of epidemic processes. When we simulate an epidemic, we can distinguish between the compartmental model and the population dynamics model. The former describes the different states in which the individual agents of a population can be found together with the transition rate between those states. Compartmental models can range from the classical Susceptible → Infected → Susceptible (SIS) and Susceptible → Infected → Recovered (SIR) to more complex models including, for instance, a more fine-grained description of the infectious process [6], compartments to described individuals going into isolation or quarantine, vaccination. On the other hand, population dynamics usually describe how the different agents interact with each other by moving contact. Models can be represented or encoded using mathematical equations or a set of predefined rules that govern the behaviour of the agents. Equation-based approaches can be grouped in deterministic models that are based on differential equations [5], and those based on stochastic processes [7]. Equation-based models can only be analytically solved for simple cases, whereas numerical simulations are the only means to analyse a model. In the case of rule-based or agent-based models results can only be obtained by numerical simulations.

In 2020, Arenas and collaborators implemented a model to simulate the COVID-19 pandemic using the Microscopic Markov Chain Approach (MMCA). The simulator is based on a library named MMCACovid19 written in the Julia programming language and allows the simulation of the spread of an infectious disease over a metapopulation connected by the recurrent mobility network. The simulator also enables the definition of different types of agents (e.g. kids, adults), their contact matrix, and their agent-dependent epidemic and parameters (rate of infection, probability of bad prognosis, etc) [5].

As depicted in Figure 23, the infectious model includes the following compartments: susceptible (S), exposed (E), asymptomatic infectious (A), symptomatic infectious (I), pre-hospitalized in ICU (PH), predeceased (PD), admitted in ICU that will recover (HR) or decease (HD), recovered (R), and deceased (D). The population is divided into three age groups: young people (Y), adults (M) and elderly people (O). On the bottom panel, the figure schematically represents the Movement-Interaction-Return model used to simulate the interaction between different regions. Finally, in the bottom right corner, the figure shows the new features added to the MMCACovid19 simulator. We have added three new modelling functionalities: i) modelling the effect of introducing vaccination; ii) considering herd immunity, and iii) a parameter that accounts for the reinfection rate. Vaccination affects the parameter γ^g (ICU probability), and the Fatality probability in ICU ω^g. These extensions are critical for achieving the scenarios defined for this project.

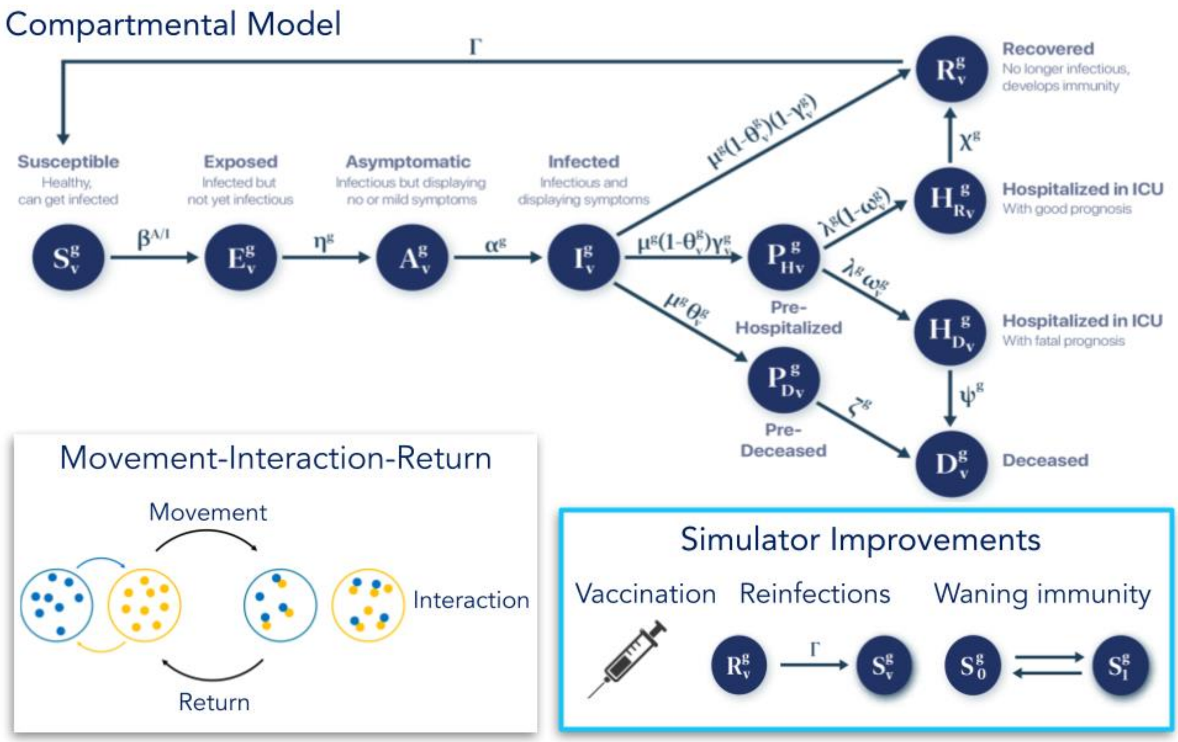


Figure 23: The extended MMCACovid19 simulator. The top panel illustrates the compartmental model, where each circle represents a different stage of the disease, and the arrows indicate the transition probability rates between states. The superscripts and subscripts g and v denote that a rate depends on the age strata and the vaccination status, respectively.

Moreover, to simplify the task of creating different instances of the model we have defined a new and simple file in JSON format to store all the model's parameters. The configuration file stores the epidemic and population demographic parameters, the vaccination strategy, the mobility reduction policies and pointers to data files. This feature allows easily creating different instances of the model by setting particular parameter values in the configuration. An example of the configuration file is shown in Figure 24. Finally, we have also defined a simple interface for instantiating and running simulations. This script works a simple interface that is invoked by the workflow to evaluate different instances of the model.

```
"simulation": {
    "first_day_simulation": "2020-02-09",
    "last_day_simulation": "2020-04-14",

    "output_format": "netcdf"
},
"data": {
    "initial_condition_filename": "initial_conditions.nc",
    "metapopulation_data_filename": "metapopulation_data.csv",
    "mobility_matrix_filename": "R_mobility_matrix.csv",
    "kappa0_filename": "kappa0_by_time.csv"
},
"epidemic_params": {
    "scale_β": 0.51,
    "βᴬ": 0.046053,
    "βᴵ": 0.0903,
    "ηᵍ": [0.2747252747252747, 0.2747252747252747, 0.2747252747252747],
    "αᵍ": [0.26595744680851063, 0.641025641025641, 0.641025641025641],
    "μᵍ": [1.0, 0.3125, 0.3125],
    "θᵍ": [0.00, 0.008, 0.047],
    "γᵍ": [0.0003, 0.003, 0.026],
    "ζᵍ": [0.12820512820512822, 0.12820512820512822, 0.12820512820512822],
    "λᵍ": [1.0, 1.0, 1.0],
    "ωᵍ": [0.30, 0.30, 0.30],
    "ψᵍ": [0.14285714285714285, 0.14285714285714285, 0.14285714285714285],
    "χᵍ": [0.047619047619047616, 0.047619047619047616, 0.047619047619047616],
    "Λ": 0.02,
    "Γ": 0.01,
    "rᵥ": [0.0, 0.6],
    "kᵥ": [0.0, 0.4],
    "risk_reduction_dd": 0.0,
    "risk_reduction_h": 0.1,
    "risk_reduction_d": 0.05
},
"population_params": {
    "age_labels": ["Y", "M", "O"],
    "C": [
            [ 0.598,  0.38486, 0.01714],
            [ 0.244,  0.721,   0.0353],
            [ 0.1919, 0.5705,  0.2376]
        ],
    "kᵍ": [11.8, 13.3, 6.76],
    "kᵍ_h": [3.15, 3.17, 3.28],
    "kᵍ_w": [1.72, 5.18, 0.0],
    "pᵍ": [0.0, 1.0, 0.0],
    "seeds_age_ratio": [0.12, 0.16, 0.72],
    "ξ": 0.01,
    "θ": 2.5
},
"vaccination":{
    "ϵᵍ": [0.1 , 0.4 , 0.5],
    "percentage_of_vacc_per_day": 0.0,
    "start_vacc": 0,
    "dur_vacc": 0
},
"NPI":{
    "κ₀s": [0.8],
    "ϕs":  [0.2],
    "δs":  [0.8],
    "tᶜs": [100]
}
```


Figure 24*:* Format for the MMCACovid19 simulator configuration file

Finally, we present a schematic representation of the model exploration workflow. Figure 25 shows the different components of the workflow as well as the different steps that are run to evaluate a single simulation instance. This representation is agnostic regarding the model exploration algorithm used which in this case is encapsulated in the arrow highlighted as the "Learning Step". The workflow has been implemented using the EMEWS framework based on previous work [8]. The model exploration workflow is fully customizable by changing the way results are aggregated (Step 3), the cost function used (Step 4) and the model exploration algorithm (Step 5). We have already tested it for calibrating epidemiological parameters using CMA-ES, and now we will extend the workflow to enable performing Active Learning to fully characterize the parameter space.

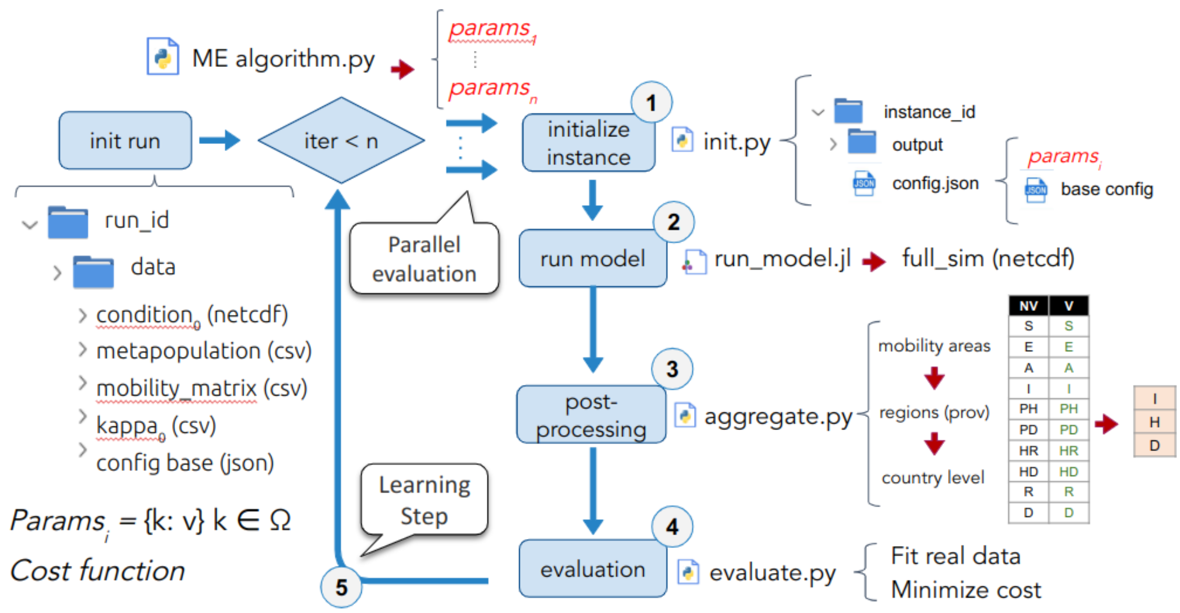


Figure 25: Model exploration workflow

4.3.2 Multi-scale infection scenario

Alya [9] is a high-performance computational mechanics code used to solve complex coupled multi-physics, multi-scale, and multi-domain problems, mostly coming from the engineering realm. It can solve a variety of physics problems, including incompressible/compressible flows, non-linear solid mechanics, chemistry, particle transport, heat transfer, turbulence modelling, electrical propagation, and more. Alya was designed for massively parallel supercomputers, and it uses high-performance computing techniques for distributed and shared memory accelerated supercomputers. Alya is proprietary software and can be accessed through its website: https://alya.gitlab.bsc.es/alya/alya/-/wikis/home.

PhysiBoSS [10] is an open-source sustainable integration of stochastic Boolean and agent-based modelling frameworks. It's a hybrid agent-based modelling framework that allows simulating signalling and regulatory networks within individual cell agents. PhysiBoSS 2.0 expands the PhysiCell functionalities by enabling the simulation of intracellular cell signalling. It's suitable for modelling diseases and studying the interplay between the microenvironment, the signalling pathways that control cellular processes, and population dynamics. PhysiBoSS repository can be found at https://github.com/PhysiBoSS/PhysiBoSS.

Both tools are part of BSC's commitment to advancing the field of computational science and engineering. They are used by researchers worldwide to tackle complex scientific problems.

EMEWS (Extreme-scale Model Exploration with Swift) [11] is a powerful open-source framework designed for the exploration and optimization of parameters in multi-scale models. It allows scientists to conduct large-scale model explorations and optimizations on high-performance computing resources. In the context of the demonstrator, EMEWS could be used to explore different parameters of the Alya and PhysiBoSS simulations, such as the initial conditions, boundary conditions, and model parameters. By running multiple simulations with different parameter settings, EMEWS together with CREXDATA technologies can help identify the parameters that lead to the best simulation results, thus optimising the multi-scale model. This could provide valuable insights into the dynamics of airflow and COVID-19 infection in the lungs, potentially informing the development of new treatment strategies or preventive measures. EMEWS repository can be found at https://github.com/bsc-life/PhysiCell-EMEWS.

4.4 Initial Use Case Evaluation

In this section we describe the evaluation that we have done of the use case in the first half of the project.

4.4.1 Epidemiology scenario

The epidemiological scenario begins with the definition of the reference model and the data to be used for calibration and exploration of effective interventions. The original model was based on a metapopulation including 7,156 areas corresponding to municipalities. This decision was primarily driven by the available mobility data at the onset of the pandemic. However, during and after the COVID-19 pandemic, the Ministry of Transportation started releasing a much more comprehensive population mobility dataset. This new dataset, referred to as MITMA zonification, is composed of 2,850 mobility areas corresponding to districts or groups of districts.

To use the MITMA mobility dataset, we need to reformulate the metapopulation model. Consequently, the reference model for this project is defined within a metapopulation composed of 2,850 MITMA mobility areas. For population demographics, we used reports from the Instituto Nacional de Estadística. The mobility network was constructed based on phone-based origin-destination matrices reported by the Ministry of Transportation, as explained in the relevant Section. We then processed the COVID-19 reports and stored them in a multi-dimensional data array to facilitate easy comparison with the aggregated simulation output. The reference model includes the same parameters as the original model, and we used the same values for all parameters. Finally, we translated the initial conditions from the original zonifications (municipalities) into the MITMA zonification. These initial conditions correspond to the first cases reported in Spain on February 9, 2020.

Next, we focused on developing the model exploration workflow (step 6) for parameter calibration and identifying effective interventions. The workflow outlines all the steps required to create an instance of the model, run the simulation, evaluate the results, and update the internal state of the model exploration algorithm (except in the case of the parameter sweep).

The workflow serves as a roadmap for the project, guiding the implementation of each step and ensuring that all necessary tasks are accounted for. The implementation of the workflow for parallel model exploration has been completed. The next step will require connecting the Model Exploration Workflow to the interactive Learning and Visual Analytic components. To connect the Model Exploration workflow to RapidMiner we will develop a REST-API so MMCACovid19 simulations can be executed as web services. The successful implementation of the workflow is key to the smooth operation of the simulations and the generation of accurate results and will be the focus of the work in the coming months.

Calibration of the epidemiological model

To test the parameter calibration workflow, we ran different experiments using the CMA-ES to evaluate the convergence and performance of the algorithm. The aim is to find a set of epidemiological parameters that produce a simulation output that minimizes an error metric to the time series of real data. Although the workflow allows the use of different error metrics and observables for fitting, here we present preliminary results using Root Mean Square Error (RMSE) as the error metric and the time series of fatalities reported at the country level as done in the original publications. For the CMA-ES we used the implementation included in the DEAP Python library. We set the population size to 191 individuals so they can be efficiently allocated into two computer nodes which contain 96 CPUs and let one free CPU for the workflow script. Additionally, we set a total of 25 generations for the algorithm to ensure convergence. For the remaining CMA-ES parameters, we used the default values.

We first focused on recalibrating the parameters from the original publication. The set includes six different parameters, four epidemiological and two related to social distancing and confinement. Table 8 shows the parameters, including the originally reported values and the new values found in the calibration procedure. Some parameters remain close to the originally reported values, while others show significant differences. These discrepancies may be due to various factors, including differences in the zonification and mobility data used, as well as the more comprehensive dataset of COVID-19 reports guiding the parameter search.

Table 8*: Set of parameters selected for testing the calibration workflow*

Parameter	Original Value	New Value	Units
β^I	0.075	0.06	unitless
β^A	0.0375	0.0451	unitless
η^g	2.444	1.474	1/days
α^g	5.671; 2.756; 2.756	4.74; 2.74; 2.74	1/days
δ	0.207	0.001174	unitless
ϕ	0.174	0.153146	unitless

We have observed that the optimal parameters demonstrate strong agreement in predicting daily hospitalizations and fatalities. However, there exists a significant disparity in the results concerning the daily case numbers. Our findings indicate that the reported cases in the first wave were 2 orders of magnitude lower than those predicted by the simulations. We assessed the convergence of the CMA-ES algorithm and determined that it typically converges to a minimum RMSE within approximately 15 generations. Figure 26 illustrates a rapid decline in the cost function value during the initial ten iterations, which stabilizes therefore, suggesting convergence. Consequently, calibrating the six selected parameters necessitates evaluating around 4,000 different simulations.

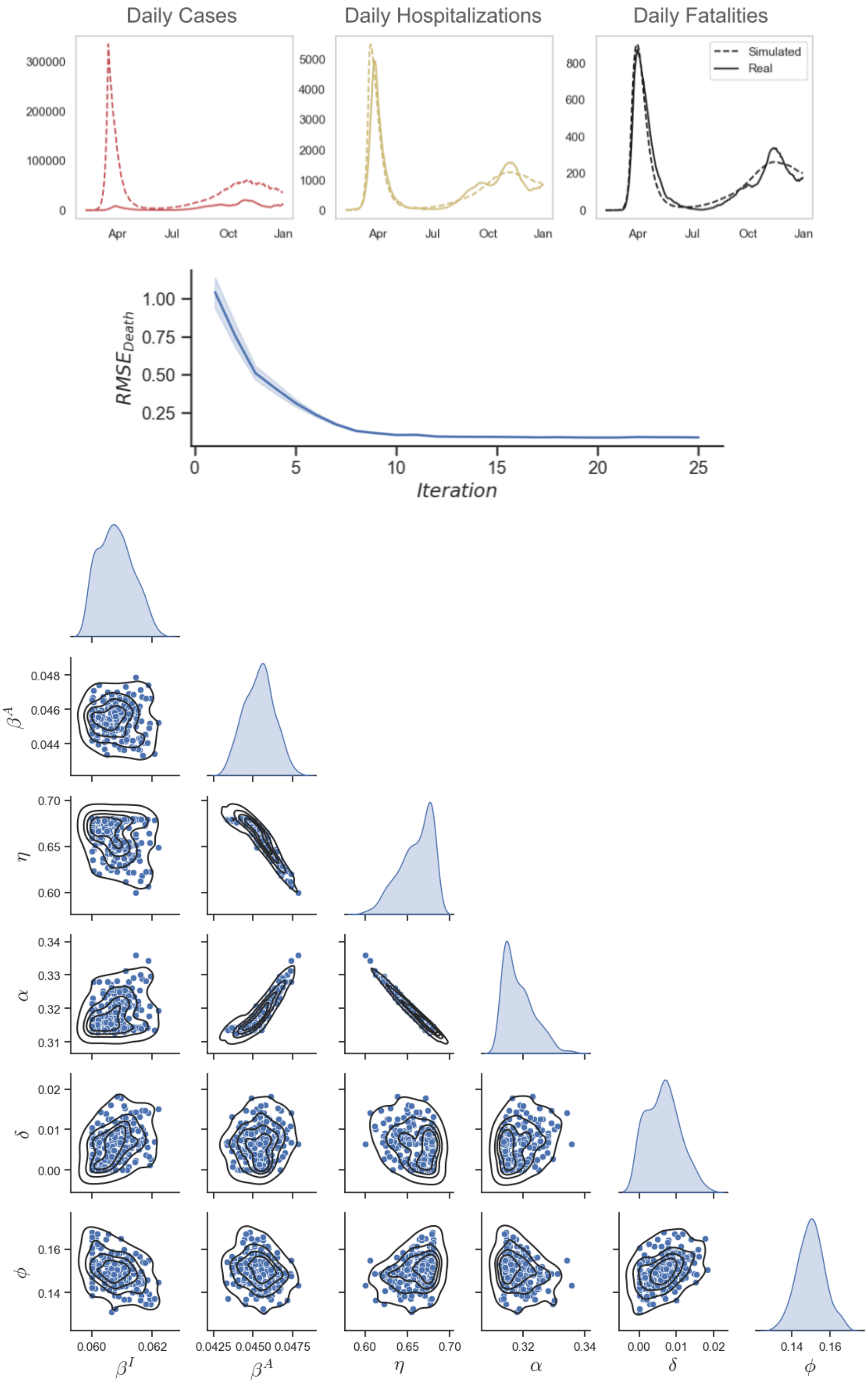


Figure 26: Parameter calibration results. The top panel shows the daily incidence, hospitalizations and fatalities simulated with the best parameter set together with the real reported data. The middle panel shows the CMA-ES algorithmic convergence for the minimization of the RMSE between simulated data and reports

on fatality rates for Spain at the country level. The panel at the bottom shows the final distribution of parameters learned by the CMA-ES algorithm.

Finally, we compare the time series of fatalities produced by the best parameter set to the real data. However, instead of comparing at the country level, we focus on the prediction at the Autonomous Community level. The results are depicted in Figure 27, which shows how the first wave is very well captured for most communities except for Castilla la Mancha and Galicia. Interestingly the second wave is also qualitatively recovered for several communities. We observed that the major discrepancies in the second wave are for the cases of the regions corresponding to islands (Canarias and Baleares) and for the autonomous cities of Ceuta and Melilla. We hypothesize that in the former case, the discrepancy could be due to changes in the mobility network that protect the island from importing cases from other regions. Currently, the simulator considers a static mobility network and therefore any change in the structure of the mobility network due to human behaviour is not considered by the model. In the cases of Ceuta and Melilla, we found that the reports are noisy and are probably subject to underreporting.

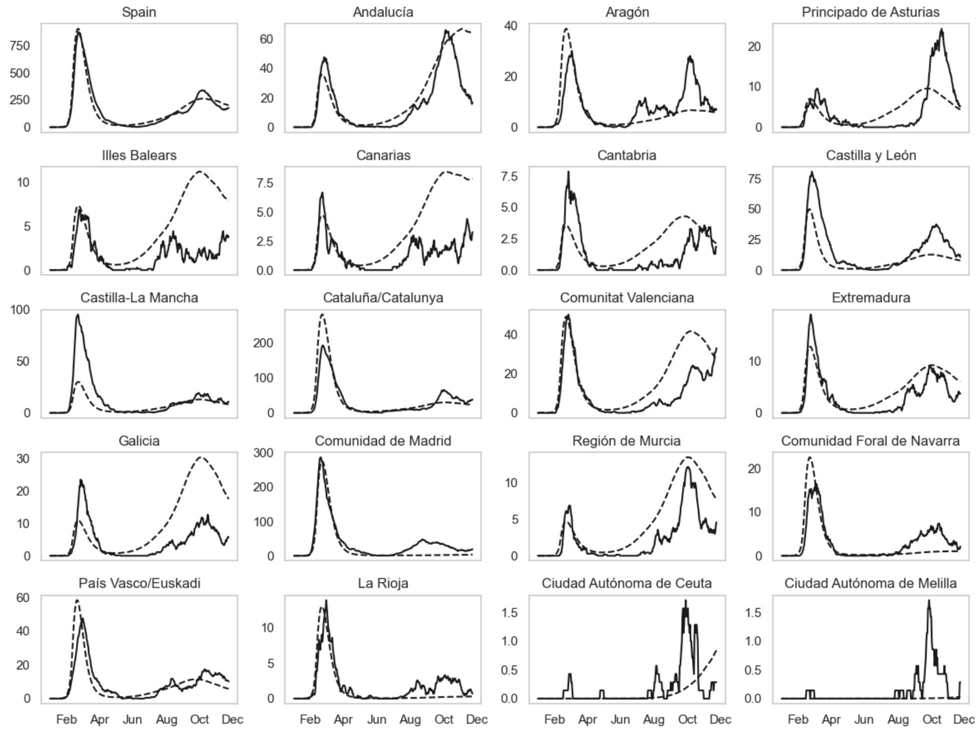


Figure 27*:* Daily fatalities in Spain during 2020 for the whole country and at the level of autonomous communities. Solid and dotted lines correspond to real reports and simulated data, respectively.

