

Project Management Plan

CHIST-ERA-19-CES-005

SUSTAINABLE WATERSHED MANAGEMENT THROUGH INTERNET OF THINGS DRIVEN ARTIFICIAL INTELLIGENCE



This work was supported by the CHIST-ERA grant.

DOCUMENT RECORD

| Deliverable Title | D1.1 Project Management Plan |
|----------