Altogether, we have tested the work for calibrating epidemiological parameters running optimization-via-simulation and found the candidate parameters fit well the time series of fatalities and hospitalization for the first and the second wave at the country level.

Interestingly, the simulations also fit the time series for several autonomous communities. In future work, we will use visual analytics (T5.3) to evaluate the uncertainty of the predictions and apply causal inference approaches to understand the patterns of spreading.

Modelling confinement measures

The MMCACovi19-vac simulator allows modelling the effect of applying non-pharmaceutical intervention in the form of confinement measures and vaccination. Confinement strategies are based on social distancing and isolating a portion of the population. This approach decreases the average number of contacts and mobility, thereby reducing the likelihood of infection and ultimately lowering the overall prevalence of the disease. Confinement policies are introduced in the formalism by different parameters described in Table 9.

Table 9: Mobility Reduction parameters

Symbol	Description
t^cs	Time steps when the containment measures will be applied
κ_0s	Mobility reduction
ϕs	Household permeability
δs	Social distancing factor

The parameter $\kappa_o^g(t)$ represents the fraction of the population within age stratum g that is under lockdown at time t. The parameter is bounded in the [0,1] interval and a $\kappa_o^g(t)$ equal to 0 or 1 represents no restriction or a total lockdown, respectively. We assume the same value for all the age strata. The parameter δ models the reduction of the number of contacts made by the non-confined population and is also bounded in the [0,1] interval. Finally, the household permeability parameter ϕ accounts for the social mixing among members from different households in those situations where members of a given household are required to go out for essential activities such as buying groceries, drugs, etc and thus interact with members of different households. Although interactions among members of different confined households influence household isolation, we assume they do not significantly alter the average number of contacts within the population. Therefore, in the current implementation, the household permeability ϕ is kept constant over time.

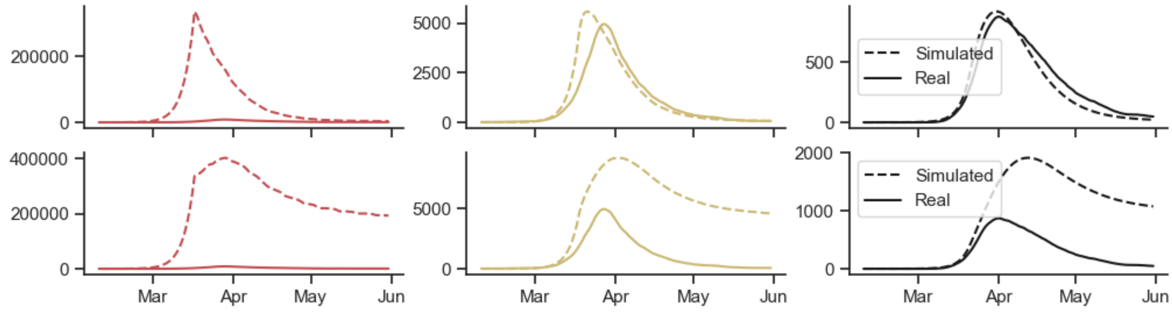


Figure 28: Daily fatalities in Spain during 2020 for the whole country and at the level of autonomous communities. Solid and dotted lines correspond to real reports and simulated data, respectively. The top panel shows the results of applying a mobility reduction policy and the bottom panel without any measurement applied.

We ran two simulations with no NPI applied and with the parameter that model, the NPI applied in Spin during the first wave. Figure 28 shows the results of the two simulations and how in the absence of NPI the number of hospitalizations at the peak of the wave is almost twice the value obtained when NPI was applied. For the case of mortality, the simulation predicts that in the absence of an NPI, the total number is ~100,000 individuals whereas with the NPI this number drops to ~25,000. These results clearly show that in the absence of vaccines, NPI can be a good approach to control epidemics.

The MMCACovi19-vac simulator also allows the modelling of the effect of applying vaccination. Vaccination is modelled by adding an extra compartment in the infectious model which has different epidemiological parameters than those used for non-vaccinated individuals. Vaccination strategies are introduced in the formalism by different parameters described in Table 10. Currently, we are working on a Reinforcement Learning (T4.2) approach to optimize the design of effective vaccination campaigns.

Table 10: Mobility Reduction parameters

Symbol	Description
ϵ^g	Total vaccinations per age strata
start_vacc	Start of the vaccination
dur_vacc	Duration of the vaccination
end_vacc	End of the vaccination
$t^v s$	[start_vacc, end_vacc, T]

An early investigation into the best-performing vaccination strategy is shown in Figure 29 where scenarios with vaccination with different starting dates, age priorities and durations were performed and then compared to a baseline scenario without vaccinations, all while trying to minimize two quantities: the relative variation of the height of the second peak and the relative variation in the number of deaths. The strategies that outperformed all the others in at least one of these measures are said to be Pareto-optimal and lie on the Pareto front. Altogether, the results show that according to our model, the best strategies reduce deaths by between 20 and 50% and reduce the second peak by 5 and 17%.

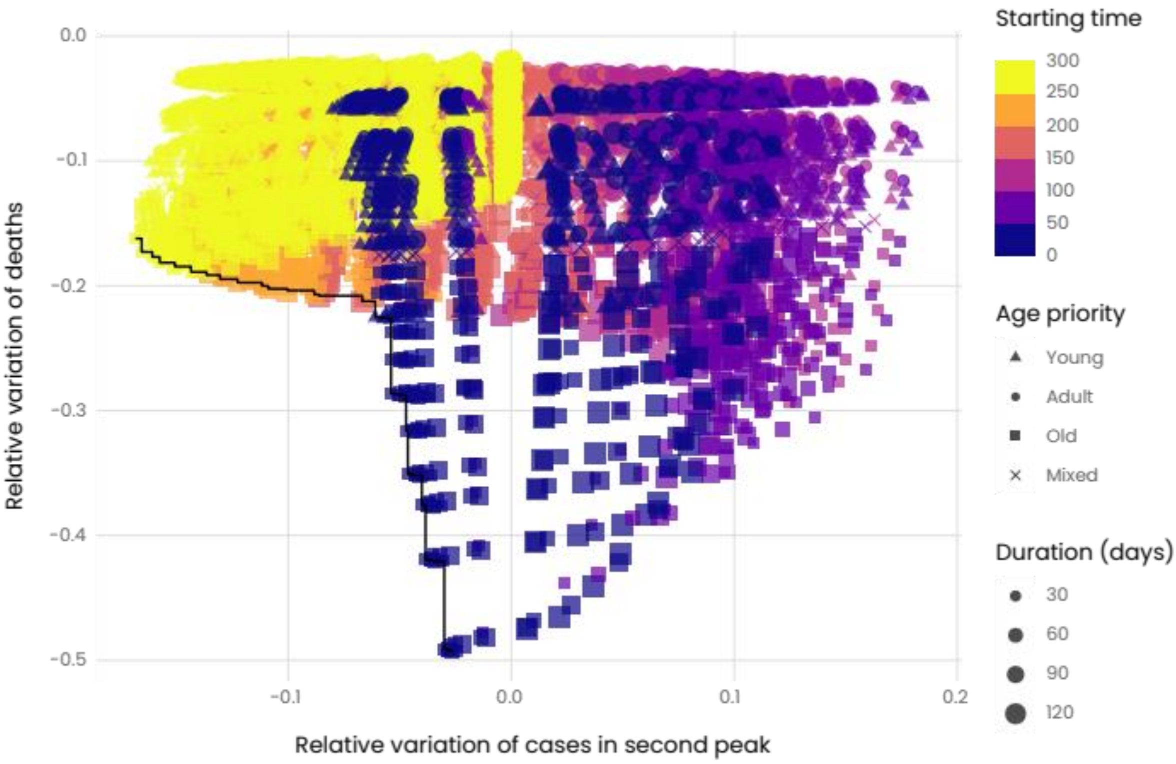


Figure 29: In this scatter plot, each point corresponds to a unique simulation based on a distinct vaccination strategy. The horizontal axis displays the relative variation in the number of cases at the next largest peak relative to the baseline case, while the vertical axis represents the relative variation in the number of cumulative deaths, both concerning a baseline without vaccinations. The black line links together all the Pareto-optimal solutions.

Two types of optimal strategies are found: the first included strategies that started early and were good at lowering the number of deaths, while the second started later and performed better at lowering the second peak. In both of these cases, the preferred age group were the elderly. Similar results were also obtained if the height of the second peak was swapped out for the height of the peak of hospitalizations.

4.4.2 Multi-scale infection scenario

The multi-scale infection scenario begins with the definition of the workflow (step 6), a crucial step that outlines the sequence of operations to be performed in the simulation. This workflow serves as a roadmap for the project, guiding the implementation of each step and ensuring that all necessary tasks are accounted for. The implementation of the workflow is currently a work in progress. This involves translating the defined workflow into executable code in RapidMiner, which will drive the Alya and PhysiBoSS simulations. The successful implementation of the workflow is key to the smooth operation of the simulations and the generation of accurate results and will be the focus of the work in the coming months.

A significant achievement of the project is the successful connection from Alya to PhysiBoSS (step 6). This allows the results of the Alya simulation, which models the airflow in the upper airways, to be fed into the PhysiBoSS simulation, which models the infection dynamics of COVID-19 in the lung epithelia. This connection enables the integration of the two simulations, allowing for a comprehensive understanding of the interplay between airflow and infection dynamics (Figure 30). The connection from PhysiBoSS back to Alya is currently a work in progress. Once completed, this will allow the results of the PhysiBoSS simulation to feedback into the Alya simulation, creating a dynamic, interactive model that can adapt to changing conditions.

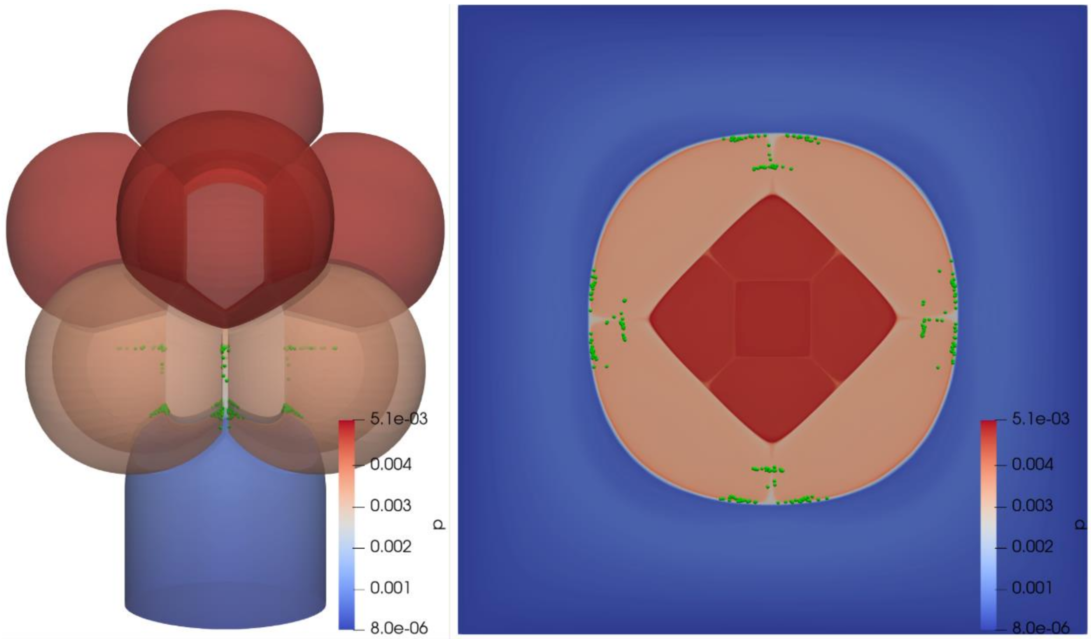


Figure 30: Conversion of Alya's 3D coordinates to PhysiBoSS' 2D coordinates to connect both simulators. Colour gradient represents air pressure (used as a proxy for oxygen content) and green spheres are virus particles.

The project has successfully conducted Alya simulations with viruses and PhysiBoSS simulations with different O2 levels and viruses (steps 2 and 3). These simulations provide valuable insights into the spread of the virus within the airways and the progression of the infection in the lung epithelia (Figure 31).

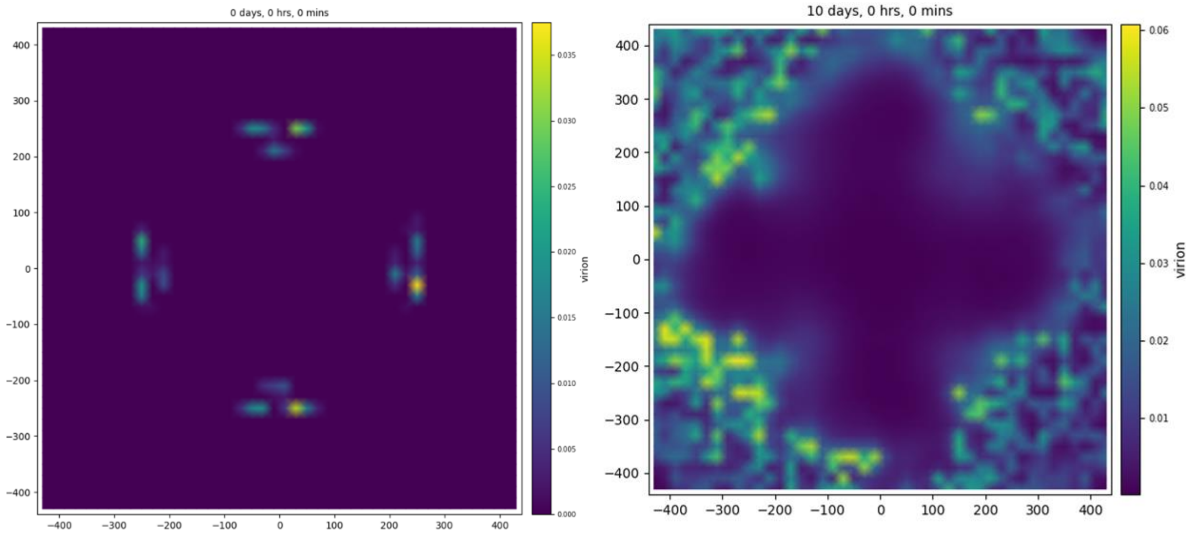


Figure 31: Results of the PhysiBoSS simulations using oxygen and virus data from Alya. Colour gradient represents density of virus particles. Left panel, initial state of the simulation; Right panel, final state of the simulation.

The project has also successfully incorporated personalised models of different patients into the PhysiBoSS simulations (step 1 and 4) and different airways from patients into the Alya simulations (step 1). This allows the simulations to consider individual variations, which can significantly influence the course of the disease and the effectiveness of interventions.

The incorporation of Early time series classification from WP4 into the workflow is currently a work in progress (step 7). Once completed, this will allow the system to classify time series data at an early stage, potentially enabling early detection and intervention of disease progression. Likewise, the incorporation of Forecasting into the workflow is in the project's scope. This will enable the system to predict future disease progression based on current and historical data, providing valuable foresight for healthcare providers and patients.

We are working on achieving the KPIs of this use cases by incorporating CREXDATA technologies to the use case and finalising the implementation of the workflow:

- Forecasting 7 parameter sets that reduce the COVID infection;
- Use the runtime adaptation of simulation trajectories to improve the outcomes of 5 scenarios or patients; and
- Characterising the space of parameters with 50% fewer simulations.

In conclusion, the demonstrator project has made significant progress in simulating the dynamics of COVID-19 infection in the lungs and the impact of airflow in the upper airways. The integration of AI technologies and multimodal data highlights the potential of this approach in understanding the different severity of COVID-19 in patients and optimising interventions. The project continues to work towards the completion of the remaining tasks, with the aim of further refining the simulations and enhancing the effectiveness of the system.

4.5 Perspectives

In this section we describe the work to be carried on this use case in the following months.

4.5.1 Epidemiology scenario

The epidemiological modelling scenario aims to understand and simulate the spatiotemporal patterns of the COVID-19 pandemic in Spain. Additionally, it should also provide a tool to support decision-making by helping to identify effective confinement and vaccination strategies that must work under different circumstances. We have successfully defined a reference model for Spain and collected a data set of COVID-19 reports and population mobility that is used as the gold standard for parameter calibration.

Furthermore, we have extended the MMCACovid1 simulator providing new modelling features, a standard file format to define model parameters and a command line script to easily run simulations. We have also completed the first version of the model exploration workflow that currently allows running optimisation-via-simulation in HPC infrastructures. We have implemented three different search strategies that include a simple parameter sweep together with two different metaheuristics from the family of evolutionary algorithms, named Genetic Algorithm and Covariance Matrix Adaptation. The model exploration workflow allows parallel evaluation of different simulation instances and can easily scale a large number of computing nodes.

Currently, we have used the Model Exploration Workflow to calibrate a subset of epidemiological parameters and found that the best candidate reproduces the observed trends of daily hospitalization and fatalities. Furthermore, we have evaluated a scenario where no confinement measurements were applied and found the impact could have been catastrophic in terms of the number of deaths and the collapse of the health system. In this way we have partially fully two of the three KPI proposed for the project. We still have to extend the calibration and implement an optimization approach based on Reinforcement Learning to design effective interventions. Additionally, we are also working in the Active Learning workflow to characterize the parameter space with a reduced number of evaluated simulations.

In addition, FMI experts participate in the Health Crisis Use Case by aiming to contribute novel results for the pandemic response during the COVID-19 pandemic. This task will deliver new information on how emergency calls are influenced by COVID-19 and critical weather situations. The task aims to develop new impact-based forecasting tools for the health sector emergency and pandemic response. Furthermore, a scientific research article focused on emergency tasks both during and outside the COVID-19 pandemic in Finland will reveal new insights into how the health sector can prepare for future risks. Research on pre-hospital service operations (ambulance calls) in Helsinki with a statistical regression model DLNM (distributed lag nonlinear model) has not been previously conducted to understand COVID-19 and weather impacts, especially temperature-related impacts. Discussions with the emergency department about the ambulance study have particularly focused on selecting research topics that have an impact on daily services and are useful while planning future services. Thus, for instance, narrowing the research question to the most frequent number of calls type will give more exact data on what type of emergencies occur during different times of the year and their possible relationship with different weather conditions. The most common emergency tasks include falling, chest pain, breathing problems, brain haemorrhage but also poisoning. These will be studied in detail together in the Emergency Use Case. Sub-daily test forecasts with the Gradient Boosting model have not shown strong weather correlations and thus the modelling is planned to be conducted with daily forecasts.

Looking forward, the next steps of the project are clear and promising. The primary focus will be on completing the remaining tasks and further refining the simulations and analyses. This includes

1. **Integration of AI Technologies:** The integration of other CREXDATA AI technologies such as Federated Machine Learning, Interactive Learning for Simulation Exploration, and Visual Analytics for Decision Making under Uncertainty, will be further evaluated and incorporated into the use case if needed.

2. **Validation:** The results of the simulation will need to be validated against experimental or clinical data. This step is crucial to ensure that the simulations accurately represent the real-world dynamics of COVID-19 infection.

3. **Usability Evaluation:** The ease of use and user-friendliness of the system will be assessed to ensure that the use case is not only effective but also user-friendly and accessible.

These next steps will bring the project closer to its goal of understanding the spatiotemporal pattern of the COVID-19 pandemic as well as what are the best strategies to control the spreading of the disease. The insights gained from this project could potentially inform the development of effective interventions such as lockdowns and vaccination campaigns.

4.5.2 Multi-scale infection scenario

The multi-scale modelling project, aimed at simulating the dynamics of COVID-19 infection in the lungs and the impact of airflow in the upper airways, has made significant progress. The successful integration of Alya and PhysiBoSS simulations, along with the incorporation of patient-specific omics data, has provided valuable insights into the different severity of COVID-19 in patients and potential optimization of interventions.

Looking forward, the next steps of the project are clear and promising. The primary focus will be on completing the remaining tasks and further refining the simulations and analyses. This includes:

1. **Completing the Workflow Implementation:** The workflow, which has been defined, now needs to be fully implemented. This will involve translating the defined workflow into executable code that will drive the Alya and PhysiBoSS simulations.
2. **Establishing the PhysiBoSS to Alya Connection:** The connection from PhysiBoSS back to Alya is currently a work in progress. Once completed, this will allow the results of the PhysiBoSS simulation to feedback into the Alya simulation, creating a dynamic, interactive model that can adapt to changing conditions.
3. **Incorporating Complex Event Forecasting and Runtime Adaptation into the Workflow:** The incorporation of forecasting and runtime adaptation into the workflow

is currently in progress. Once completed, this will allow the system to classify time series data at an early stage, predict future disease progression based on current and historical data, and adapt the simulation outcomes in real-time. This could potentially enable early detection and intervention of disease progression and improve the outcomes of the simulations based on real-time data.

4. **Validation:** The results of the simulation will need to be validated against experimental or clinical data. This step is crucial to ensure that the simulations accurately represent the real-world dynamics of COVID-19 infection.
5. **Usability Evaluation:** The ease of use and user-friendliness of the system will be assessed to ensure that the use case is not only effective but also user-friendly and accessible.
6. **Integration of AI Technologies:** The integration of other CREXDATA AI technologies such as Federated Machine Learning, Interactive Learning for Simulation Exploration, Optimised Distributed "Analytics as a Service", Explainable AI, Visual Analytics supporting XAI, and Visual Analytics for Decision Making under Uncertainty, will be further evaluated and incorporated to the use case if needed.

These next steps will bring the project closer to its goal of understanding the different severity of COVID-19 in patients and optimising interventions. The insights gained from this project could potentially inform the development of new treatment strategies or preventive measures, ultimately contributing to the fight against COVID-19.

5 Maritime Use Case

5.1 Scenario description

The maritime use case involves the testing of components from the CREXDATA platform and related algorithms for forecasting hazardous events, such as collisions or areas of hazardous weather, and rerouting vessels to safety. This research is relevant to various stakeholders, including Vessel Traffic System operators, deck officers or vessel crew, coastal authorities, vessel remote operators and others, as outlined in D2.1. With the increasing introduction of autonomous ships and related technologies, these components are becoming even more important especially for remote operators. The Maritime Use Case aims to enable early detection and forecasting of maritime events, allowing for proactive measures to prevent incidents.

In the context of this use case, the following objectives will be pursued:

a) Development of a novel IoT device to access real-time data from the "black box" of a vessel.
b) Fusion of vessel data with global data sources (e.g., AIS and Copernicus data) to create reliable digital twins of vessels.
c) Development of highly scalable route forecasting algorithms.

The end-users will have the capability to view forecast vessel motion and predicted routes, along with confidence levels for each route. Key performance indicators (KPIs) include achieving at least 80% accuracy in route forecasting/weather routing and hazardous event detection/forecasting, sub-second latency in route forecasting/weather routing and event detection/forecasting over streaming data, and forecasting maritime hazardous events 15 minutes before they occur. The pilot will be validated through a sea trial experiment.

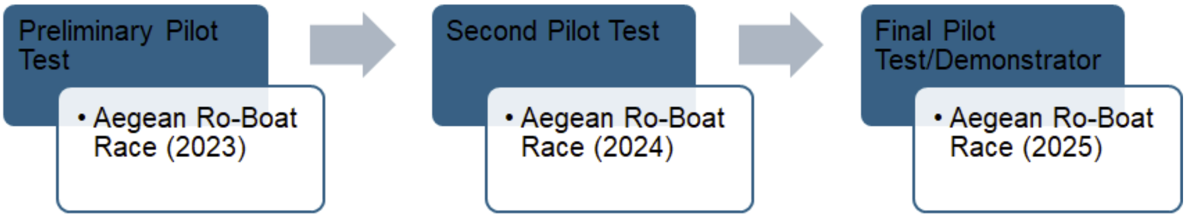
The purpose of this use case is to develop a comprehensive solution that combines hardware and software development to utilize vessel data, fused with global data, in order to create reliable digital twins of vessels. Additionally, it aims to develop weather and emergency routing solutions that can be performed for all vessels of a fleet simultaneously, leveraging big data and AI technologies.

5.2 Pilot definition and demonstrator

The use case involves the comprehensive development and testing of various components and functionalities for both manned and unmanned vessels. This includes conducting multiple sea tests and making preparations for upcoming tests. These tests are essential for evaluating the performance and capabilities of the vessels in different scenarios and conditions at sea.

Within this context the Aegean Ro-boat Race[9] is directly related to the pilot tests as it provides a real-world environment for testing the vessels' capabilities. The tests and competition are crucial for the project as they offer opportunities to assess and validate the performance of the autonomous vessels in challenging and dynamic maritime conditions.

9 https://smartmove.aegean.gr/events/1st-aegean-roboat-race/



1. The Preliminary Pilot Test conducted in 2023 was focused on receiving data from the competing vessels while they were at sea. This involved the installation of the IOT box on several autonomous vessels participating in the Aegean Race in July 2023. The data collected during this test will be used to evaluate and refine the collision detection algorithms.
2. In 2024, the Second Pilot Test will expand its focus to include not only receiving data but also sending commands to the vessels (through the IOT Box), allowing for the alteration of their behaviour. This test will provide valuable insights into the vessels' responsiveness to external commands and their ability to adapt to different operational instructions. This will take place during the Aegean Ro-boat Race in the summer of 2024.
3. The Final Pilot Test/Demonstrator scheduled for 2025 represents the culmination of the maritime use case, where both receiving and sending complex commands and behaviours to the vessels while at sea will be tested. This final test will serve as a demonstrator of the advanced capabilities of the autonomous vessels, showcasing their readiness for real-world maritime applications. This will take place during the Aegean Ro-boat Race in July 2025.

In summary, the Aegean Ro-boat Race and the pilot tests are closely interconnected, with the race providing a practical platform for testing the vessels' performance, and the pilot tests serving as essential stages for evaluating and refining their capabilities for future maritime operations.

The Aegean Ro-boat Race is an international university level competition, challenging teams to design and develop innovative autonomous robotics systems that can perform at sea in real world conditions. The main aim of this competition is to foster innovation, while strengthening ties between academic institutions and industry. The goal of the Aegean Ro-Boat Race Is to push the current state of the art technology and approaches in marine robotics to their limits, as would happen in real-world ocean environments. The autonomous boats would have to demonstrate “seaworthiness”, that is fit for the normal perils of the sea. More specifically:

1) Each student team should design and built their own autonomous vessels from scratch. Teams were not limited by being given identical vessels to equip and race, so as to give room for completely innovative designs and constructions, as opposed to other maritime competitions where commercial solutions were utilized. Teams were supported by industrial partners but could not compete with an off the shelf solution.
2) The race had to take place in a completely dynamic and unknown environment, even if this meant that there would be a high failure rate of competing teams. To build reliable, near turn-key solutions capable of a long duration performance it is necessary to push the capabilities and performance of autonomous systems.

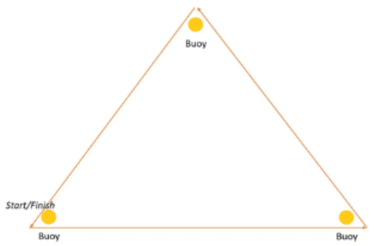
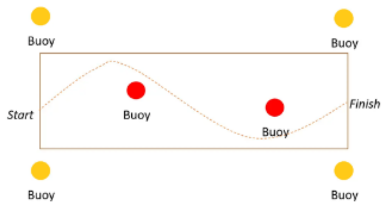
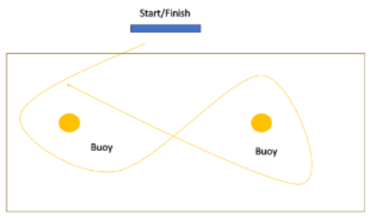
For research to progress in the field of autonomous maritime systems, the availability of real-world data is key. To realize the full value of the research data it needs to be accessible to the wider research community (not only to involved teams/organizers), under the Findability, Accessibility, Interoperability, and Reusability (FAIR) principles. For this, all participants were committed to publishing the data under open and FAIR principles after the completion of the race (Under review in data in brief).

The 2023 competition consisted of three mission tasks focused on: high speed and performance, collision avoidance, and endurance. Similar to any yacht race, the team with the best overall win/loss record at the end, won the regatta. Points were won for each successful race while points were reduced for collisions between the vessels, piers, and obstacles. The overall field dimensions were (100×70) meters and it was delimited by triangular external large (at least 1 meter high) bright yellow buoys. The obstacles that were dispersed within the field were smaller orange buoys (used in the collision avoidance race). Prior to the race day, a collection of images of the buoys were provided to all competing teams to assist in object recognition (Figure 32). Before each race, the position of the external yellow buoys were provided.

Figure 32: Collection of images of the buoys provided to all competing teams to assist in object recognition in the context of Aegean Ro-boat Race 2023.

The contest was composed of three races, which were held consecutively on the day of the competition. As opposed to other competitions, the teams raced the track together for two of the races (collision avoidance and endurance). Thus, creating a completely dynamic environment where boats had to consider the current traffic and avoid other boats while recognizing the static obstacles. At the beginning of each scoring run, the vessel was remotely controlled to reach the start point. Once the boat left the start line, no interaction with the boat by human operators was allowed during the entire run. More details for the three challenges are given in Table 11.

Table 11: Objectives and buoys setup for speed-race, collision-avoidance and endurance race

High-speed race: Each boat had to race across the track at the highest possible speed for one lap. The aim of this race was to cross the finish line first, while testing the overall performance of the vessel. The racetrack was configured as an equilateral triangle with the yellow buoys marking the vertices. The starting line was near the pier.	
Collision avoidance: The boats needed to travel throughout the course avoiding dispersed static obstacles and the other boats that were racing at the same time. The positions of the obstacles were unknown to the teams. For each collision, 2 points (boat or buoy) were removed.	
Endurance race: The boats needed to perform as many laps as possible around the track. This would ultimately test the performance and reliability of the systems. The last boat going or the boat with the most laps would win the race (10 points), while points would be reduced for each collision with other boats or drifting out of the boundaries.	

The Teams

In total, seven university teams registered for the 2023 competition, with five of them being from Greece and 2 from the countries of Portugal and Latvia (Figure 33 and Figure 34).

Figure 33: The participants of the Aegean Ro-boat race 2023

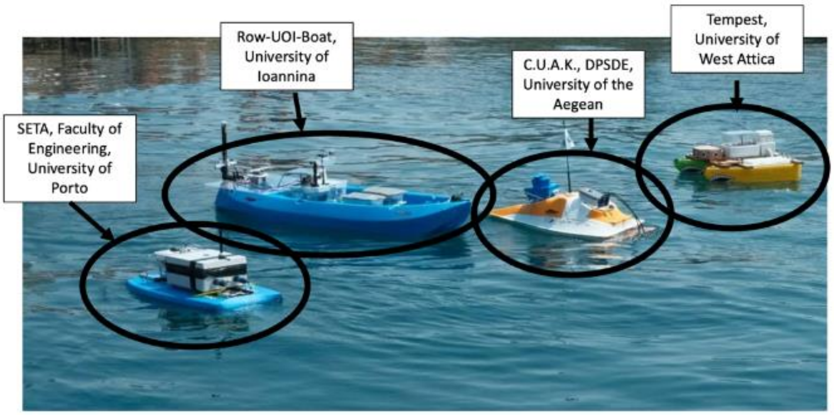


Figure 34: Autonomous boats that competed in the Aegean Ro-boat race 2023

5.3 Simulation tools

5.3.1 Synthetic dataset for testing collisions

The Automatic Identification System (AIS) enables ships to share their identification, characteristics, and location data through self-reporting mechanisms. This information is periodically broadcast and can be intercepted by other vessels equipped with AIS transceivers, as well as ground-based or satellite sensors. Following the International Maritime Organization (IMO) mandate for AIS installation on vessels exceeding 300 gross tonnage, vast datasets have been generated, serving as a crucial resource for maritime intelligence.

Maritime collisions occur when two vessels collide or when a vessel strikes a floating or stationary object, such as an iceberg. These collisions are particularly significant in the context of marine accidents for several reasons:

1. Injuries and fatalities among crew members and passengers.

2. Environmental damage, especially in incidents involving, shipwrecks, large tanker ships and oil spills.
3. Economic losses, both direct and indirect, impacting local communities near the accident site.
4. Financial repercussions for ship owners, insurance companies, and cargo owners, including vessel loss and penalties.

With the increase in sea traffic and vessel speeds, the likelihood of major accidents during a ship's operational life has risen. The growing congestion on sea routes heightens the risk of accidents, particularly collisions between vessels.

Developing solutions and models for analyzing, detecting early, and mitigating vessel collision events is crucial for future maritime safety. In this context, a synthetic vessel proximity event dataset has been created by Kpler using actual vessel AIS messages originating from the MarineTraffic AIS terrestrial network. This synthetic dataset of trajectories, with reconstructed timestamps, is designed so that two vessel trajectories intersect simultaneously, simulating an unintended proximity event (collision close call). The dataset aims to provide a foundation for developing methods to detect and mitigate maritime collisions and proximity events, as well as for training vessel crews in simulator environments. So far, the Synthetic AIS Dataset of Vessel Proximity Events has been used for the evaluation of the vessel collision forecasting approach developed by Kpler [1]. In the context of the CREXDATA Maritime Use Case, the synthetic AIS proximity dataset has been used for developing, finetuning and testing the vessel collision avoidance algorithm.

The dataset includes 4658 samples/AIS messages from 213 unique vessels in the Aegean Sea. The process to create the collision dataset involved the following steps: Given two vessels, X (vessel_id1) and Y (vessel_id2), with their known locations (LATITUDE [lat], LONGITUDE [lon]):

1. Verify if the trajectories of vessels X and Y intersect spatially.
2. If the trajectories intersect, align the timestamp of vessel Y at the interSection point with the timestamp of vessel X at the same point. This temporal alignment ensures that the spatial interSection (nearest proximity point) occurs simultaneously for both vessels.
3. Ensure that for each vessel pair, the timestamp of a proximity event is unique and does not overlap with other proximity events occurring later, preventing temporal overlap of different vessel trajectory pairs.

The synthetic collision dataset has been open-sourced and is available under a Creative Commons Attribution 4.0 International License at the Zenodo platform [2].

Specifically, two CSV files are provided. vessel_positions.csv contains the AIS positions, vessel_id, t, lon, lat, heading, course, and speed of all vessels. Simulated_vessel_proximity_events.csv includes the id, position, and timestamp of each identified proximity event, along with the vessel_id numbers of the associated vessels. The final count of unintended proximity events in the dataset is 237. An overview of the sea area and a sample of multiple simulated vessel collision events is presented in Figure 35.

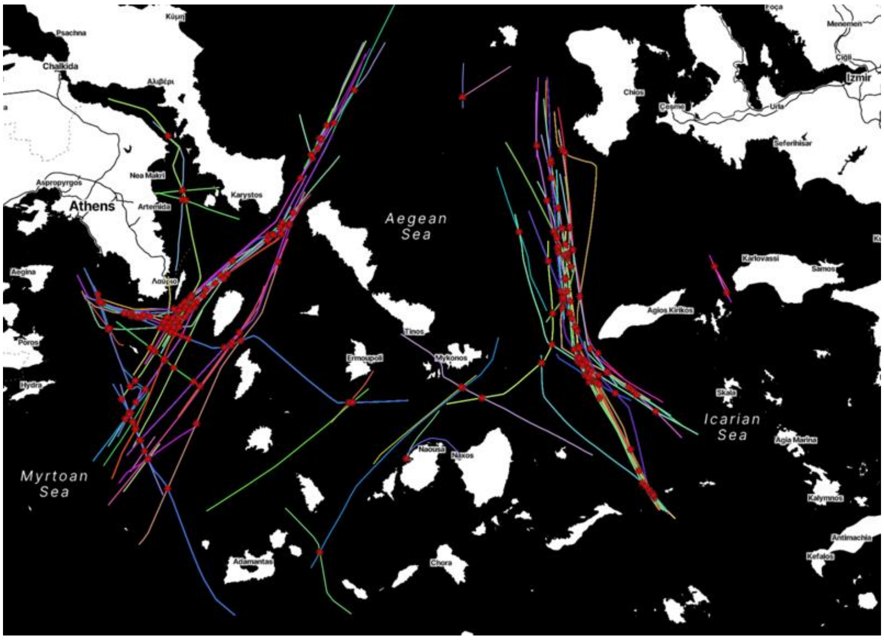


Figure 35: Example of simulated vessel proximity events in the Synthetic Vessel AIS Dataset

5.3.2 Algorithms and system component evaluation

Two distinct components for forecasting and resolving critical complex maritime events are being developed by Kpler:

1. MAR_1: Collision forecasting and rerouting
2. MAR_2: Hazardous weather routing

Until M18, the internal development and the evaluation of both MAR_1 and MAR_2 has been finalized. The methodological approach followed for the development of MAR_1 and MAR_2 is presented in detail in Deliverable 4.1 sub-Section 2.4. Finally, the system requirements for the integration of MAR_1 and MAR_2 with the CREXDATA platform and the setup of the RapidMiner Operators have been defined. Figure 36 presents the Kpler system architecture supporting the deployment of the Maritime Use Case pilot.

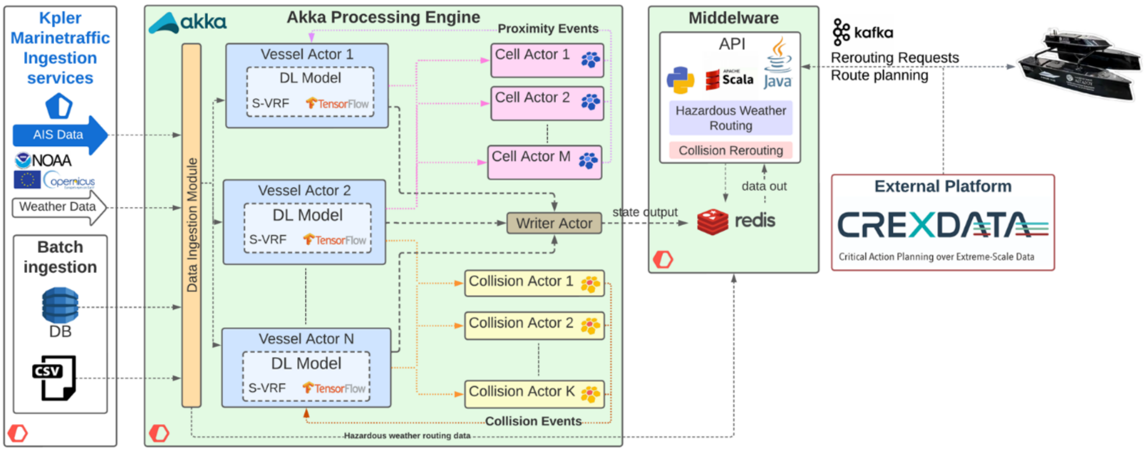


Figure 36: Kpler system architecture for the Maritime Use Case

MAR_1: Collision forecasting and rerouting initial component evaluation

The synthetic AIS proximity dataset has been used for experimentation, testing and first evaluation of the collision avoidance solution [2]. Example evaluation results for each category of COLREG interaction are presented in Figure 37. First results indicate that the the vessel collision avoidance algorithm developed in the context of the CREXDATA project is able to generate COLREG compliant routes with respect to a forecasted proximity event with another AIS-equipped vessel. The developed solution is able to identify the type of COLREG interaction based on the forecasted short-term route and course of the respective vessels and select the optimal frenet path that complies with the COLREG rules and the defined collision avoidance safety radius.

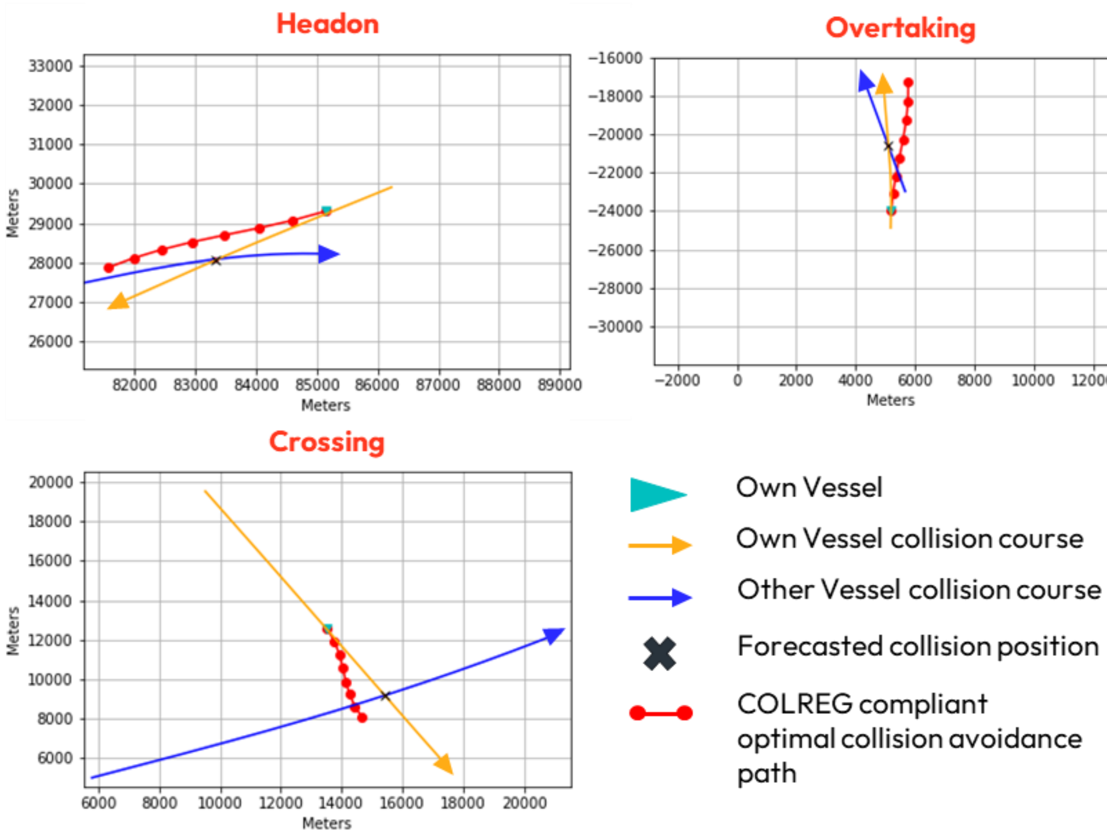


Figure 37: Collision Avoidance Results using the synthetic AIS proximity dataset [12]

Error! Reference source not found.Error! Reference source not found. *MAR_2: Hazardous weather routing initial component evaluation*

A first evaluation of the proposed hazardous weather routing solution has been completed until M18. For the purposes of testing the hazardous weather routing solution a weather dataset that fulfils certain requirements had to be extracted.

In order to simulate long-trips that last more than a few hours, a large sea area should be considered. Additionally, for experimentation with extreme weather conditions the dataset has to include a significant number of extreme weather events in terms of significance, size and duration for efficiently evaluating the hazardous weather routing solution that focuses primarily on extreme events. Therefore, specific sea regions are more appropriate than others.

After evaluating historical weather reports the sea area of the Gulf of Mexico and the US East Coast has been selected. Historically, in this area a high frequency and intensity of “Atlantic hurricanes” (i.e. tropical cyclones that occur on the Atlantic Ocean, primarily between the months of June and November) is observed. In this context the month of September 2022 has been selected for the evaluation of the hazardous weather routing solution, where according to statistical data, the highest number of Atlantic Hurricanes for 2022 has been observed [3].

40 origin and destination queries have been manually executed across the specific sea area. An effort has been made to select origin and destination pairs so that the formedtest routes pass through sea areas with diverse weather conditions. As the proposed solution is developed primarily for alerting and automatically rerouting vessels at areas of extreme weather events it is expected that in the case of normal weather conditions generated routes would still closely follow the shortest route to destination. Thus, the main aim of the evaluation is to determine whether the hazardous weather routing approach consistently generates routes for vessels that minimize the total weather penalty across the route in comparison to state of the art approaches such as the shortest path. Additionally, through visual inspection of the results in cases of extreme weather events it is important to evaluate that the hazardous weather routing solution routes vessels away from sea areas with extreme weather conditions as thisthese are considered through the expected forecasted values for the critical weather features.

Results for all 40 queries with respect to the mean wave height achieved through the shortest path and the hazardous weather routing approaches (Pmean). Pmean considers the average forecasted wave heights across the entire route of each routing method. At the same time, Pmax considers the maximum wave height experienced along each generated route. In an effort to evaluate that the hazardous weather routing method diverts vessels away from areas with extreme weather conditions. Results are presented in Table 12. Results showcase that the hazardous weather routing method effectively reduces the wave height along the route, indicating that the hazardous weather routing method follows routes with calmer weather conditions.

Table 12: Initial evaluation results for the hazardous weather routing solution against state-of-the-art routing methods (shortest path)

Routing method	count	Pmean Wave Height (m)		Pmax Wave Height (m)	
		mean	std	mean	std
shortest	40	2.64	0.90	6.02	3.06
weather	40	2.48	1.08	5.66	3.52
% Difference		-6%	19%	-6%	15%

In a second step, an origin destination query is selected to evaluate the hazardous weather routing methodapproach in the case of a registered extreme weather event. Due to the scarcity of such events the origin and destination position of the query were selected in such a way in order to ensure that the vessel route would be significantly impacted by the extreme weather conditions. Specifically, a query and the respective route across the Hurricane Fiona formed in September 2022 is selected for evaluation. Results are presented in Figure 38. Results demonstrate under consideration of the results from Table 12 the increased performance achieved by the hazardous weather routing solution specifically in the case of very extreme weather events (Figure 38a). Moreover, Figure 38b and Figure 38c demonstrate that the hazardous weather routing solution is capable of entirely avoiding areas with forecasted extreme weather conditions. This is specifically highlighted in Figure 38c where the route does not pass through sea areas with high waves (≥ 6m).

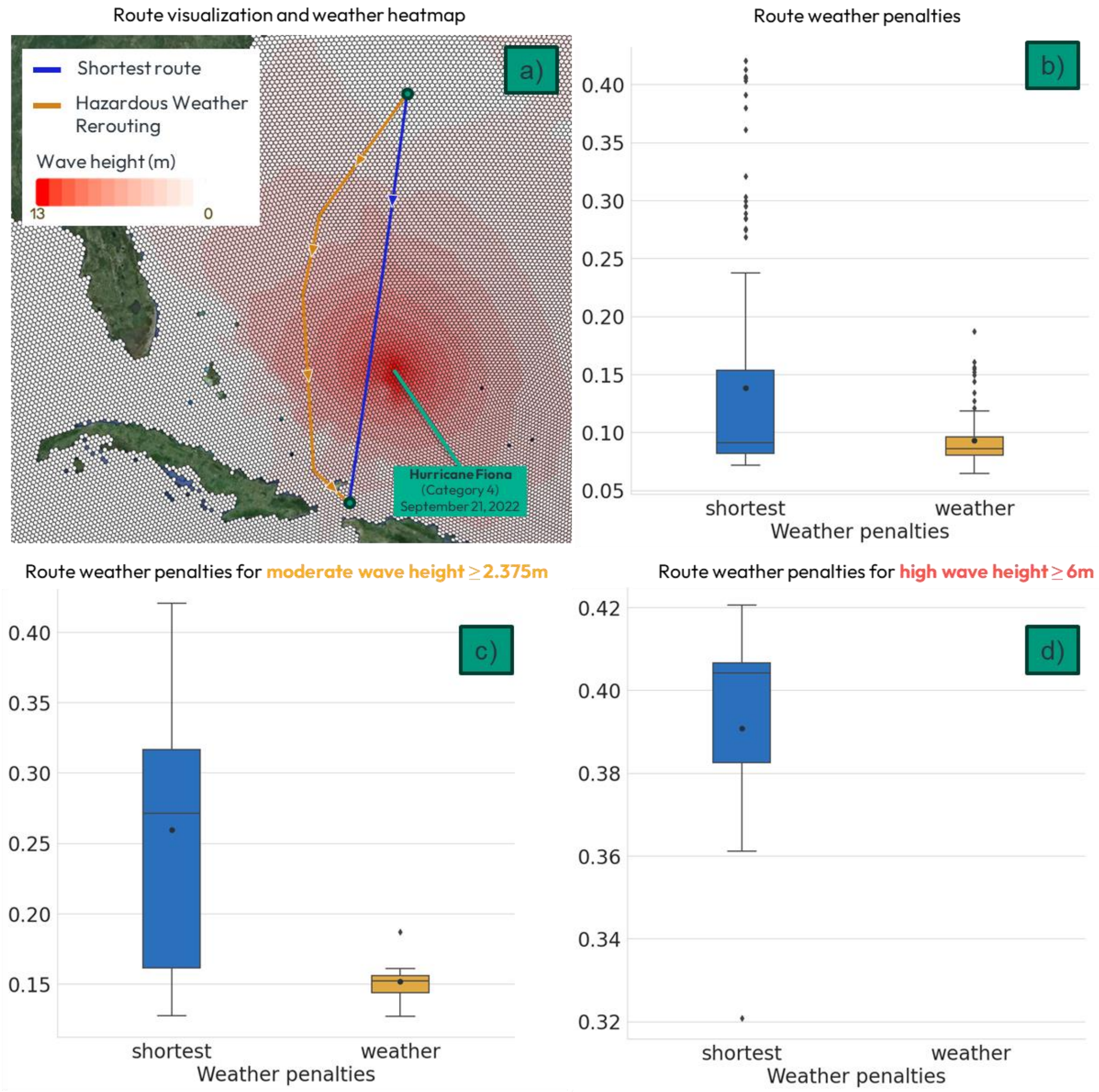


Figure 38: Vessel route evaluation along the sea area of Category 4 hurricane Fiona in September 2022. a) Route visualization and wave height heatmap along the vessel route from origin to destination. b) Weather penalties across both routes. c) Weather penalties for moderate and greater wave heights (≥ 2.375m). d) Weather penalties for high wave heights (≥ 6m).

The initial evaluation results showcase that the hazardous weather routing method successfully routes vessels through areas of calmer and preferred weather conditions compared to state of the art methods. More importantly the performance of the hazardous weather routing method significantly increases in the case of very extreme weather events. In subsequent steps the full implementation of the hazardous weather routing solution within the internal Kpler software stack and system architecture, as well as with the CREXDATA platform are planned.

In the context of the Maritime Use Case evaluation, two key scenarios will be tested:

1) Communicating with the vessel (sending and receiving data), while at sea though the IoT-Voyage Data Streamer (VDS);
2) Remotely commanding the vessel when hazardous situations have been forecasted (collision or weather). In such cases the vessel will be rerouted to safety.

Both scenarios will be tested in real-world conditions at sea in the dedicated sea trial testbed of the University of the Aegean in Syros, Greece, utilizing multiple vessels.

5.3.3 Rerouting Simulation System

The system consists of two components, a simulator and the collision avoidance service that communicate through a kafka cluster, which is a core architectural component of CREXDATA (Figure 39). This simulation setup is intended to ensure communication between components as well as measure message exchange times introduced into the system from transmitting information over the internet.

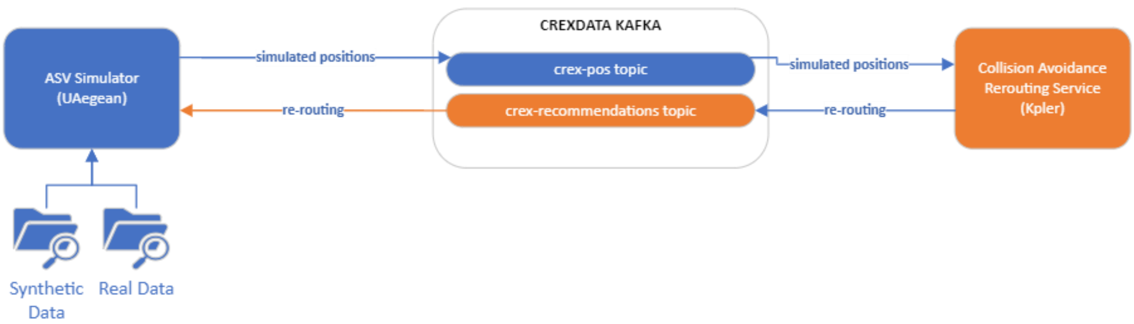


Figure 39: Rerouting Simulation System core components

On the one hand, the simulator component reads the vessel proximity dataset and its corresponding collision location from csv files, then it formats all positional records according to the service specification and publishes them to a kafka topic as a stream. The simulator component ensures that the order and transmission intervals of consecutive messages of the original dataset are preserved.

On the other hand, the collision avoidance service consumes the simulation topic, and sends back a new route to follow (re-routing) through a predefined “recommendation” topic.

The rerouting recommendations topic is consumed from the simulator component, completing a full communication cycle between the components. Both positional broadcasting and recommendation messages are broadcasted and received in real time .

5.4 Initial Use Case Evaluation

5.4.1 Involved Data

The resulting data set consists of positional and mobility data of 3 vessels during all 3 challenges. The data set has high temporal precision, with positions. recorded almost every second, resulting in over 6900 positional reports (Table 13).

Table 13: Initial use case evaluation dataset summary table

Feature	Description	Units
Identifier	Including the vessel's team name, the challenge in question and a unique increasing number	
Timestamp	Reported time of each positional message	UNIX epoch format in seconds
Coordinates	Reported longitude and latitude of each message	EPSG: 4326
Speed	Reported speed of the vehicle at each point	kilometers per hour (kph)
Heading	Reported heading of the vessel's bow	degrees (0-360)
	Challenge	**Positions (#)**
	Speed	987
	Avoidance	1871
	Endurance	4071
	Total	**6929**

Figure 40: Race data visualizations. From left to right: the speed, collision avoidance and the endurance race, with the pink trajectory is the University of Ioannina, blue the University of the Aegean, while red is the University of Porto.

5.4.2 Visualisations Evaluation

Computer-based visualisation systems provide visual representations of datasets designed to help people carry out tasks more effectively. Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

Vessel operators and teams require to perform a number of challenging tasks

- Investigate movement characteristics and sensor measurement recordings from a single boat in space and time. This includes detection of anomalies and unwanted behaviours, such as boat malfunctions or weather-related disruptions of its movement.
- Assess the degree of stability in performing repeated movements and/or operations.
- Detect and examine potential collision situations, in particular, during simultaneous movement of several vessels.

Overall, these tasks involve data exploration and analysis tasks, namely exploration of single trajectories, analysis of repeated parts of movement, and analysis of collective movement. Additional focus is on detecting and exploring events of close approach of vessels to other static or moving objects; such events as interactions. Within the first part of the evaluation a number of data visualisation were created to explore their ability to convey the important information to an operator. These included simple maps and more complex visualizations such as space-time cubes.

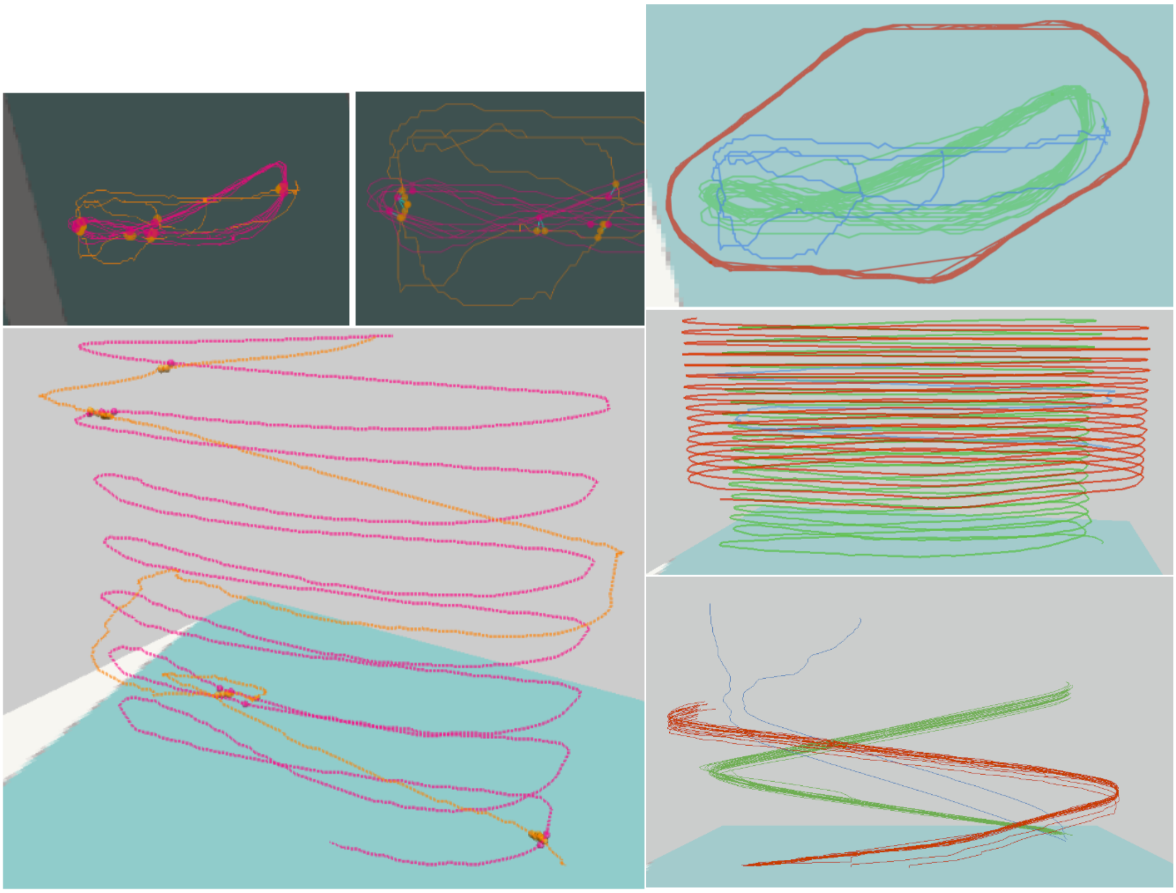


Figure 41: Advanced visual representations of the data from the endurance race. Images on the left focus on the Interactions between autonomous vessels presented on map (top) and space-time cube (bottom). Images on the right depict trajectory evolution during the race on map (top) and space-time cube (middle-bottom).

Operators and race team members were interviewed and involved in the follow up survey (see Section 9) to comment on the multiple alternative data visualisations.

5.5 Perspectives

Currently we have collected data and tested several of the algorithmic and system components of the CREXDATA system. A key activity of the project has been the Aegean Ro-Boat Race that gave us access to real world data originating from several heterogenous autonomous vessels and system platforms. Additionally, we were able to validate the requirements from D2.1 and test several visualization idioms to explore which were capable of assisting decision making. In the follow-up survey we collected responses from the team that took part in the race and analysed the result to draw some interesting conclusions.

Looking forward, as we progress into the second half of the project, the steps are clear and promising. During the summer of 2024, in the second edition of the Aegean Ro-Boat Race more complex test will be performed. Apart from achieving data communications to and from the vessel, commands will be pushed to the vessel allowing it to alter its behaviour. In 2025 during the final pilot demonstration even more complex behaviour will be tested with vessel showcasing advanced autonomous behaviours (received remotely).

Finally, FMI researchers have been in active collaboration with the Marine Use Case partners to find the optimal showcases for the Marine Use Case and give support in the analysis of the results. In addition to these, FMI will produce coastal maritime use case focusing on ‘on-demand’ high-resolution wave forecasts for the coastal fairways. High-resolution grids will be produced to selected coastal Finnish fairways, where accurate information on wave conditions has been identified as important for maritime safety. The event-based statistics method will be utilized to identify situations where conditions would require more detailed wave forecasts to ensure safety at sea. Based on this evaluation, the high-resolution forecast would be automatically launched by the forecast system. Reanalysis data and a-posteriori weather forecasts will be used to build and test the system.

6 Conclusions

This report has presented an interim evaluation of the use cases, pilots, demonstrators, and simulation models and tools developed within the CREXDATA project. The evaluation has been conducted using a common methodology, ensuring consistency across the different use cases.

The EmCase, Life Science Use Case, and Maritime Use Case have all been evaluated in terms of their fulfilment of requirements and usability. The initial evaluations have provided valuable insights into the effectiveness of the use cases and the potential areas for improvement.

The pilots and demonstrators developed for each use case have shown the practical applicability of our research. They have served as tangible proof of the project's impact, demonstrating the potential of our research in real-world settings.

The simulation models and tools developed for each use case have played a crucial role in testing our hypotheses and validating our solutions. They have allowed us to assess the performance of our solutions in a controlled environment, providing valuable feedback for further development.

The final scenario definitions for each use case have provided a clear and detailed description of the situations our project aims to address. These scenarios have guided our development work, ensuring that our solutions are tailored to the specific needs of each use case.

The initial conclusions drawn from the evaluation results across the use cases have provided valuable feedback for our future work. They have highlighted the strengths of our approach, as well as the areas where further development is needed.

In summary, the work presented in this report is a key milestone in the CREXDATA project, as it provides a comprehensive overview of our progress to date and sets the direction for our future work. As we move forward, we will continue to refine our solutions, guided by the feedback from our evaluations and the needs of our use cases.

7 Acronyms and Abbreviations

Each term should be bulleted with a definition.

Below is an initial list that should be adapted to the given deliverable.

- CA – Consortium Agreement
- D – deliverable
- DoA – Description of Action (Annex 1 of the Grant Agreement)
- EB – Executive Board
- EC – European Commission
- GA – General Assembly / Grant Agreement
- HPC – High Performance Computing
- IPR – Intellectual Property Right
- KPI – Key Performance Indicator
- M – Month
- MS – Milestones
- PM – Person month / Project manager
- WP – Work Package
- WPL – Work Package Leader

8 References

[1] J. Brooke, ‘SUS: A “Quick and Dirty” Usability Scale’, in *Usability Evaluation In Industry*, CRC Press, 1996.

[2] J. Pottebaum, C. Rechberger, M. Hieb, I. Gräßler, and C. Resch, ‘Extremwettersituationen in alpinen Gebieten: Management kritischer Situationen in Echtzeit bei extremen und komplexen Daten’, *Tagungsband Fachtag. Katastrophenforschung 2023*, 2023, Accessed: Jun. 03, 2024. [Online]. Available: https://ris.uni-paderborn.de/record/46988

[3] J. Pottebaum, M. Ebel, and I. Gräßler, ‘Uncovering Impact of Innovation: Continuous Stakeholder Engagement through Scenario-based Systems Engineering’, ISCRAM, May 24, 2024. doi: 10.5281/zenodo.11298333.

[4] M. Ponce-de-Leon *et al.*, ‘COVID19 Flow-Maps Daily Cases Reports’. Zenodo, Aug. 18, 2021. doi: 10.5281/zenodo.5217386.

[5] A. Arenas *et al.*, ‘Modeling the Spatiotemporal Epidemic Spreading of COVID-19 and the Impact of Mobility and Social Distancing Interventions’, *Phys. Rev. X*, vol. 10, no. 4, p. 041055, Dec. 2020, doi: 10.1103/PhysRevX.10.041055.

[6] V. Colizza, M. Barthélemy, A. Barrat, and A. Vespignani, ‘Epidemic modeling in complex realities’, *C. R. Biol.*, vol. 330, no. 4, pp. 364–374, 2007, doi: 10.1016/j.crvi.2007.02.014.

[7] D. Soriano-Paños, W. Cota, S. C. Ferreira, G. Ghoshal, A. Arenas, and J. Gómez-Gardeñes, ‘Modeling Communicable Diseases, Human Mobility, and Epidemics: A Review’, *Ann. Phys.*, vol. 534, no. 6, p. 2100482, 2022, doi: 10.1002/andp.202100482.

[8] C. Akasiadis *et al.*, ‘Parallel model exploration for tumor treatment simulations’, *Comput. Intell.*, vol. 38, no. 4, pp. 1–23, 2022, doi: 10.1111/coin.12515.

[9] M. Vázquez *et al.*, ‘Alya: Multiphysics engineering simulation toward exascale’, *J. Comput. Sci.*, vol. 14, pp. 15–27, May 2016, doi: 10.1016/j.jocs.2015.12.007.

[10] M. Ponce-de-Leon *et al.*, ‘PhysiBoSS 2.0: a sustainable integration of stochastic Boolean and agent-based modelling frameworks’, *Npj Syst. Biol. Appl.*, vol. 9, no. 1, Art. no. 1, Oct. 2023, doi: 10.1038/s41540-023-00314-4.

[11] J. Ozik *et al.*, ‘High-throughput cancer hypothesis testing with an integrated PhysiCell-EMEWS workflow’, *BMC Bioinformatics*, vol. 19, no. 18, p. 483, Dec. 2018, doi: 10.1186/s12859-018-2510-x.

[12] G. Grigoropoulos *et al.*, ‘A Scalable System for Maritime Route and Event Forecasting’. OpenProceedings.org, 2024. doi: 10.48786/EDBT.2024.66.

9 Appendix 1: Expert users' questionnaire.

9.1 Weather Emergencies Use Case (EmCase):

1. **Generic user background information**
    1.1. What is your job title?
    1.2. In which sector do you work in?
    1.3. How many years of experience do you have in the sector?
    1.4. What is your educational background (e.g. university degree major, apprenticeship)?
    1.5. What are your main job tasks?
    1.6. To whom are you responsible for performing these tasks?

2. **Specific background information**
A set of questions were added per technology to be investigated in the evaluation.

    2.1. Text mining
        2.1.1. Have you used or do you use social media analysis software for mission-related purposes?
        Never, Rarely, Sometimes, Often, Always
            2.1.1.1. If you have, for what purpose? Examples: to gain insights into the mission area and about observable dangers, to search for specific objects or persons, to see what you can only assume from communication, for documentation.
            2.1.1.2. If you have, which ones?
        2.1.2. How would you describe your experience with map-based incident management systems? What works well, what doesn't?

    2.2. ARGOS
        2.2.1. Have you used or do you use a map-based incident management system including weather-related data like ARGOS for mission-related purposes?
        Never, Rarely, Sometimes, Often, Always
            2.2.1.1. If you have, which ones?
        2.2.2. How would you describe your experience with map-based incident management systems? What works well, what doesn't?

    2.3. Complex Event Forecasting
        2.3.1. Have you used or do you use some sort of machine learning or AI system for mission related purposes?
        Never, Rarely, Sometimes, Often, Always
        2.3.2. If you have, for what purpose and what kind of system do/did you use?

2.4. NeRF

2.4.1. Have you used or do you use drone imagery (i.e., photos, videos) for mission-related purposes? Note: It does not matter if you have flown a drone yourself, only if you have used the gathered photos or videos for any mission-related purposes.
Never, Rarely, Sometimes, Often, Always

2.4.1.1. If you have, for what purpose? Examples: to gain a general overview of the mission area, to search for specific objects or persons, for documentation.

2.4.2. Have you used or do you use photogrammetry tools (e.g., WebODM, DJI Terra, Pix4D, Agisoft Metashape) to create 3D models from drone imagery?
Never, Rarely, Sometimes, Often, Always

2.4.2.1. If you have, which ones?
ebODM, Terra, Metashape, Pix4D, Other

2.4.3. Have you used or do you use 3D models from photogrammetry tools for mission-related purposes?
Never, Rarely, Sometimes, Often, Always

2.4.3.1. If you have used photogrammetry tools and/or the created 3D models, for what purposes?

2.4.4. How would you describe your experience with map-based incident management systems? What works well, what doesn't?

2.4.5. Have you heard of "Neural Radiance Fields (NeRFs)" or "Gaussian splatting" before today? Yes/No

2.5. Augmented Reality

2.5.1. Have you used or do you use Augmented Reality applications for mission-related purposes?
Never, Rarely, Sometimes, Often, Always

2.5.1.1. If you have, for what purpose? Examples: to visualize the mission area, to indicate points of interests, to view simulation results, to prepare for recovery after an incident.

2.5.1.2. If you have, which ones?

2.5.2. How would you describe your experience with map-based incident management systems? What works well, what doesn't?

2.5.3. Have you heard of Augmented Reality applications for emergency management before today? Yes/No

1. Please rate these objectives of the CREXDATA project according to your background and present and future needs (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful):

1.1. Being able to have extreme-scale data ingestion/generation, fusion and exploitation.

1	2	3	4	5

1.1.1. Ingesting multimodal data (images, simulations, social media publications, etc).

1	2	3	4	5

1.1.2. Using dynamic modelling to predict the systems' behaviour.

1	2	3	4	5

1.1.3. Handling multilingual social data in real-time.

1	2	3	4	5

1.2. Having real-time predictive knowledge and forecasts.

1	2	3	4	5

1.2.1. Using online federated learning.

1	2	3	4	5

1.2.2. Having multiresolution complex event forecasting under uncertainty.

1	2	3	4	5

1.2.3. Using optimization techniques for Prediction-as-a-Service (PaaS).

1	2	3	4	5

1.3. Reducing the perceived complexity.

1	2	3	4	5

1.3.1. Using graphical workflow design.

1	2	3	4	5

1.3.2. Using visual analytics coupled with XAI for understanding complexity and reasoning under uncertainty.

1	2	3	4	5

1.3.3. Using augmented reality under uncertainty on-site & remotely.

1	2	3	4	5

4. Specific aspects of technologies in the EmCase

4.1. Text Mining
You were shown the text mining functionality of the CREXDATA system. It is intended to indicate relevant social media postings especially with a focus on highlight relevant media that helps to get visual insights into the situation.

4.1.1. Please rate the following statements on a scale from 1 (strongly disagree) to 5 (strongly agree).
4.1.1.1. The tool is suited to inform mission-relevant decisions.
4.1.1.2. The tool is helpful to gather relevant textual statements.
4.1.1.3. The tool is helpful for indicating relevant information in postings.
4.1.1.4. I would use the tool, in its current state, for future missions.
4.1.2. What did you especially like about the tool?
4.1.3. What did you dislike?
4.1.4. In which types of scenarios (flooding, wildfire, earthquake, ...) and for what decisions would this tool be especially advantageous?
4.1.5. Feature requests
4.1.5.1. Are there other improvements you would like to see?
4.1.6. Is there anything else you would like to tell us?

4.2. ARGOS
You were shown the ARGOS system with its wide variety of information layers, covering data streams from sensors like weather stations and traffic cameras, but especially also satellite images and all types of weather nowcasts and forecasts. You also saw the opportunity to incorporate further data sources like emergency calls or social media postings.
4.2.1. Please rate the following statements on a scale from 1 (strongly disagree) to 5 (strongly agree).
4.2.1.1. The tool is suited to inform mission-relevant decisions.
4.2.1.2. The tool is helpful for presenting all types of information that make up a situational picture (German: Lagebild).
4.2.1.3. The categorisation of threshold values for different vulnerable elements, is helpful to take measures correctly and in time.

4.2.1.4. The possibility of alerting the responsible persons is helpful to take measures correctly and in time
4.2.1.5. A graphical illustration of real-time flood simulations be helpful for decision-making?
4.2.1.6. I would use the tool, in its current state, for future missions.

4.2.2. What did you especially like about the tool?
4.2.3. What did you dislike?
4.2.4. In which types of scenarios (flooding, wildfire, earthquake, ...) and for what decisions would this tool be especially advantageous?
4.2.5. Feature requests: Please consider the following ideas for improving the helpfulness of such tools in your line of work. Rate the desirability of each idea on a scale from 1 (not interesting) to 5 (very interesting).

4.2.5.1. Incorporate simulation results
4.2.5.2. Visualize locations and status of own resources (like fire engines, ambulances etc.)
4.2.5.3. Add communication functionality

4.2.6. Are there other improvements you would like to see?
4.2.7. Is there anything else you would like to tell us?

4.3. Complex Event Forecasting

You were shown the first results of the event forecasting technology. Based on precipitation and monitoring data in the sewer network, critical situations / events (i.e. water levels in the sewer exceeding a certain threshold value) were identified for Innsbruck within the years 2021-2023. Based on this approach, the technology can predict future events during real-time data transmission.

4.3.1. Please rate the following statements on a scale from 1 (strongly disagree) to 5 (strongly agree).

4.3.1.1. The prediction of critical events, i.e. water level in the sewer exceeds a certain threshold value, is suited to inform mission-relevant decisions.
4.3.1.2. The technology is helpful to make decisions such as opening and closing parts of the sewer network or switching on flood pumping stations.
4.3.1.3. The threshold values well chosen.
4.3.1.4. A gradual increase of threshold values, i.e. categorisation into threshold classes, is helpful.

4.3.2. What did you especially like about the tool?
4.3.3. What did you dislike?
4.3.4. Are there other improvements you would like to see?
4.3.5. Is there anything else you would like to tell us?

4.4. NeRF

You were shown 3D models generated by different photogrammetry tools on three datasets, each with a different number of images (small, medium and large set).

4.4.1.1. Please rate the following statements on a scale from 1 (strongly disagree) to 5 (strongly agree).

4.4.1.2. The model is suited to inform mission-relevant decisions.

4.4.1.3. The model is helpful for identifying relevant structures, such as building entrances (doors and windows), road blockages, or hazardous objects (e.g., power lines).

4.4.1.4. In an “average” flooding scenario, the required time between takeoff and model availability would be adequate.

4.4.1.5. This time would also be adequate in an “average” wildfire scenario.

4.4.1.6. I would use the tool, in its current state, for future missions.

4.4.2. What did you especially like about the tool?

4.4.3. What did you dislike?

4.4.4. In which types of scenarios (flooding, wildfire, earthquake, ...) and for what decisions would a 3D model be especially advantageous? When is an orthophoto (2D top-down image), or a single photo taken from a large height, sufficient?

4.4.5. In different scenarios, what is more important: a) Time from drone start to model availability, or b) model quality?
If possible, please provide concrete limits. E.g., “during a flood, the model must be good enough to identify text on street signs”, or “during a wildfire, I need to have a usable model no later than 15 minutes after the drone has landed”.

4.4.6. For areas which the drone has not (or only briefly) seen during the survey flight, which is better: 1) A gap in the 3D model (as seen in WebODM and other “classical” photogrammetry tools), or 2) a guess based on the surroundings (NeRFs and other AI-based approaches)?

4.4.6.1. Does it depend on the scenario, or does it not matter at all?

4.4.7. Feature requests: Please consider the following ideas for improving the helpfulness of photogrammetry tools in your line of work. Rate the desirability of each idea on a scale from 1 (not interesting) to 5 (very interesting).

4.4.7.1. Generate the model “live”. I.e., the drone transmits each image right after it’s taken and the model is updated with each new image, gradually expanding and improving over time.

4.4.7.2. Combine multiple small 3D models from different sections of the incident site into one global model.

4.4.7.3. Visualize the mission’s progression by creating and comparing models of the same area(s) at different points in time.

4.4.7.4. Visualize additional data as overlays to the 3D model, such as temperature measurements.
4.4.7.5. Automatically identify and highlight possibly interesting objects, such as persons or doors, ...
4.4.7.6. Automatically identify and remove dynamic objects (e.g., people or cars) from the model.
4.4.7.7. Create predicted hypothetical models based on text input, e.g., “increase the water level by 2 meters”.
4.4.7.8. Reduce the required time for model generation.
4.4.7.9. Reduce the CPU and memory requirements.
4.4.7.10. Improve the “responsiveness” of the visualization. For example, by reducing the time it takes to render a new image when changing the desired perspective.

4.4.8. Are there other improvements you would like to see?
4.4.9. Is there anything else you would like to tell us?

4.5. Augmented Reality

You were shown the Augmented Reality application. It is an alternative frontend to visualize data processing results, offering situational awareness close to the incident scene. The application implements various functionalities like map view, highlighting of points of interests (e.g., critical nodes within a sewer network), fatigue status of first responders based on biometric data and flooding simulations.

4.5.1. Please rate the following statements on a scale from 1 (strongly disagree) to 5 (strongly agree).
4.5.1.1. The illustration of the flooding level is suited to inform mission-relevant decisions.
4.5.1.2. The illustration of vulnerable elements is suited to inform mission-relevant decisions.
4.5.1.3. The illustration of manholes is suited to inform mission-relevant decisions.
4.5.1.4. The illustration of evacuation routes / routing is suited to inform mission-relevant decisions.
4.5.1.5. The tool is helpful for routing of forces
4.5.1.6. The tool is helpful to advice people in danger.
4.5.1.7. The tool is helpful for orientation on site.
4.5.1.8. The tool is helpful for indicating critical points and areas beyond the typical map-based visualisation.
4.5.1.9. I would use the tool, in its current state, for future missions.

4.5.2. What did you especially like about the tool?
4.5.3. What did you dislike?
4.5.4. In which types of scenarios (flooding, wildfire, earthquake, ...) and for what decisions would this tool be especially advantageous?
4.5.5. Feature requests: Please consider the following ideas for improving the helpfulness of such tools in your line of work. Rate the

desirability of each idea on a scale from 1 (not interesting) to 5 (very interesting).

- 4.5.5.1. Incorporate simulation results
- 4.5.5.2. Visualize locations and status of own resources (like fire engines, ambulances etc.)
- 4.5.5.3. Add communication functionality
- 4.5.5.4. Detail biometric data
- 4.5.5.5. Incorporate object plans and technical information about buildings, infrastructure etc.
- 4.5.5.6. Incorporate weather-related information like wind direction/speed, rain gauges and forecasts etc.
- 4.5.5.7. Add functionality to monitor and control robots

4.5.6. Are there other improvements you would like to see?

4.5.7. Is there anything else you would like to tell us?

5. System Usability Scale (SUS) questions

The System Usability Scale (SUS) [1] is a widely used and validated approach to usability assessment. SUS is easy to administer, easy to analyse and quick. The questionnaire results are analysed to obtain an aggregated score for the usability of a product. SUS is a questionnaire of ten items to which participants need to answer using a five-point Likert scale with verbal anchors at the extremes. Following [2] the answers are transformed and aggregated to obtain a score between 0-100, where 0 is the minimum usability score possible for the system and scores above 68 are considered above average [2].

To calculate the SUS score you should follow these steps:

- Calculate the single item score contribution which will range between 0 and 4:
  - for items 1,3,5,7 and 9 the score contribution is the position marked on the scale by the expert minus 1;
  - for items 2,4,6,8 and 10, the contribution is 5 minus the position marked on the scale by the expert;
- Sum the score contributions
- Multiply by 2.5 the sum of the scores (point 2)

5.1. SUS questions (1= Strongly disagree; 5 Strongly Agree):

1. I think that I would like to use this system frequently

1	2	3	4	5

2. I found the system unnecessarily complex.

1	2	3	4	5

3. I thought the system was easy to use.

1	2	3	4	5

4. I think that I would need the support of a technical person to be able to use this system.

1	2	3	4	5

5. I found the various functions in this system were well integrated.

1	2	3	4	5

6. I thought there was too much inconsistency in this system.

1	2	3	4	5

7. I would imagine that most people would learn to use this system very quickly.

1	2	3	4	5

8. I found the system very cumbersome to use.

1	2	3	4	5

9. I felt very confident using the system.

1	2	3	4	5

10. I needed to learn a lot of things before I could get going with this system.

1	2	3	4	5

6. Further comments from expert user

Thank you for your expertice and participation in the CREXDATA EmCase Field Trials!

9.2 Life Sciences Use Case:

Test user evaluation questionnaire - CREXDATA project

Questionnaire

2. **User background information**
   - 2.1. What is your job title?
   - 2.2. Do you work in academia or in industry?
   - 2.3. Years of experience
   - 2.4. Background studies (university degree major, etc.)
   - 2.5. What are your main job tasks?
   - 2.6. To whom are you responsible for performing these tasks?

3. **Modelling background of the user**
   - 3.1. Are you a model developer or do you use models already developed?
   - 3.2. How relevant is in your research the use of models and simulation for addressing questions (low / medium / very high)
   - 3.3. What kind of models and simulations do you use to analyse systems?
   - 3.4. How much would you regard yourself as a model developer?
   - 3.5. Which modelling and simulation approaches (e.g. ODEs, Stochastic, ABMs) and tools (e.g. programming/modelling languages, libraries, frameworks) do you use in your research?
   - 3.6. Do you work with real-time data? If not, would you like to work with this kind of data?
     - 3.6.1. Are you tools and workflows able to work with such data?
     - 3.6.2. Would you be interested in using a data processing workflow that would allow using real-time data?
     - 3.6.3. Would using this data allow you to address different problems than the ones you are currently addressing?
   - 3.7. Do you currently use forecasting techniques? Are there specific events that you would like to forecast in real-time, which you currently cannot forecast?

The vision of CREXDATA is to develop a generic platform for real-time critical situation management including flexible action planning and agile decision making over data of extreme scale and complexity. CREXDATA develops the algorithmic apparatus, software architectures and tools for federated predictive analytics and forecasting under uncertainty. The envisioned framework boosts proactive decision making providing highly accurate and *transparent short- and long-term forecasts to end-users, explainable via advanced visual analytics and accurate, real-time, off and on-site augmented reality facilities.*

4. **Please rate these objectives of the CREXDATA project according to your background and present and future needs (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful):**
    4.1. Being able to have extreme-scale data ingestion/generation, fusion and exploitation.
        4.1.1. Ingesting multimodal data (images, simulations, social media publications, etc).
        4.1.2. Using dynamic modelling to predict the systems' behaviour.
        4.1.3. Handling multilingual social data in real-time.
    4.2. Having real-time predictive knowledge and forecasts.
        4.2.1. Using online federated learning.
        4.2.2. Having multiresolution complex event forecasting under uncertainty.
        4.2.3. Using optimization techniques for Prediction-as-a-Service (PaaS).
    4.3. Reducing the perceived complexity.
        4.3.1. Using graphical workflow design.
        4.3.2. Using visual analytics coupled with XAI for understanding complexity and reasoning under uncertainty.
        4.3.3. Using augmented reality under uncertainty on-site & remotely.

6. **Specific aspects of the health emergency use case**
    6.1. Rate the following components of the health emergency use case according to your interest on the foreseen results (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful)
        i. Using a graphical user interphase
        ii. Being able to have parameter calibration
        iii. Using early time series characterisation
        iv. Online model exploration
        v. Real-time/online forecasting simulation trajectories
    6.2. Rate the following Key Performance Indicators according to your interest on the foreseen results. (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful)
        i. KPI 1: Forecasting 7 parameter sets that reduce the COVID infection of the simulation (multiscale infection or epidemiologic scenario).
        ii. KPI 2: Use the runtime adaptation of simulation trajectories to improve the outcomes of 5 scenarios or patients of the multiscale infection scenario.
        iii. KPI 3: Calibration of the epidemiological parameter to fit incidence time series.
        iv. KPI 4: Characterizing the space of parameters with 50% fewer simulations.

7. System Usability Scale questions

The System Usability Scale (SUS) [1] is a widely used and validated approach to usability assessment. SUS is easy to administer, easy to analyse and quick. The questionnaire results are analysed to obtain an aggregated score for the usability of a product. SUS is a questionnaire of ten items to which participants need to answer using a five-point Likert scale with verbal anchors at the extremes. Following [2] the answers are transformed and aggregated to obtain a score between 0-100, where 0 is the minimum usability score possible for the system and scores above 68 are considered above average [2].

To calculate the SUS score you should follow these steps:

- *Calculate the single item score contribution which will range between 0 and 4:*
  - *for items 1,3,5,7 and 9 the score contribution is the position marked on the scale by the expert minus 1;*
  - *for items 2,4,6,8 and 10, the contribution is 5 minus the position marked on the scale by the expert;*
- *Sum the score contributions*
- *Multiply by 2.5 the sum of the scores (point 2)*

7.1. SUS questions (1= Strongly disagree; 5 Strongly Agree):

1. I think that I would like to use this system frequently

2. I found the system unnecessarily complex

3. I thought the system was easy to use

4. I think that I would need the support of a technical person to be able to use this system

5. I found the various functions in this system were well integrated

6. I thought there was too much inconsistency in this system

7. I would imagine that most people would learn to use this system very quickly

8. I found the system very cumbersome to use

9. I felt very confident using the system

10. I needed to learn a lot of things before I could get going with this system

7. Further comments from expert user

REFERENCES

[1] Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

[2] Sauro, J. (2011). SUStisfied? Little-known System Usability Scale facts. User Experience: The Magazine of the User Experience Professionals Association, 10(3)

9.1 Maritime Use Case:

The questionnaire of the maritime use case has been distributed to stakeholders using the EU Survey platform. The online questionnaire is available at : https://ec.europa.eu/eusurvey/runner/62f59865-6fe9-7a5d-7248-249333d32d3c.

CREXDATA - Maritime Use Case

The vision of CREXDATA is to develop a generic platform for real-time critical situation management including flexible action planning and agile decision making over data of extreme scale and complexity. CREXDATA develops the algorithmic apparatus, software architectures and tools for federated predictive analytics and forecasting under uncertainty. The envisioned framework boosts proactive decision making providing highly accurate and transparent short- and long-term forecasts to end-users, explainable via advanced visual analytics and accurate, real-time, off and on-site augmented reality facilities.

User Background Information

What is your Job Title

* Do you work in academia or in industry?

- ○ Industry
- ○ Academia

* Years of experience

Only values between 1 and 15 are allowed

* Background studies (university degree major, etc.)

- ☐ University degree
- ☐ Master's
- ☐ PhD

* What are your main job tasks?

* To whom are you responsible for performing these tasks?

Did you participate in the 2023 Aegean Ro-Boat Race?

- ○ Yes
- ○ No

Figure 42: Maritime Use case questions : User Background Information.

Modeling/Data Analysis Background

* Are you a data scientist (model developer or similar) or do you use models already developed?

- ○ I build new models /systems myself
- ○ No / Not sure

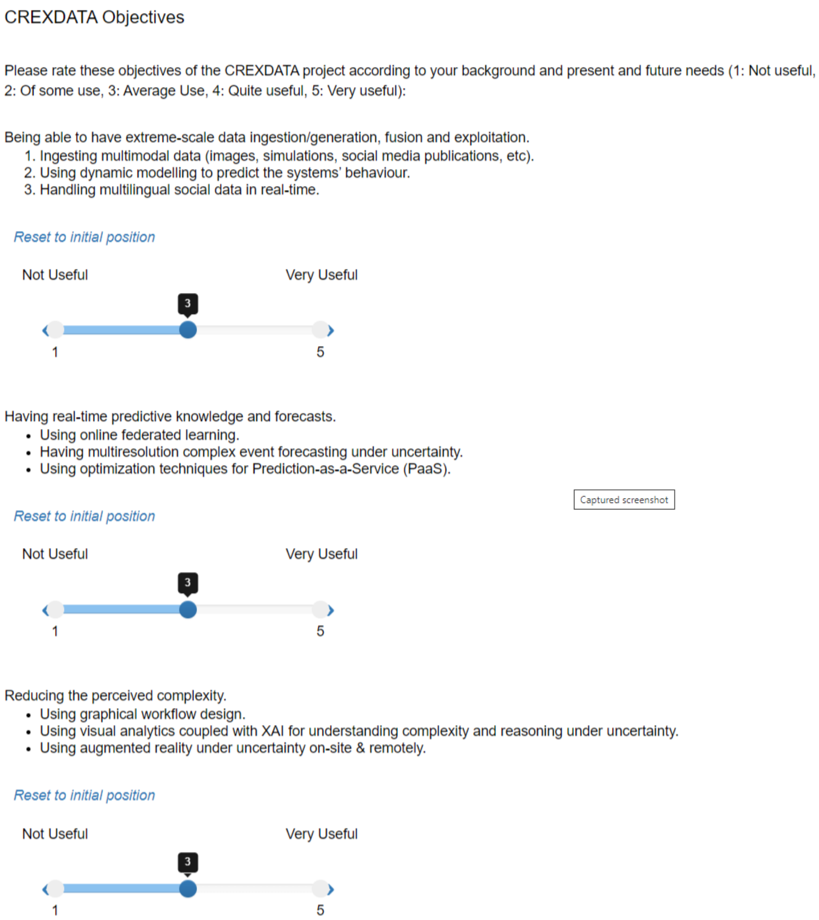
Do you try to forecast things into the future as part of your job? e.g forecast future vessel positions, or weather.
Do you currently use forecasting techniques?

Captured screenshot

- ○ Yes
- ○ No

Please select **any** of these data analysis tasks that are important for completing your tasks.

- ☐ Investigate movement characteristics and sensor measurement recordings from a single boat in space and time. This includes detection of anomalies and unwanted behaviours, such as boat malfunctions or weather-related disruptions of its movement.
- ☐ Assess the degree of stability in performing repeated movements and/or operations.
- ☐ Detect and examine potential collision situations, in particular, during simultaneous movement of several vessels.
- ☐ Forecast potential dangerous situations so that you gain manual control of the vessel



CREXDATA Objectives

Please rate these objectives of the CREXDATA project according to your background and present and future needs (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful):

Being able to have extreme-scale data ingestion/generation, fusion and exploitation.

1. Ingesting multimodal data (images, simulations, social media publications, etc).
2. Using dynamic modelling to predict the systems' behaviour.
3. Handling multilingual social data in real-time.

Reset to initial position

Not Useful Very Useful

3

1 5

Having real-time predictive knowledge and forecasts.

- Using online federated learning.
- Having multiresolution complex event forecasting under uncertainty.
- Using optimization techniques for Prediction-as-a-Service (PaaS).

Captured screenshot

Reset to initial position

Not Useful Very Useful

3

1 5

Reducing the perceived complexity.

- Using graphical workflow design.
- Using visual analytics coupled with XAI for understanding complexity and reasoning under uncertainty.
- Using augmented reality under uncertainty on-site & remotely.

Reset to initial position

Not Useful Very Useful

3

1 5

Figure 43: Maritime use case questions: Modelling background and CREXDATA objectives

Collision avoidance service

Example

The pilot of vessel A is alerted of a possible collision event with vessel B, as their current routes will intersect. The pilot evaluates the emergency of the forecasted event and is provided with a set of alternative routes for vessel A to follow in order to avoid colliding with vessel B. Based on the experience of the vessel pilot, he/she may opt to accept the proposed route, correct the proposed suggestion or follow an entirely different route in order to avoid collision with vessel B

How useful is a collision avoidance rerouting service for your day-to-day work?
Reset to initial position

Not Useful — Very Useful

3

1 — 5

How important is a collision avoidance and rerouting service to provide COLREG compliant recommendations?
Reset to initial position

Not Important — Very Important

3

1 — 5

* How often should a collision avoidance service provide re-routing recommendations?

- ○ sub-second
- ○ every 1-minute
- ○ every 5-minutes
- ○ every 30 minutes

How important is the existence of a graphical interface for a collision avoidance rerouting service?
Reset to initial position

Not Important — Very Important

3

1 — 5

* Would the existence of an Augmented Reality interface improve usability of the collision avoidance rerouting service?

- ○ No improvement
- ○ Minor improvement
- ○ Major improvement

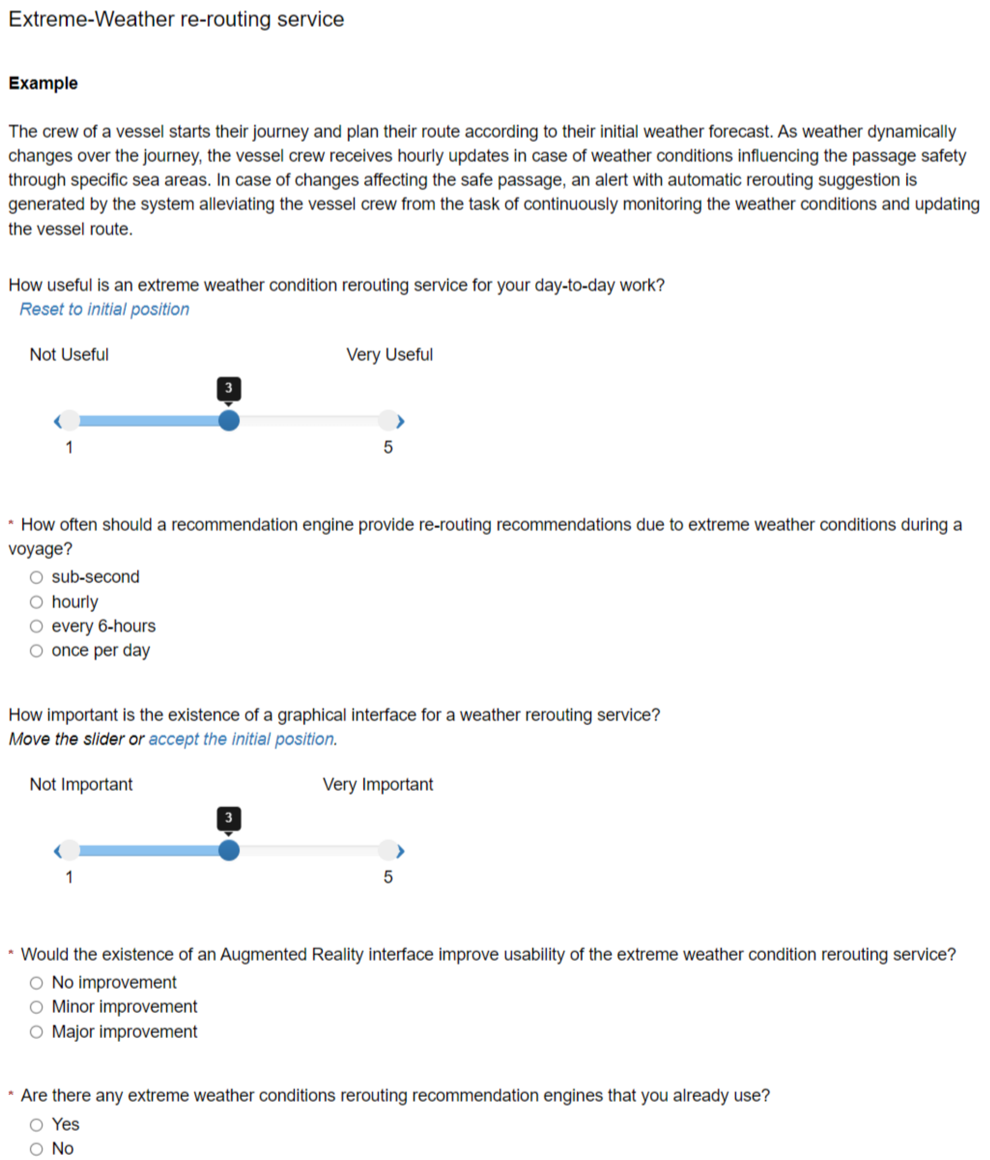
* Are there any collision avoidance rerouting recommendation engines that you already use?

- ○ Yes
- ○ No

In a crossing situation that your vessel should give priority to the other vessel. Which recommendations you would consider safe to follow.

- ☐ Pass from stern side, keeping at least 10m distance
- ☐ Pass from stern side
- ☐ Pass from stern side, keeping at least one length distance
- ☐ Pass from stern side, keeping at least 1m distance
- ☐ Pass from stern side, keeping at least two lengths distance
- ☐ Pass from bow side, keeping at least two lengths distance
- ☐ Pass from bow side

Figure 44: Maritime use case questions: Collision avoidance service



Extreme-Weather re-routing service

Example

The crew of a vessel starts their journey and plan their route according to their initial weather forecast. As weather dynamically changes over the journey, the vessel crew receives hourly updates in case of weather conditions influencing the passage safety through specific sea areas. In case of changes affecting the safe passage, an alert with automatic rerouting suggestion is generated by the system alleviating the vessel crew from the task of continuously monitoring the weather conditions and updating the vessel route.

How useful is an extreme weather condition rerouting service for your day-to-day work?
Reset to initial position

Not Useful Very Useful
3
1 5

* How often should a recommendation engine provide re-routing recommendations due to extreme weather conditions during a voyage?

- sub-second
- hourly
- every 6-hours
- once per day

How important is the existence of a graphical interface for a weather rerouting service?
Move the slider or accept the initial position.

Not Important Very Important
3
1 5

* Would the existence of an Augmented Reality interface improve usability of the extreme weather condition rerouting service?

- No improvement
- Minor improvement
- Major improvement

* Are there any extreme weather conditions rerouting recommendation engines that you already use?

- Yes
- No

Figure 45: Maritime use case questions: Extreme-Weather re-routing service

System Usability/Scale questions

A single trajectory represented on a map (left) and in a space-time cube (right). The time axis in the space-time cube is oriented upwards. Which is more useful for you (in real time)?

- ○ Map projection (Left)
- ○ Space-time cube (Right)

Regarding the previous, if analysis was performed at a later date on historical data, would your selection change?

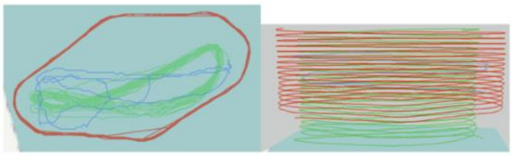
- ○ Yes
- ○ No

Captured screenshot

Both these data visualisation show the data from one of the races for three competing boats. In the left the visualisation involves only spatial data in a 2D map while the right is temporospatial (space-time cube)

Which one is more useful for you in real time decision making?

- ○ Map Projection (left)
- ○ Space time cube (right)



Similarly which is easier for you to spot the interactions between two vessels?

- ○ Map projection (left)
- ○ Space time cube (right)

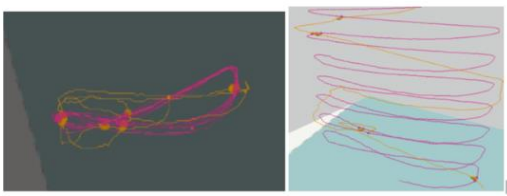


Figure 46: Maritime use case questions: system usability/scale questions

10 Appendix 2: Questionnaires filled by the expert users.

10.1 Weather Emergencies Use Case

For the EmCase, the survey was structured according to the time schedule presented in 3.5.1. As soon as they entered their first session either in Innsbruck or later Dortmund, each participants answered the individual background questionnaire. In each session, the corresponding "specific background" questionnaire was filled at the start and the "specific aspects of technologies" questionnaire was completed at the end. For AR, four participants in Dortmund conducted a technology-specific SUS survey. After all sessions, all participants responded to an overall SUS survey. Therefore, from 14 experts in total we have gathered 10 SUS for the overall assessment, plus 23 filled questionnaires for the specific questions on technologies. These filled questionnaires are available upon request.

User background information

- 1.1. What is your job title?
  - 1.1.1. senior specialist
  - 1.1.2. Employee in business intelligence and quality management, former emergency dispatcher*
  - 1.1.3. senior engineer
  - 1.1.4. environmental engineer
  - 1.1.5. Strategy and science coordination
  - 1.1.6. Chief fire officer
  - 1.1.7. GIS Expert
  - 1.1.8. Aviation engineer (AUS, Flight control)
  - 1.1.9. Research associate
  - 1.1.10. Incident commander, press spokesman
  - 1.1.11. Research assistant
  - 1.1.12. Incident commander
  - 1.1.13. Chief fire inspector
  - 1.1.14. Fire councillor
- 1.2. In which sector do you work in?
  - 1.2.1. Governmental Organization
  - 1.2.2. other
  - 1.2.3. other
  - 1.2.4. other
  - 1.2.5. other
  - 1.2.6. governmental Organization
  - 1.2.7. Governmental Organization
  - 1.2.8. Academia, Government Organization
  - 1.2.9. Governmental Organization
  - 1.2.10. Governmental Organization
  - 1.2.11. Governmental institution

1.2.12. Governmental organization
1.2.13. Governmental organization
1.2.14. Governmental organization

1.3. How many years of experience do you have in the sector?
1.3.1. 10
1.3.2. 3, 22*
1.3.3. 32
1.3.4. 10
1.3.5. 30
1.3.6. 8
1.3.7. 3
1.3.8. 6
1.3.9. 3
1.3.10. 19
1.3.11. 12
1.3.12. 21
1.3.13. 3
1.3.14. 34

1.4. What is your educational background (e.g. university degree major, apprenticeship)?
1.4.1. Master of Philosophy (Geography), Bachelor degree in geography, Phd student (ongoing, political science)
1.4.2. unfinished university degree in geodesy
1.4.3. Master 2 in Crisis Management
1.4.4. M.Sc.
1.4.5. university degree in computer science
1.4.6. Masters degree
1.4.7. M.S.c. Geography
1.4.8. University(ongoing)
1.4.9. M.S.c. Industrial Engineering
1.4.10. Abitur
1.4.11. Diplom
1.4.12. Abitur
1.4.13. Physics M.Sc.
1.4.14. Technical baccalaureate, craft apprenticeship

1.5. What are your main job tasks?
1.5.1. CREXDATA & another project CLIMAXX
1.5.2. Data modelling
1.5.3. Network of french firefighters
1.5.4. hydronamic sewer network calculations
1.5.5. evaluation of measurement data in the sewer network "
1.5.6. Combining science with firefighting
1.5.7. Leading a fire brigade; HR, finance, incidents, training, planning of resources
1.5.8. GIS, drone scenarios, disaster protection

1.5.9. UAS-research operating UAS
1.5.10. identifying potential improvements in the field of public safety. Derive technical specifications for future developments. Document and analyse evaluations of developments
1.5.11. press spokesman of the fire department
1.5.12. Research, special advisor, extinguishing fire
1.5.13. Calls of fire department
1.5.14. Disaster prevention planning (networking of various departments and offices in the city) including risk analyses
1.5.15. Working with geoinformation"
1.5.16. Officer in civil protection and commander in Fire Dept. B-level

1.6. To whom are you responsible for performing these tasks?
1.6.1. The ministry of the interior
1.6.2. CEO
1.6.3. civil defense director
1.6.4. Head of Department
1.6.5. President of firebrigade(ÖBFV)
1.6.6. Government of city of Insbruck
1.6.7. Other technical experts like Geologists or foresters
1.6.8. Senior researcher/Department head
1.6.9. Director of Fire Department of Dortmund
1.6.10. The chief of the fire department and the press spokesman of the city
1.6.11. UBA, Diplom chemist for fire department
1.6.12. Firefighter
1.6.13. Population protection team leader, fire councillor
1.6.14. Teamleader Civil protection

10.2 Life Sciences Use Case:

1. User background information

1.1. What is your job title?
1.1.1. Full Professor
1.1.2. Head of genome informatics
1.1.3. Professor
1.1.4. Sub-Director for Surveillance and Emergency Response of Public Health

1.2. Do you work in academia or in industry?
1.2.1. Academia
1.2.2. Academia
1.2.3. Academia
1.2.4. Public Administration

1.3. Years of experience
1.3.1. 15
1.3.2. 20
1.3.3. 22

1.3.4. 21

1.4. Background studies (university degree major, etc.)

1.4.1. PhD

1.4.2. PhD Computer Science

1.4.3. PhD Theoretical Physics

1.4.4. University degree Major (MD)

1.5. What are your main job tasks?

1.5.1. Lecturing and Research

1.5.2. Developing workflow for genomics and IA

1.5.3. Research director

1.5.4. Develop the main strategic lines aimed at the surveillance, prevention, and control of infectious diseases. Coordinate the surveillance team and the territorial units to ensure accurate collection of epidemiological information on outbreaks, notifiable diseases, and alerts. Direct, coordinate and promote the systematic monitoring of notifiable diseases and the microbiological notification system of Catalonia. Coordinate the actions of epidemiological control with the Spanish authorities and international institutions. Direct, coordinate and promote actions aimed at detecting and responding to alert situations and public health emergencies that require a rapid response. Promote the implementation of the different disease surveillance systems and determinants of health. Develop new monitoring systems and tools for epidemiological response. Promote research projects aimed at improving biomedical knowledge. Develop new lines of research in the field of epidemiology applied to public health policies. Ensuring resource levels, budget, health care professional competencies, skills and knowledge in order to achieve main goals. Coordinate training activities. Plan the communication activities both during outbreaks and in specific programs.

1.6. To whom are you responsible for performing these tasks?

1.6.1. University

1.6.2. Head of department, Alfonso Valencia

1.6.3. None in research, Department chair in the University

1.6.4. Catalonian Secretary of Public Health

2. Modelling background of the user

2.1. Are you a model developer or do you use models already developed?

2.1.1. Both (more developer than user)

2.1.2. Model developer

2.1.3. Both

2.1.4. I am a user

2.2. How relevant is in your research the use of models and simulation for addressing questions (low / medium / very high)

2.2.1. Very high

2.2.2. With Astrid Laegrid, very high. For other not so much.

2.2.3. Exclusively, very high
2.2.4. Very high

2.3. What kind of models and simulations do you use to analyse systems?
2.3.1. Epidemiological models
2.3.2. Statistical models of gene regulation using Decoupler
2.3.3. FEM, network, agents, ML
2.3.4. Forecasts, predictive models, scenarios simulation

2.4. How much would you regard yourself as a model developer?
2.4.1. Very much
2.4.2. Not so much
2.4.3. Very much
2.4.4. Not much

2.5. Which modelling and simulation approaches (e.g. ODEs, Stochastic, ABMs) and tools (e.g. programming/modelling languages, libraries, frameworks) do you use in your research?
2.5.1. Agent-based models, compartmental, empirical (data-driven). Matlab/Python/R
2.5.2. Statistical models of gene regulation using Decoupler
2.5.3. FEM, network, agents, ML. ODE, PDE, stochastics. Commercial software (abaqus), moFEM, Alya (collabo). Fortran, Python, matlab, C++. Libs: repast, vtk.
2.5.4. As a user, mainly ODE

2.6. Do you work with real-time data? If not, would you like to work with this kind of data?
2.6.1.1. I work with weekly reported data (not real time)
2.6.1.2. Yes, I would like.
2.6.1.3. No, but would like to.
2.6.1.4. Yes, real time data

2.6.2. Are you tools and workflows able to work with such data?
2.6.2.1. Currently yes, but it can be further improved
2.6.2.2. Yes
2.6.2.3. They could through the ML-based meta modelling.
2.6.2.4. Yes

2.6.3. Would you be interested in using a data processing workflow that would allow using real-time data?
2.6.3.1. Yes, that would be useful.
2.6.3.2. Yes
2.6.3.3. Yes
2.6.3.4. Yes

2.6.4. Would using this data allow you to address different problems than the ones you are currently addressing?
2.6.4.1. Not necessarily, but we would be more efficient analysing the current data we have.
2.6.4.2. Probably yes.
2.6.4.3. Yes

2.6.4.4. Yes

2.7. Do you currently use forecasting techniques? Are there specific events that you would like to forecast in real-time, which you currently cannot forecast?

2.7.1.1. Yes, we apply techniques to perform weekly forecast of seasonal respiratory diseases. Yes we would like to forecast number of hospitalization or primary care attendances, emergency departments calls, to ultimately forecast their impact over the health system.

2.7.1.2. Not currently, but could be interesting

2.7.1.3. I would like to, for onsite planification and circadian variations.

2.7.1.4. Yes, specifically related to the impact of respiratory diseases

The vision of CREXDATA is to develop a generic platform for real-time critical situation management including flexible action planning and agile decision making over data of extreme scale and complexity. CREXDATA develops the algorithmic apparatus, software architectures and tools for federated predictive analytics and forecasting under uncertainty. The envisioned framework boosts proactive decision making providing highly accurate and transparent short- and long-term forecasts to end-users, explainable via advanced visual analytics and accurate, real-time, off and on-site augmented reality facilities.

3. Please rate these objectives of the CREXDATA project according to your background and present and future needs (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful):

3.1. Being able to have extreme-scale data ingestion/generation, fusion and exploitation.

3.1.1. Ingesting multimodal data (images, simulations, social media publications, etc).

3.1.1.1. 5, 4, 3, 2

3.1.2. Using dynamic modelling to predict the systems' behaviour.

3.1.2.1. 4, 4, 5, 4

3.1.3. Handling multilingual social data in real-time.

3.1.3.1. 2, 2, 3, 2

3.2. Having real-time predictive knowledge and forecasts.

3.2.1. Using online federated learning.

3.2.1.1. 3, 3, 5, 4

3.2.2. Having multiresolution complex event forecasting under uncertainty.

3.2.2.1. 4, 4, 5, 5

3.2.3. Using optimization techniques for Prediction-as-a-Service (PaaS).

3.2.3.1. 4, 4, 5, 5

3.3. Reducing the perceived complexity.

3.3.1. Using graphical workflow design.

3.3.1.1. 4-5, 2, 2-3, 3

3.3.2. Using visual analytics coupled with XAI for understanding complexity and reasoning under uncertainty.

3.3.2.1. 4, 4, 5, 4

3.3.3. Using augmented reality under uncertainty on-site & remotely.

3.3.3.1. 1, 3, 5, 2

4. Specific aspects of the health emergency use case

4.1. Rate the following components of the health emergency use case according to your interest on the foreseen results (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful)

i. Using a graphical user interphase
   a. 5, 3, 5, 3

ii. Being able to have parameter calibration
   a. 5, 4, 5, 4

iii. Using early time series characterisation
   a. 5, 5, 5, 4

iv. Online model exploration
   a. 4, 5, 5, 3

v. Real-time/online forecasting simulation trajectories
   a. 3, 5, 5, 5

4.2. Rate the following Key Performance Indicators according to your interest on the foreseen results. (1: Not useful, 2: Of some use, 3: Average Use, 4: Quite useful, 5: Very useful)

i. KPI 1: Forecasting 7 parameter sets (interventions) that reduce the COVID infection of the simulation (multiscale infection or epidemiologic scenario).
   a. 4, 5, 2, 4

ii. KPI 2: Use the runtime adaptation of simulation trajectories to improve the outcomes of 5 scenarios or patients of the multiscale infection scenario.
   a. 4, 5, 4, 4

iii. KPI 3: Calibration of the epidemiological parameter to fit incidence time series.
   a. 5, NA, NA, 5

iv. KPI 4: Characterizing the space of parameters with 50% fewer simulations.
   a. 5, 4, 4, 3

5. System Usability Scale questions

5.1. SUS questions (1= Strongly disagree; 5 Strongly Agree): Totals: 60, 47.5, 10. Mean 39.167

1. I think that I would like to use this system frequently
   4, 4, 3, 5
2. I found the system unnecessarily complex
   2, 2, 5, 2
3. I thought the system was easy to use
   2, 4, 1, 4

4. I think that I would need the support of a technical person to be able to use this system
   4, 2, 5, 3
5. I found the various functions in this system were well integrated
   5, NA, NA, 4
6. I thought there was too much inconsistency in this system
   1, 3, 1, 2
7. I would imagine that most people would learn to use this system very quickly
   1, 3, 1, 2
8. I found the system very cumbersome to use
   3, NA, NA, 2
9. I felt very confident using the system
   5, NA, NA, 4
10. I needed to learn a lot of things before I could get going with this system
   3, 2, 5, 2

6. Further comments from expert user

"If this system is in production, there might be problem with MN5 use from a researcher/server outside of MN5. This aim is not compatible with current MN5 usage (slurm rules, security, etc). Real time data in MN5 is a problem as it is air-gapped."

"Concerns about use of agent-based modelling. Everything under the FEM would facilitate research and use. Maybe the dependency with HPC is a limiting factor. If the use case needs HPC the user base will be limited."

10.3 Maritime Use Case :

Regarding SUS calculation we consider positive the questions numbered 4.1,4.2,4.4, 5.1 and 5.3 while in lack of negative questions we will consider the default answer (3) for the five negative questions.

We provide SUS scores for each one of the experts in our questionnaire below:

1. 70
2. 70
3. 62.5
4. 60
5. 55
6. 52.5
7. 52.5

The average SUS score is 60.36

1. User Background Information

1.1. What is your Job Title
- 1.1.1. Researcher in Autonomous Systems
- 1.1.2. Software Engineer - Data Scientist
- 1.1.3. –
- 1.1.4. Research Associate
- 1.1.5. Master mariner
- 1.1.6. Associate Professor
- 1.1.7. Professor

1.2. Do you work in academia or in industry?
  1.2.1. Academia
  1.2.2. Industry
  1.2.3. Academia
  1.2.4. Academia
  1.2.5. Industry
  1.2.6. Academia
  1.2.7. Academia
1.3. Years of experience
  1.3.1. 5
  1.3.2. 15+
  1.3.3. 12
  1.3.4. 6
  1.3.5. 7
  1.3.6. 15
  1.3.7. 15
1.4. Background studies (university degree major, etc.)
  1.4.1. Master's
  1.4.2. Master's
  1.4.3. PhD
  1.4.4. Master's
  1.4.5. PhD
  1.4.6. PhD
1.5. What are your main job tasks?
  1.5.1. Research
  1.5.2. Software development & Big data Analysis & Mobility analytics
  1.5.3. Research on intelligent transportation systems and on robot motion planning, teaching.
  1.5.4. Teaching undergraduate courses (computer science/software engineering), participating in research projects, website development/maintenance, other IT-related work.
  1.5.5. Casualty investigation & Claims handling
  1.5.6. Energy and fuels
  1.5.7. Teaching, conducting research and administration.
1.6. To whom are you responsible for performing these tasks?
  1.6.1. Head of Research Group
  1.6.2. Contract owners
  1.6.3. Students
  1.6.4. University professors
  1.6.5. Partners
  1.6.6. Academia
  1.6.7. Dean and Superintendent of the Helleic Naval Academy
1.7. Did you participate in the 2023 Aegean Ro-Boat Race?
  1.7.1. Yes
  1.7.2. No
  1.7.3. Yes
  1.7.4. Yes
  1.7.5. No
  1.7.6. No

1.7.7. No

2. Modelling/Data Analysis Background

2.1. Are you a data scientist (model developer or similar) or do you use models already developed?
2.1.1. I build new models/ systems myself
2.1.2. I build new models/ systems myself
2.1.3. I build new models/ systems myself
2.1.4. No/ Not sure
2.1.5. No/ Not sure
2.1.6. No/ Not sure
2.1.7. No/ Not sure

2.2. How relevant is in your work the use of models and simulation for addressing questions (low / medium / very high) (optional - visibility depends on 2.1 question) (1: Not relevant, 2: Somewhat relevant, 3: Average relevant, 4: Quite relevant, 5: Very relevant)
2.2.1. 5
2.2.2. 4
2.2.3. 3
2.2.4. –
2.2.5. –
2.2.6. –
2.2.7. -

2.3. What kind of models and simulations do you use to analyse systems? (optional - visibility depends on 2.1 question)
2.3.1. Forecasting
2.3.2. Physics mobility models, custom
2.3.3. –
2.3.4. –
2.3.5. –
2.3.6. –
2.3.7. -

2.4. Which modelling and simulation approaches (e.g. ODEs, Stochastic, ABMs) and tools (e.g. programming/modelling languages, libraries, frameworks) do you use in your research? (optional - visibility depends on 2.1 question)
2.4.1. Deep Learning
2.4.2. Stochastic and ABMs
2.4.3. I am using Matlab for my research
2.4.4. –
2.4.5. –
2.4.6. –
2.4.7. -

2.5. Do you work with real-time data? (optional - visibility depends on 2.1 question)
2.5.1. No
2.5.2. Yes
2.5.3. No
2.5.4. –
2.5.5. –
2.5.6. –
2.5.7. -

2.6. Are your tools and workflows able to work with such data? (optional - visibility depends on 2.1 question)
   2.6.1. Yes
   2.6.2. Yes
   2.6.3. Yes
   2.6.4. –
   2.6.5. –
   2.6.6. –
   2.6.7. -
2.7. Would you be interested in using a data processing workflow that would allow using real-time data? (optional - visibility depends on 2.1 question)
   2.7.1. Yes
   2.7.2. Yes
   2.7.3. Yes
   2.7.4. –
   2.7.5. –
   2.7.6. –
   2.7.7. -
2.8. Would using this data allow you to address different problems than the ones you are currently addressing? (optional - visibility depends on 2.1 question)
   2.8.1. Yes
   2.8.2. Yes
   2.8.3. Yes
   2.8.4. –
   2.8.5. –
   2.8.6. –
   2.8.7. -
2.9. Do you try to forecast things into the future as part of your job? e.g forecast future vessel positions, or weather.
   2.9.1. Yes
   2.9.2. Yes
   2.9.3. Yes
   2.9.4. No
   2.9.5. Yes
   2.9.6. No
   2.9.7. Yes
2.10. Are there specific events that you would like to forecast in real-time, which you currently cannot forecast? (optional - visibility depends on 2.1 & 2.9 questions)
   2.10.1. Movement forecasting
   2.10.2. –
   2.10.3. Collision Risk assessment
   2.10.4. –
   2.10.5. Not that I know of.
   2.10.6. –
   2.10.7. Investigate detection probability of navy radars in any location, altitude and bearing
2.11. Do you currently use forecasting techniques?
   2.11.1. Yes
   2.11.2. No

2.12. Please select **any** of these data analysis tasks that are important for completing your tasks.

2.12.1.
+Investigate movement characteristics and sensor measurement recordings from a single boat in space and time. This includes detection of anomalies and unwanted behaviours, such as boat malfunctions or weather-related disruptions of its movement.;
+Assess the degree of stability in performing repeated movements and/or operations.;
+Detect and examine potential collision situations, in particular, during simultaneous movement of several vessels.;
+Forecast potential dangerous situations so that you gain manual control of the vessel

2.12.2.
+Investigate movement characteristics and sensor measurement recordings from a single boat in space and time. This includes detection of anomalies and unwanted behaviours, such as boat malfunctions or weather-related disruptions of its movement.;
+Assess the degree of stability in performing repeated movements and/or operations.;
+Detect and examine potential collision situations, in particular, during simultaneous movement of several vessels.;

2.12.3.
+Detect and examine potential collision situations, in particular, during simultaneous movement of several vessels.;

2.12.4.
+Investigate movement characteristics and sensor measurement recordings from a single boat in space and time. This includes detection of anomalies and unwanted behaviours, such as boat malfunctions or weather-related disruptions of its movement.;
+Assess the degree of stability in performing repeated movements and/or operations.;
+Detect and examine potential collision situations, in particular, during simultaneous movement of several vessels.;
+Forecast potential dangerous situations so that you gain manual control of the vessel

2.12.5.
+Detect and examine potential collision situations, in particular, during simultaneous movement of several vessels.;

2.12.6.
+Investigate movement characteristics and sensor measurement recordings from a single boat in space and time. This includes detection of anomalies and unwanted behaviours, such as boat malfunctions or weather-related disruptions of its movement.

2.12.7.
+Investigate movement characteristics and sensor measurement recordings from a single boat in space and time. This includes detection of anomalies and unwanted behaviours, such as boat malfunctions or weather-related disruptions of its movement.

+Forecast potential dangerous situations so that you gain manual control of the vessel

3. CREXDATA Objectives

3.1. Being able to have extreme-scale data ingestion/generation, fusion and exploitation.

3.1.1. 5
3.1.2. 5
3.1.3. 3
3.1.4. 3
3.1.5. 3 (default)
3.1.6. 3(default)
3.1.7. 2

3.2. Having real-time predictive knowledge and forecasts.

3.2.1. 5
3.2.2. 4
3.2.3. 3
3.2.4. 3
3.2.5. 3 (default)
3.2.6. 3(default)
3.2.7. 3(default)

3.3. Reducing the perceived complexity.

3.3.1. 5
3.3.2. 4
3.3.3. 3
3.3.4. 4
3.3.5. 3 (default)
3.3.6. 3(default)
3.3.7. 4

4. Maritime use case

Collision avoidance service

4.1. How useful is a collision avoidance rerouting service for your day-to-day work?

4.1.1. 5
4.1.2. 4
4.1.3. 4
4.1.4. 4
4.1.5. 3
4.1.6. 1
4.1.7. 2

4.2. How important is a collision avoidance and rerouting service to provide COLREG compliant recommendations?
(1: Not Important, 2: Of some importance, 3: Average Importance, 4: Quite Important, 5: Very Important)

4.2.1. 5
4.2.2. 5
4.2.3. 5
4.2.4. 3 (default)

- 4.2.5. 3 (default)
- 4.2.6. 4
- 4.2.7. 2

4.3. How often should a collision avoidance service provide re-routing recommendations?
- 4.3.1. Every 1-minute
- 4.3.2. Sub-second
- 4.3.3. Every 5-minutes
- 4.3.4. Every 1-minute
- 4.3.5. Every 5-minutes
- 4.3.6. Every 30-minutes
- 4.3.7. Every 5-minutes

4.4. How important is the existence of a graphical interface for a collision avoidance rerouting service?
- 4.4.1. 5
- 4.4.2. 5
- 4.4.3. 4
- 4.4.4. 5
- 4.4.5. 5
- 4.4.6. 4
- 4.4.7. 4

4.5. Would the existence of an Augmented Reality interface improve usability of the collision avoidance rerouting service?

- 4.5.1. Minor improvement
- 4.5.2. Minor improvement
- 4.5.3. Major improvement
- 4.5.4. Minor improvement
- 4.5.5. Minor improvement
- 4.5.6. No improvement
- 4.5.7. Major improvement

4.6. Are there any collision avoidance rerouting recommendation engines that you already use?
- 4.6.1. No
- 4.6.2. No
- 4.6.3. No
- 4.6.4. No
- 4.6.5. No
- 4.6.6. No
- 4.6.7. No

4.7. Which collision avoidance recommendation engines you use? (visibility depends on 4.6)
- 4.7.1. – no answer received

4.8. In a crossing situation that your vessel should give priority to the other vessel. Which recommendations you would consider safe to follow.
- 4.8.1.
  Pass from stern side, keeping at least two lengths distance

4.8.2.
Pass from stern side;
Pass from stern side, keeping at least one length distance;
Pass from stern side, keeping at least two lengths distance
4.8.3.
Pass from bow side
4.8.4.
Pass from stern side, keeping at least two lengths distance
4.8.5.
Pass from stern side;
Pass from stern side, keeping at least one length distance;
Pass from stern side, keeping at least two lengths distance
4.8.6. Pass from stern side, keeping at least two lengths distance
4.8.7. Pass from stern side, keeping at least two lengths distance

5. Extreme-Weather re-routing service

Example:The crew of a vessel starts their journey and plan their route according to their initial weather forecast. As weather dynamically changes over the journey, the vessel crew receives hourly updates in case of weather conditions influencing the passage safety through specific sea areas. In case of changes affecting the safe passage, an alert with automatic rerouting suggestion is generated by the system alleviating the vessel crew from the task of continuously monitoring the weather conditions and updating the vessel route.

5.1. How useful is an extreme weather condition rerouting service for your day-to-day work?
5.1.1. 5
5.1.2. 4
5.1.3. 3(default)
5.1.4. 3(default)
5.1.5. 3(default)
5.1.6. 4
5.1.7. 4

5.2. How often should a recommendation engine provide re-routing recommendations due to extreme weather conditions during a voyage?
5.2.1. Hourly
5.2.2. Every 6-hours
5.2.3. Every 6-hours
5.2.4. Hourly
5.2.5. Every 6-hours
5.2.6. Sub-second
5.2.7. Hourly

5.3. How important is the existence of a graphical interface for a weather rerouting service?
5.3.1. 4
5.3.2. 5
5.3.3. 4
5.3.4. 4
5.3.5. 3
5.3.6. 4
5.3.7. 4

5.4. Would the existence of an Augmented Reality interface improve usability of the extreme weather condition rerouting service?
  5.4.1. No improvement
  5.4.2. Minor improvement
  5.4.3. Minor improvement
  5.4.4. Minor improvement
  5.4.5. No improvement
  5.4.6. No improvement
  5.4.7. Major improvement

5.5. Are there any extreme weather conditions rerouting recommendation engines that you already use?
  5.5.1. No
  5.5.2. No
  5.5.3. No
  5.5.4. No
  5.5.5. No
  5.5.6. No
  5.5.7. No

5.6. Which collision avoidance recommendation engines you use? (visibility depends on 5.5)
  5.6.1. – no answer received

6. System Usability/Scale questions

6.1. A single trajectory represented on a map (left) and in a space-time cube (right). The time axis in the space-time cube is oriented upwards. Which is more useful for you (in real time)?
  6.1.1. Map projection (left)
  6.1.2. Map projection (left)
  6.1.3. Space-time cube (Right)
  6.1.4. Map projection (left)
  6.1.5. Map projection (left)
  6.1.6. Map projection (left)
  6.1.7. Map projection (left)

6.2. Regarding the previous, if analysis was performed at a later date on historical data, would your selection change?
  6.2.1. Yes
  6.2.2. Yes
  6.2.3. No
  6.2.4. No
  6.2.5. No
  6.2.6. No
  6.2.7. No

6.3. Both these data visualisation show the data from one of the races for three competing boats. In the left the visualisation involves only spatial data in a 2D map while the right is temporospatial (space-time cube).
Which one is more useful for you in real time decision making?
  6.3.1. Space-time cube (Right)
  6.3.2. Space-time cube (Right)
  6.3.3. Space-time cube (Right)

- 6.3.4. Map projection (left)
- 6.3.5. Map projection (left)
- 6.3.6. Map projection (left)
- 6.3.7. Map projection (left)

6.4. Similarly, which is easier for you to spot the interactions between two vessels?

- 6.4.1. Space-time cube (Right)
- 6.4.2. Space-time cube (Right)
- 6.4.3. Space-time cube (Right)
- 6.4.4. Map projection (left)
- 6.4.5. Map projection (left)
- 6.4.6. Map projection (left)
- 6.4.7. Map projection (left)

6.5. Comments: Please write your comments below:

- 6.5.1. No comments received.

11 Appendix 3: EmCase detailed test cases and settings

11.1 Test case specifications

Table 14: TC_001 ARGOS demonstration Innsbruck

Test_Case_Name	ARGOS demonstration Innsbruck
Test_Case_ID	TC_001
Test_Item	ARGOS system
Test_Procedure	Demonstration (with in-situ data input of a reference event: heavy rain with flooding on 2nd July 2016 in Innsbruck)
CREXDATA_system_level	sub-system (ARGOS with T2.1)
Test_Case_Owner	DCNA (HYDS)
Use_Case_underlaying	Emergency_UC_01
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	Hybrid demonstration with technical partner online and stakeholders present in a workshop setting
System_State	ARGOS system active
User	A level commander
Input Parameter	• Water levels from HYDRO Tyrol • Precipitation data from GeoSphere Austria ○ Station data - 10min precipitation ○ Areal data – INCA hourly data ○ Areal data – Spartacus daily data ○ Forecasted data – short-cast forecast ○ Forecasted data – warnings • Emergency protocols • Geodata Services ○ Background layers (DGM, Orthophoto, rivers, lakes) ○ vulnerable elements (schools, museums, etc.) ○ infrastructure (transport lines, bridges, tunnels, airport, power lines) ○ hazard zones (flooding, water levels, debris flows, avalanche) ○ industrial hazard zones ○ contaminated areas (waste management zones) ○ sewage treatment plants, waste water discharge plants ○ drinking water structures
Test_Case_Sequence	1. Starting ARGOS system 2. Choose ARGOS project (city of Innsbruck) 3. Define area of interest 4. Demonstrate GUI 5. Demonstrate functionality of the early warning system 6. Visualize different layers in different situations 7. Evaluation
Expected_Result	Evaluation of system usability

Table 15: TC_002 ARGOS demonstration Dortmund

Test_Case_Name	**ARGOS demonstration Dortmund**
Test_Case_ID	TC_002
Test_Item	ARGOS system
Test_Procedure	Demonstration
CREXDATA_system_level	sub-system (ARGOS with T2.1)
Test_Case_Owner	FDDO (HYDS)
Use_Case_underlaying	Emergency_UC_01
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	Hybrid demonstration with technical partner online and stakeholders present in a workshop setting
System_State	ARGOS system active
User	A level commander

Input Parameter	• Water levels from HYDRO Tyrol
	• Precipitation data from GeoSphere Austria
	  ○ Station data - 10min precipitation
	  ○ Areal data – INCA hourly data
	  ○ Areal data – Spartacus daily data
	  ○ Forecasted data – short-cast forecast
	  ○ Forecasted data – warnings
	• Emergency protocols
	• Geodata Services
	  ○ Background layers (DGM, Orthophoto, rivers, lakes)
	  ○ vulnerable elements (schools, museums, etc.)
	  ○ infrastructure (transport lines, bridges, tunnels, airport, power lines)
	  ○ hazard zones (flooding, water levels, debris flows, avalanche)
	  ○ industrial hazard zones
	  ○ contaminated areas (waste management zones)
	  ○ sewage treatment plants, waste water discharge plants

Test_Case_Sequence	1. Starting ARGOS system 2. Choose ARGOS project (city of Dortmund) 3. Define area of interest 4. Demonstrate GUI 5. Demonstrate functionality of the early warning system 6. Visualize different layers in different situations 7. Evaluation
Expected_Result	Evaluation of system usability

Table 16: TC_003 CEF sewer network

Test_Case_Name	**CEF sewer network**
Test_Case_ID	TC_003
Test_Item	CEF sewer network
Test_Procedure	Demonstration of CEF for discharge events in the Innsbruck sewer network during and after a heavy rain event (reference event: heavy rain with flooding on 2nd July 2016 in Innsbruck)
CREXDATA_system_level	sub-system (T4.1)

Test_Case_Owner	DCNA (NCSR)
Use_Case_underlaying	Emergency_UC_10, Emergency_UC_12
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	Precipitation data from IKB of 5 different locations with coordinates from 01.07.-04.07. for every minute
System_State	CEF sewer network is active and producing forecasts
User	IKB - Communal services

Input Parameter	• Weather data from GeoSphere Austria
	  ○ Station data - 10min precipitation
	  ○ Areal data – INCA hourly data
	  ○ Areal data – Spartacus daily data
	  ○ Forecasted data – short-cast forecast
	• Water levels of 5 different locations with coordinates in the canals from 01.07.-04.07. for every minute
	• Water levels of 3 different locations with coordinates of rivers and streams from 01.07.-04.07. for every minute
	• Discharge Events of flood pumping stations 01.07-.04.07.

Test_Case_Sequence	Final definiton under discussion with technology partner.
Expected_Result	Evaluation of system usability and relevance

Table 17: TC_006 3D mapping using UAV imagery

Test_Case_Name	3D mapping using UAV imagery
Test_Case_ID	TC_006
Test_Item	UAV, RobLW, NeRF processing
Test_Procedure	Physical_Test
CREXDATA_system_level	sub-system (T2.1, T4.3)
Test_Case_Owner	DRZ (TUC)
Use_Case_underlaying	Emergency_UC_02, Emergency_UC_12, Emergency_UC_20, Emergency_UC_21
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	Brightness in daylight: 25,000 lumen
System_State	UAV deployed and in operation
User	UAV operator, C level commander
Input Parameter	Number of photos required: 120; area of interest: 100x100 meters; flight altitude: 60 meters; image resolution: 12 MP: Number of repeats: 1
Test_Case_Sequence	1. Prepare the UAV and data transmission/processing 2. Create a survey mission 3. Launch the survey mission 4. Send the images to RobLW server 5. Calculate a NeRF 6. Visualize NeRF on GUI A pilot uses a consumer drone to survey the designated area along a programmed trajectory automatically collecting images at regular intervals. The acquired images are transmitted to the RobLW server at DRZ for processing. The server is used to create a *Neural Radiance Field (NeRF)* of

	the site-a relatively new, AI-based approach for representing 3D scenes. For comparison purposes, the same image set is processed in parallel on another server using the well-known photogrammetry tool *Open Drone Map (ODM)*, which uses longer-established methods to create orthophotos and 3D models from UAV surveys. Resource and time consumption of both tools are recorded.
Expected_Result	Successful creation of a 3D model of the target area Experts fill in questionnaires and TUC and DRZ conduct semi-structured expert interviews with the attending end-users to obtain qualitative feedback on the achieved model quality of each tool, their usability, and other relevant factors with respect to real-life emergency scenarios. Proposed improvements of the NeRF technology (and the even newer, related technology of Gaussian splatting) for these scenarios, using the CREXDATA system, are discussed, refined, and prioritized. This test case establishes a baseline to measure developments towards said improvements against in future trials.

Table 18: TC_007 Text mining Innsbruck

Test_Case_Name	Text mining Innsbruck
Test_Case_ID	TC_007
Test_Item	Text mining, Twitter/X
Test_Procedure	Demonstration, Usability Test
CREXDATA_system_level	sub-system (T4.5)
Test_Case_Owner	DCNA (BSC/NCSR)
Use_Case_underlaying	Emergency_UC_13
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	• Heavy rain in the area of interest • Information quality of the situation picture not sufficient
System_State	sub-system “Text mining” active
User	B-/C-Level fire officers, who coordinate forces
Input Parameter	• Updated list of relevant keywords • Results of the manual search for Twitter posts for the event on 2 July 2016 as a reference dataset
Test_Case_Sequence	• start text mining system • select search terms • start collection of relevant posts based on selection of keywords (filter algorithm) • automated identification of flooding events in demand of emergency response • providing decision makers with categorized/prioritized list of locations/events with possibility of own exploration of underlying information included in posts • visualize relevant postings
Expected_Result	Selected social media posts are relevant for decision-makers and broaden or clarify situational awareness

Table 19: TC_008 Text mining Dortmund

Test_Case_Name	Text mining Dortmund
Test_Case_ID	TC_008
Test_Item	Text mining, Twitter/X
Test_Procedure	Demonstration, Usability Test
CREXDATA_system_level	sub-system (T4.5)
Test_Case_Owner	FDDO (BSC/NCSR)
Use_Case_underlaying	Emergency_UC_13
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	• Heavy rain in the area of interest • Information quality of the situation picture not sufficient
System_State	sub-system “Text mining” active
User	B-/C-Level fire officers, who coordinate forces
Input Parameter	Updated list of relevant keywords
Test_Case_Sequence	• start text mining system • select search terms • start collection of relevant posts based on selection of keywords (filter algorithm) • automated identification of flooding events in demand of emergency response • providing decision makers with categorized/prioritized list of locations/events with possibility of own exploration of underlying information included in posts • visualize relevant postings
Expected_Result	Suggestion of current flooding event which demands deployment of rescue forces (with additional information: location, severity, impact on infrastructure/people); Optimization of filters and keyword sets

Table 20: TC_011 AR routing Innsbruck

Test_Case_Name	AR routing Innsbruck
Test_Case_ID	TC_011
Test_Item	Augmented Reality device: HoloLens 2 and its application
Test_Procedure	Physical_Test
CREXDATA_system_level	sub-system (T5.4)
Test_Case_Owner	DCNA (TUC)
Use_Case_underlaying	Emergency_UC_52
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	• certain water level within the operation area • global coordinates • closed paths and roads
System_State	• AR device active and in use • AR application running
User	C-Level operator
Input Parameter	• GPS coordinates as a location of the operator

	• Nodes and links in the sewer network • Catchments and sewer network data • Digital terrain and surface model • Coordinates of POIs in the area of interest • grid elements • precipitation data
Test_Case_Sequence	1. Activate Augmented Reality device 2. Define area of interest 3. Load input data 4. Request simulated water level through CREXDATA system and ARGOS a. Simulation configuration via ARGOS b. Configuration file written by ARGOS c. Start simulation with batch process in MIKE+ d. ARGOS reading simulation output file 5. Pull simulation output from ARGOS system using CREXDATA system 6. Visualize water level in the area of interest 7. Visualize points of interest 8. Move thorugh area of interest 9. Turn off AR device and its application
Expected_Result	• user-friendly visualization of POIs when moving through the area under consideration • continuous, trouble-free visualization of the predicted water level • feedback from C-Level operators regarding usability and further developments

Table 21: TC_012 AR visualizing points of interest Innsbruck

Test_Case_Name	AR visualizing points of interest Innsbruck
Test_Case_ID	TC_012
Test_Item	Augmented Reality device: HoloLens 2 and its application
Test_Procedure	Physical_test
CREXDATA_system_level	sub-system (T5.4)
Test_Case_Owner	FDDO (TUC)
Use_Case_underlaying	Emergency_UC_51, Emergency_UC_52
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	• certain water level within the operation area • global coordinates
System_State	• AR device active and in use • AR application running
User	C-Level operator
Input Parameter	• GPS coordinates as a location of the operator • Nodes and links in the sewer network • Catchments and sewer network data • Digital terrain and surface model • Coordinates of POIs in the area of interest • grid elements • precipitation data

Test_Case_Sequence	1. Activate Augmented Reality device 2. Define area of interest 3. Load input data 4. Request simulated water level through CREXDATA system and ARGOS a. Simulation configuration via ARGOS b. Configuration file written by ARGOS c. Start simulation with batch process in MIKE+ d. ARGOS reading simulation output file 5. Pull simulation output from ARGOS system using CREXDATA system 6. Visualize water level in the area of interest 7. Visualize points of interest 8. Move thorugh area of interest 9. Turn off AR device and its application
Expected_Result	• user-friendly visualization of POIs when moving through the area under consideration • continuous, trouble-free visualization of the predicted water level • feedback from C-Level operators regarding usability and further developments

Table 22: TC_013 AR routing Dortmund

Test_Case_Name	AR routing Dortmund
Test_Case_ID	TC_013
Test_Item	Augmented Reality device: HoloLens 2 and its application
Test_Procedure	Physical_test
CREXDATA_system_level	sub-system (T5.4)
Test_Case_Owner	FDDO (TUC)
Use_Case_underlaying	Emergency_UC_52
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	• certain water level within the operation area • global coordinates • closed paths and roads
System_State	• AR device active and in use • AR application running
User	C-Level operator
Input Parameter	Final definiton under discussion with technology partner.
Test_Case_Sequence	1. Activate Augmented Reality device 2. Define area of interest 3. Load input data 4. Request simulated water level through CREXDATA system and ARGOS a. Simulation configuration via ARGOS b. Configuration file written by ARGOS c. Start simulation with batch process in MIKE+ d. ARGOS reading simulation output file 5. Pull simulation output from ARGOS system using CREXDATA system 6. Visualize water level in the area of interest

	7. Visualize points of interest 8. Move through area of interest 9. Turn off AR device and its application
Expected_Result	• user-friendly visualization of POIs when moving through the area under consideration • continuous, trouble-free visualization of the predicted water level • feedback from C-Level operators regarding usability and further developments

Table 23: TC_014 AR visualizing points of interest Dortmund

Test_Case_Name	AR visualizing points of interest Dortmund
Test_Case_ID	TC_014
Test_Item	Augmented Reality device: HoloLens 2 and its application
Test_Procedure	Physical_test
CREXDATA_system_level	sub-system (T5.4)
Test_Case_Owner	FDDO (TUC)
Use_Case_underlaying	Emergency_UC_51, Emergency_UC_52
Detailed_Description	
Environmental_Conditions	
Test_Case_ Environmental_Parameter	• certain water level within the operation area • global coordinates
System_State	• AR device active and in use • AR application running
User	C-Level operator
Input Parameter	Final definiton under discussion with technology partner.
Test_Case_Sequence	1. Activate Augmented Reality device 2. Define area of interest 3. Load input data 4. Request simulated water level through CREXDATA system and ARGOS a. Simulation configuration via ARGOS b. Configuration file written by ARGOS c. Start simulation with batch process in MIKE+ d. ARGOS reading simulation output file 5. Pull simulation output from ARGOS system using CREXDATA system 6. Visualize water level in the area of interest 7. Visualize points of interest 8. Move thorugh area of interest 9. Turn off AR device and its application
Expected_Result	• user-friendly visualization of POIs when moving through the area under consideration • continuous, trouble-free visualization of the predicted water level • feedback from C-Level operators regarding usability and further developments

11.2 Themes available in open Tyrol data

Table 24: OGD Data by the province of Tyrol via https://gis.tirol.gv.at/arcgis

Themes	Data [GERMAN]	Data [ENGLISH]
Basis	basis_dgm	DTM
	basis_dom	DSM
	basis_karte	BASIS MAP
	basis_karte_label	BASIS MAP description
	basis_oek	Austrian Map 1:50000 Basis Map
	basis_ortho	Orthophoto
	basis_ortho_label	Orthophoto description
	basis_overview	Basis map overview
HIK	HIK_MD	Old map
	HIK_Metadaten	Old map Metadata
	HIK_Metadaten	Old map Metadata
INSPIRE	INSPIRE/AT_0024_05_Adressen	Addresses
	INSPIRE/AT_0024_07_Verkehrsnetze	Transport network
	INSPIRE/AT_0024_09_Schutzgebiete	Protected areas
	INSPIRE/AT_0024_10_Hoehe	height
	INSPIRE/AT_0024_15_Gebaeude	buildings
	INSPIRE/AT_0024_17_Bodennutzung	land use
	INSPIRE/AT_0024_19_Daseinsvorsorge	public services
	INSPIRE/AT_0024_20_Umweltueberwachung	environmental monitoring
	INSPIRE/AT_0024_21_Seveso	landfills
	INSPIRE/AT_0024_24_Bewirtschaftungsgebiete	Cultivation areas
	INSPIRE/AT_0024_25_Naturgefahren	Natural hazards
	INSPIRE/AT_0024_31_Biotope	biotopes
	INSPIRE/AT_0024_33_Energiequellen	energy sources
Service Public	See Table 5	

Table 25: OGD Data - Folder Service_Public

Theme: Service_Public	GERMAN		ENGLISH	
	subgroup	sub-subgroup	subgroup	sub-subgroup
/fahrverbote_reisetage	Fahrverbote Reisetage		Driving bans travelling days	
/geoland	Polygondecker		Polygon cover of neighbouring countries	
	Kirchen		Churches	
	Museen		museums	
	Schulstandorte		School locations	
	Burgschloss		Castle	
	Wasserschongebiet_allgemein		Water protection area_general	
	Ueberflutungsflaechen		Floodplains	
	Habitatrichtlinie_SCI		Habitats Directive_SCI	
	Vogelschutzrichtlinie_SPA		Bird Protection Directive_SPA	
	Geschuetzter_Landschaftsteil		Protected_landscape_part	
	Landschaftsschutzgebiet		Protected landscape area	
	Naturpark		Nature park	
	Naturschutzgebiet		Nature reserve	
	Sonstiges_Schutzgebiet		Other_protected_area	
	Widmungskategorie		Dedication category	

	Waldentwicklungsplan		Forest development plan	
/NPHT_Gl	Labels		Labels	
	Glacier Inventory_Gl3		Glacier Inventory_Gl3	
	Glacier Inventory_Gl2		Glacier Inventory_Gl2	
	Glacier Inventory_Gl1		Glacier Inventory_Gl1	
	Glacier Inventory_LIA		Glacier Inventory_LIA	
/OGD	service stopped			
/ogd_agrar	Höfe und Almen		Farms and pastures	
	Grundzusammenlegungen		Land consolidation	
	Landwirtschaftliche Böden		Agricultural soils	
/ogd_basis	Beschriftung		Labelling	
	Adressen		Addresses	
	Kataster		Cadastre	
	Verwaltungsgrenzen		Administrative boundaries	
	Blattschnitte		Sheet Sections	
	Datenstand Orthofoto		Data status orthophoto	
	Datenstand Höhenmodell		Data status elevation model	
/ogd_bildung	Schulen		Schools	
	Kinderbetreuung		Childcare	
/ogd_infrastruktur	Öffentlicher Verkehr	ÖV Haltestellen	Public transport	Public transport stops
		ÖV Haltestellensteige		Public transport stop platforms
		ÖV Linienrouten		Public transport routes
	Verkehrswege	Verkehrswege Übersicht	Traffic routes	Routes overview
		Verkehrswege Beschriftung		Routes Labelling
		Verkehrswege Detail		Traffic routes detail
	Hochrangige Straßennetz	Grundstücke LSV B+L	High-ranking road network	Properties LSV B+L
		Kilometer		Kilometres
		Brücken und Tunnel		Bridges and tunnels
		Straßen		Roads
		Ortsgebiete		Local areas
		Polizeirayone		Police zones
		Straßenmeistereien		Road maintenance depots
	Bahn	Kilometer	Railway	Kilometres
		Brücken und Tunnel		Bridges and tunnels
		Bahn		Railways
	Behindertenparkplätze		Disabled parking spaces	
	Flugverkehr	Flughafen	Air traffic	Airport
		Flughindernisse		Airport obstacles
	Versorgung	Strom Leitungen	Utilities	Electricity lines
		Gasversorgung Leitungen		Gas supply Pipelines
/ogd_natur	-			
/ogd_naturgefahren	Bundesbauverwaltung (BWV)	Bearbeitungsstand	Federal Building Administration (BWV)	Processing status
		Besondere Gefährdung		Special hazards
		Gefahrenzonen und Funktionsbereiche		Hazard zones and functional areas
		Überflutungsflächen		Flood areas
		Wassertiefen		Water depths
	Wildbach- und Lawinenverbauung (WLV)	Planungsbereich	Torrent and Avalanche Control (WLV)	Planning area
		Gefahrenzone Lawine		Avalanche danger zone
		Beschriftung Lawine		Avalanche labelling
		Gefahrenzone Wildbach		Torrent danger zone
		Beschriftung Wildbach		Torrent labelling
		Brauner Hinweisbereich		Brown indication area
		Blauer Vorbehaltsbereich		Blue restricted area

		Violetter Hinweisbereich		Purple warning area
		Einzugsgebiete Lawine		Avalanche catchment area
/ogd_raumordnung	Örtliches Raumordnungskonzept		Local spatial planning concet	
	Flächenwidmung		Zoning	
	Überörtliche Raumordnung	Freihaltegebiete	Supra-local spatial planning	Open areas
		Einkaufszentren		Shopping centres
		Schigebietsprogramm		Ski area programme
	Sonstige örtliche Raumordnung		Other local spatial planning	
	Industriegefahren	Seveso II Betrieb	Industrial hazards	Seveso II operation
		Seveso II Betrieb -		Seveso II operation - hazard area
	Versorgungseinrichtungen		Supply facilities	
/ogd_sport				
/ogd_umwelt	Abfallwirtschaft	Altablagerungen (Punkte)	Waste management	Old Contaminated sites (points)
		Altlasten		Contaminated sites
	Luftgüte		Air quality	
	Umgebungslärm		Environmental noise	
/ogd_wald			-	
/ogd_wasser	Hydrogeologie	Erdwärmesonde	Hydrogeology	Geothermal probe
		Baugrundaufschluss		Soil exploration
		Bewilligungspflicht Erdwärmesonde		Authorisation requirement for geothermal probes
	Abwasser	Kläranlage	Waste water	Sewage treatment plant
		Abwassersicherung		Wastewater protection
		Entlastungsanlage		Discharge plant
		Einleitung WB		Discharge WB
		Entsorgungsgebiet		Disposal area
	Wasserkraft /Beschneiung	Wehranlage WB	Hydropower / Snow making	weir system WB
		Kraftwerk WB		Power station WB
		Mühle WB		Mill WB
		Speicher		Reservoir
		Rückleitung		Return line
	Wasserversorung / Grundwassernutzung	Quelle	Water supply / groundwater utilisation	Source
		Grundwasserrückgabe		Groundwater return
		Grundwasserentnahme		Groundwater withdrawal
		Grundwassersonde		Groundwater probe
		GW Schichtlinien hoch		GW stratification lines high
		GW Schichtlinien mittel		GW stratification lines medium
		GW Schichtlinien tief		GW stratigraphic lines low
		Trinkwasserbauwerk		Drinking water structure
		Schutz- und Schongebiet		Protection and protection zone
		Trinkwasserversorgungszone		Drinking water supply zone
	Sonstige Wasserinformationen	Alle Wasserinformationen (Übersicht)	Other water information	All water information (overview)
		Betrieb, Gebäude		Operation, buildings
		Indirekte Wassernutzung		Indirect water utilisation
		Sonderanlage		Special facility
	Hydrographische Messstellen	Messstelle - Quelle	Hydrographic measuring points	Measuring point - source
		Messstelle - Niederschlag		Measuring point - precipitation
		Messstelle - Grundwasser		Measuring point - groundwater
		Messstelle - Gewässer		Measuring point - water bodies
	Flüsse / Seen	Gewässerzustand	Rivers / lakes	Water status
		Öffentliches Wassergut		Public water resources
		Flüsse km		Rivers km
		Flüsse		Rivers
		Seen		Lakes

		Zuständigkeit BWV		Responsibility BWV
		Einzugsgebiete Gewässer		Catchment areas Water bodies
		Einzugsgebiete WLV		Catchment areas WLV
		Fischereireviere		Fishing districts
/OPH_Sturmschaden_2018	-			
/orthofoto	Raster data			
/terrain	Raster data			

11.3 Possible locations in Dortmund

1. Mackenrothweg 15 44328 Dortmund (low traffic)

Two Schools

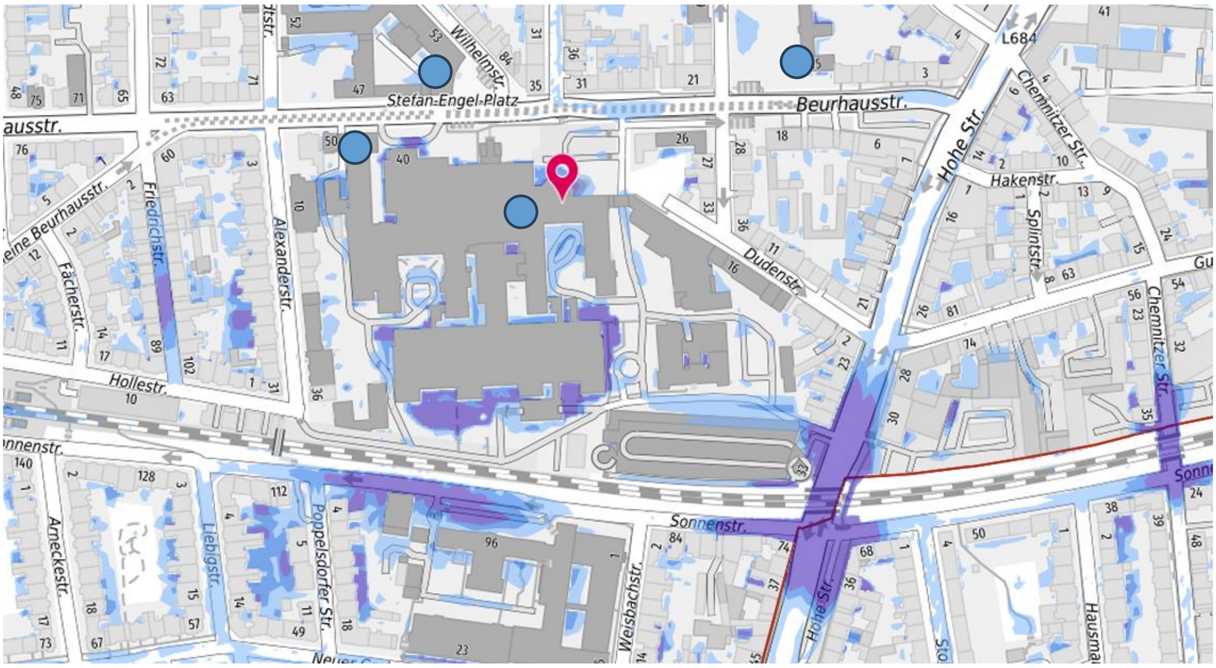
Five nursery schools

2. Beurhausstraße 40 44137 Dortmund (high traffic)

One Nursery school

Two Hospitals

One School

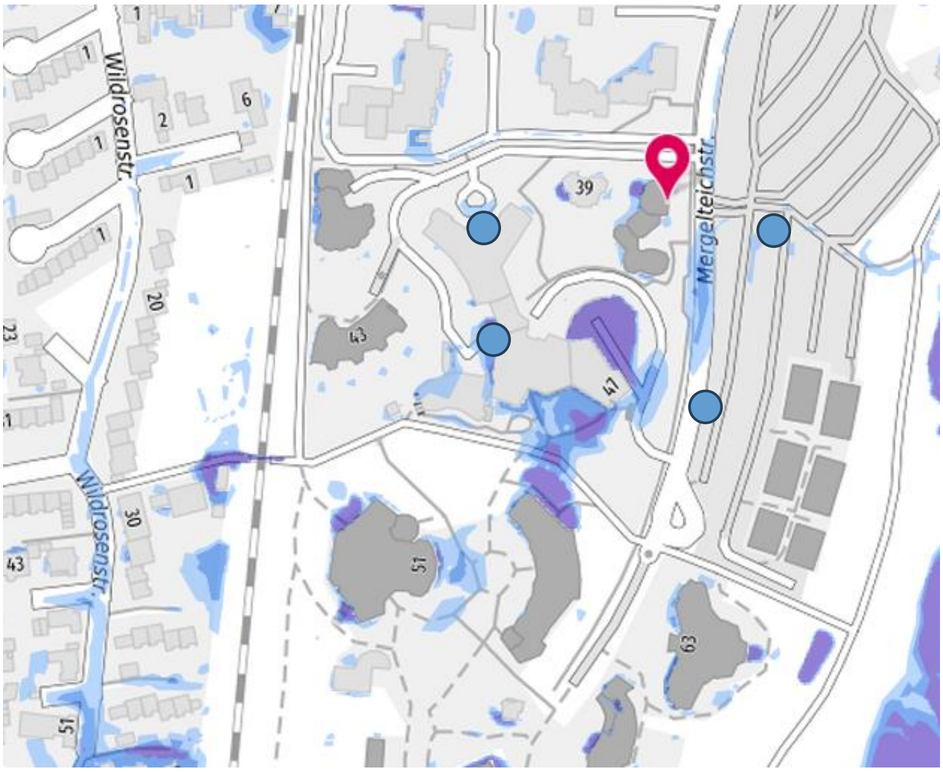


3. Mergelteichstraße 45 44225 Dortmund (medium traffic)

One School

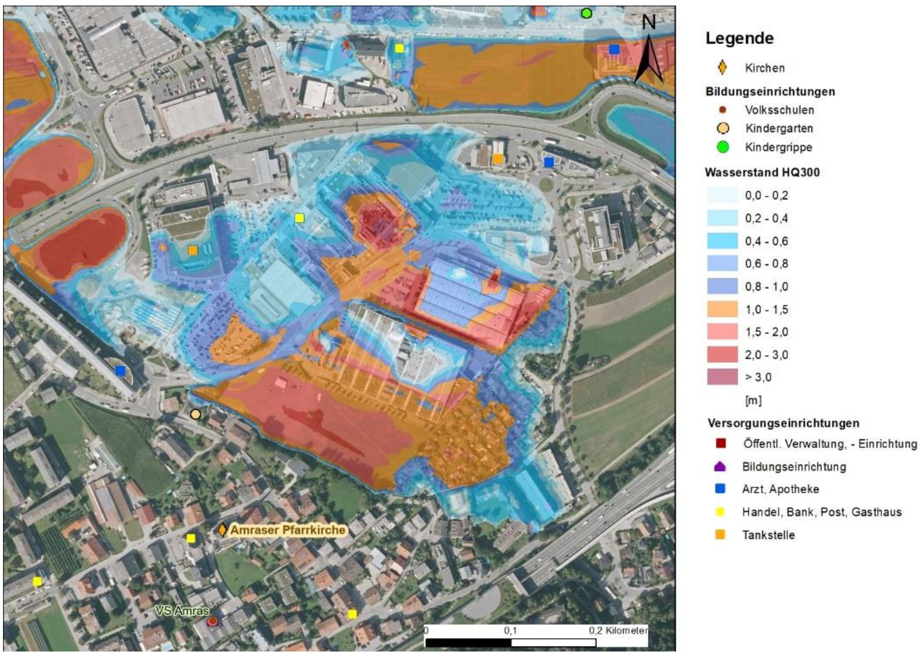
Two Nursery Schools

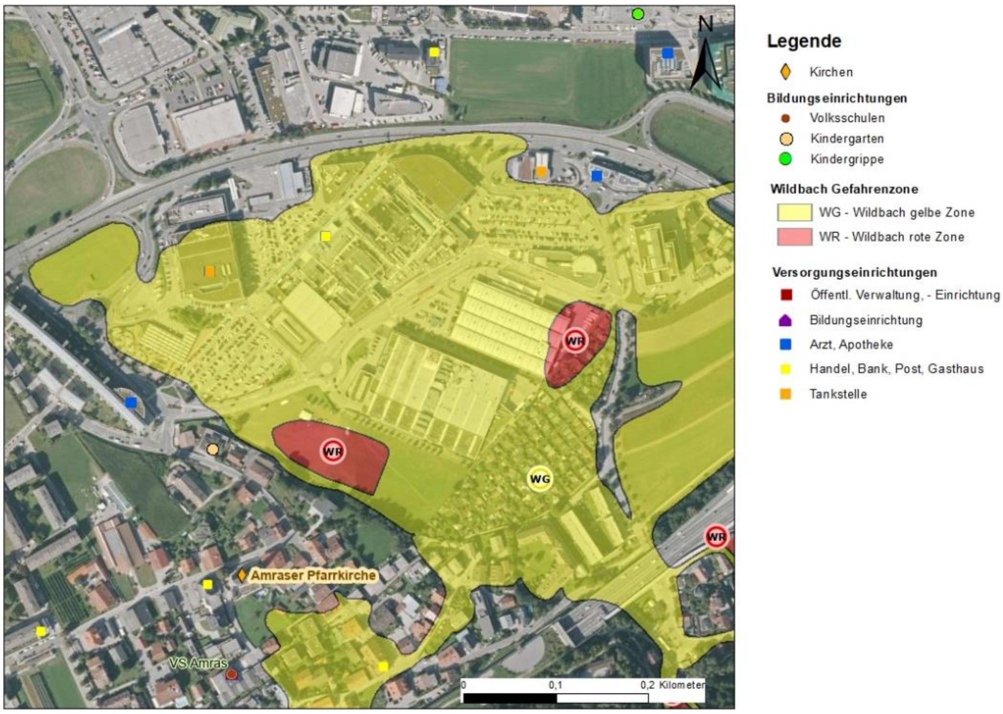
One Retirement Home



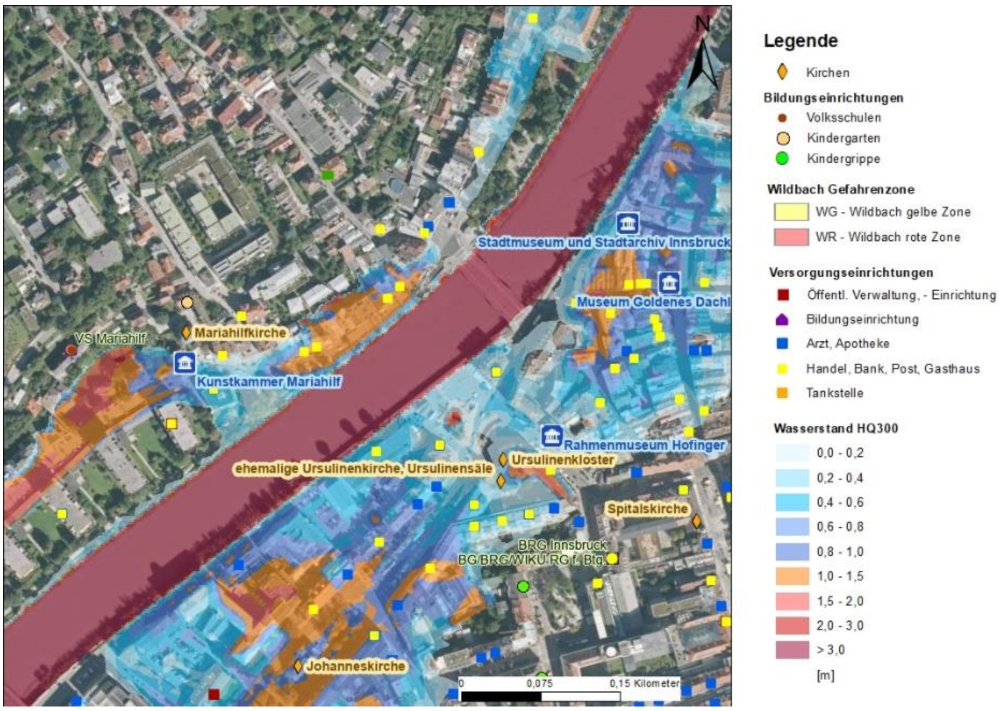
11.4 Possible locations in Innsbruck

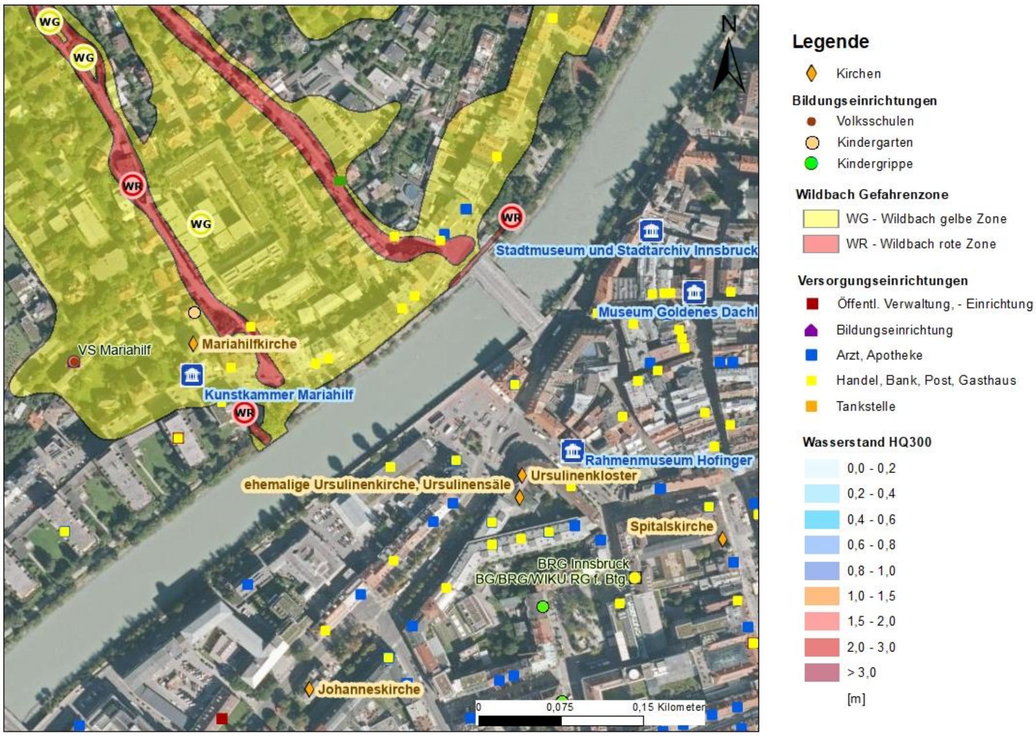
1. Amras district





2. Innsbruck Old town





11.5 Pilot definition and demonstrator details:

11.5.1 Pre-study: Acquired data based on use case narratives and lab tests

Detailed data collection, pre-processing and fine-tuning with technology partners is understood as a pre-study with specific technologies and referring to specific use cases at one or more pilot sites. Therefore, use case narratives and especially their "test settings" were utilized from Deliverable D2.1.

11.5.1.1 Data in Innsbruck

Data in Innsbruck has been compiled from the following sources:

Open-Source Datahub of Austria (data.gv.at)

- *Digital Terrain Model (DTM) / Digital Surface Model (DSM)*
  DTM is the elevation-based description of the terrain, represented by a georeferenced raster grid with a spatial resolution of 1 to 5m. The Digital Surface Model (DSM) includes, in contrast to a DTM, the objects on the earth's surface in three dimensions[10]. (raster file)
- *Orthophoto 2022*
  Digital orthophotos are distortion-free and georeferenced aerial images. They are true to scale and can therefore be combined directly with maps or specialized data. The surveying authorities of the federal state of Tyrol produce Orthophotos with a ground resolution of 50cm. (raster file)

[10] Both comply with the definition of Digital Elevation Model (DEM).

- *Building Information Data & Infrastructure*
  The data set comprises point information (X, Y coordinates, .csv or .shp file) of (a) Education Facilities: Academies & Universities, Higher Education Schools, private and state Kindergarden, private and state Child Day-care, Secondary schools, Polytechnic schools, private and state After-school care centres, Special schools, Other private child care centers , Other private schools, Elementary schools, and (b) Municipal Facilities: Fire brigade locations, Cemeteries, Dog meadows, Churches, Playgrounds, Sports facilities, City tree register, Underground car parks, Drinking fountains, Pharmacy, Disabled parking spaces, Parking zones; as well as polyline information (.csv or .shp) of the Street List of Tyrol.
- *Urban Atlas 2012 - Innsbruck*
  Land cover and land use in Innsbruck with an MMU of 0.25 ha within the city and 1 ha in the countryside in 2006 and 2012; 17 urban classes and 10 rural classes; COPERNICUS Service (.shp)
- *API Interfaces*
  - Weather stations of Tyrol:
    Measured values and location information from the weather stations of HD Tyrol and LWD Tyrol. The data is updated hourly.
  - Precipitation data of HD Tyrol:
    The data record contains the measured values of the last 24 hours for the respective measuring point, including the name of the water body, location information and altitude reference. For further master data, see apps.tirol.gv.at/hydro Availability of data: minus 24 hours to current.
  - Water level data of HYDRO Tyrol:
    The data set contains the measured values of the last 24 hours for the respective measuring point including water body name, location information and elevation reference (water level zero point). For further master data, see Hydro Online at https://apps.tirol.gv.at/hydro/#/Wasserstand?station=201525 Period of the data provided: minus 24 hours to current.
  - Topographic Map:
    The topographic map of Tyrol and the alternative view of the aerial map depict the provincial territory and neighboring areas with high cartographic precision together with the geographical names.
- *Geodata-Services*
  As part of the Open Government Data (OGD) initiative, the province of Tyrol provides geodata services of the following types:
  - ArcGIS-Server: Visualization service for vector and raster data (various themes from TirisMaps)
  - WMS (Web Map Service): Raster data visualization service (orthophotos and terrain data)
  - WCS (Web Coverage Service): Download service for raster data (orthophotos and terrain data)
  - WFS (Web Feature Service): Visualization and download service for vector data
  - WMTS (Web Map Tile Service): Raster data tiled display service (base map of Tyrol)

The available themes are enlisted in the appendix Section 11.2.

Hydrographic Service (HD Tyrol)

Data by HD Tyrol such as Springs, Groundwater level, water level, precipitation is available by the Open-Source Datahub Austria (see above).

Sewer network of Innsbruck city (Innsbrucker Kommunalbetriebe, IKB)

- *General data set:*
  - Precipitation data
    Precipitation data for 7 different IKB measurement gauges from 2021, 2022 and 2023 for every minute (.csv).
  - Monitoring data sewer network
    Water Level of locations in the canals (manholes) from 2021, 2022 and 2023 for every minute (.csv)
  - General sewer network data
    - Blueprints of Mixed System Structures (.pdf)
    - Plans of sewer network
      - Site plan of catchment area (current status) (.dwg and .pdf)
      - Site plan of canal system and structures (current & 2016 status) (.dwg and .pdf)
      - Geo-network plan (current status) (.dwg)
      - Schematic plan of relief structures (.pdf)
      - Catchment areas (.shp)
    - Coordinates of measuring locations within the canal system (.xlsx)
    - Thresholds / Measuring reference of locations within the canal system (.xlsx)
- *Data from the Reference Event: Heavy Rain in Amras, 02.07.2016*
  - Rain data of 5 different IKB measurement gauges from 01.07.-04.07. for every minute (.csv).
  - Water levels of 5 different locations in the canals (manholes) from 01.07.-04.07. for every minute (.csv).
  - Water levels of 3 different locations of rivers and streams from 01.07.-04.07. for every minute (.csv).
  - Discharge / Inflow measurement of 2 locations from 01.07.-04.07. for every minute (.csv).
  - Discharge / Release Events of Mixed Water Structures of 01.07-.04.07 (.pdf, converted by DCNA into .xlsx)
  - Pump operation protocol of 4 flood pumping stations (.csv).

Emergency call protocols by Leitstelle Tirol

The emergency call protocols for the heavy rain event in 2016 in Innsbruck were provided by Leitstelle Tirol who receive all emergency calls in the federal state of Tyrol. The data consists of an excel list containing the time of the alert, deployment keyword, city, street, and xy-coordinates of location and was translated into English. Additionally, the Leitstelle provided all alerts in Innsbruck from January 2014 to August 2016.

Weather data of GeoSphere Austria (GSA)

Open-Source GSA Datahub (https://data.hub.geosphere.at/):

- Station data (API Interface)
  - Measuring stations ten-minute data v2
    This data set contains measurement data from 1992 to the present day in 10-minute resolution. The weather stations of GeoSphere Austria form the only comprehensive meteorological measurement network in Austria. They form the backbone of weather forecasting, climate products, climate research and dispersion calculations at GeoSphere Austria. The approximately 260 measuring stations cover all climatic regions and altitude levels in Austria. Most of them are semi-automatic weather stations (TAWES), which record the basic weather elements and transmit them in real time to the Hohe Warte in Vienna. There they are checked for quality and archived in databases. The majority of the measurement data in the dataset has been quality-checked since 2006. The quality status of individual parameters can be found in the quality flags with the name ending _flag. The data set is updated every 10 minutes based on the archived databases, whereby retroactive changes may also occur due to the data check.
- Areal data (API Interface)
  - INCA hourly data
    The INCA analysis and nowcasting system uses all available data sources - station observations, remote sensing data, numerical weather prediction models and a high-resolution terrain model - to provide the best possible analysis of the current state of the (near-surface) atmosphere. The combination of these data sources can best and most simply be interpreted as the correction of a gridded background field with observational data. INCA is used, among other things, in flood warning and forecasting and as a basis for internet portals with spatially and temporally detailed meteorological information. The analyses contained in this data set have a spatial resolution of 1 km x 1 km and a temporal resolution of 1 hour.
  - SPARTACUS v2.1 Daily data
    The gridded data set describes the spatial distribution of the observed air temperature (minimum temperature (TN) and maximum temperature (TX)), precipitation (RR) and absolute sunshine duration (SA) from day to day since 1961 with a resolution of one kilometre over Austria. In addition to the original daily resolution, monthly, seasonal and annual aggregates are also available. The data set is operationally expanded on a daily basis.
- Forecast data (API Interface)
  - Short-term forecast
    The short-term forecasts are based on the INCA analysis and nowcasting system. INCA uses data sources such as station observations, remote sensing data, numerical weather prediction models and a high-resolution terrain model to provide the best possible analysis of the current state of the (near-surface) atmosphere and short-term forecast. The combination of these data sources can best and most simply be interpreted as the correction of a gridded background field with observational data. The forecasts contained in this data set have a spatial resolution of 1 km x 1 km and a temporal resolution of 15 minutes with a forecast horizon of 3 hours.
  - Warnings
    The GeoSphere Austria warning system provides information to protect the population from potential weather hazards. In addition to meteorological information, such as the course of the warning: Warning progression, wind peaks, rainfall

amounts or snow depths, the possible effects of the expected weather situation are pointed out. The issued warning level (yellow, orange or red) is not exclusively based on the meteorological parameters, but also on factors that can intensify or weaken the effects (e.g. traffic volume, daytime or night-time temperatures): Traffic volume, time of day or season, vegetation status, previous pollution of a region). In addition, recommendations for action are given to minimise or prevent the damage to people, property and possessions caused by the expected warning situation as far as possible. The weather warnings of GeoSphere Austria apply to the permanent settlement area. The high alpine regions of Austria are not covered by a warning. However, the meteorological description of a warning can sometimes contain information about the situation in the high mountains.

Detailed data of a heavy rain event in Amras (Innsbruck) on 2nd of July in 2016:

- Precipitation data of 3 different measurement gauges in Innsbruck area in a 1-minute interval from 02.07.2016 12:00 to 23:00 (.xlsx file)
- INCA forecast data of 9 raster cells covering the Amras district in a 15-minutes interval from 02.07.2016 12:00 to 23:00 (.csv file)
- RAPIDINCA analysis data of 9 raster cells covering the Amras district in a 15-minute interval from 02.07.2016 00:00 to 23:00 (.csv file)
- ATNT warnings of the Innsbruck area from 02.07.2016 (.txt file)

Social Media Data

- *Heavy rain event - keywords:*
  Keywords relating to heavy rain events were brainstormed and defined with experts/stakeholders and were made available to T 4.5.
- *Manual analysis of twitter/X posts:*
  A manual analysis of twitter posts for the heavy rain event on the 2nd of July 2016 in Amras (district of Innsbruck) was carried out. The search window was selected from 2nd to 5th of July 2016 with the following keywords: Innsbruck, Tirol, Hochwasser, Überschwemmung, Starkregen, Unwetter, Gewitter, Sturm, Pradl, Amras, Reichenau, Aldrans, Ampass. In total, 30 relevant posts were found and summarised in an excel file which was made available to T4.5.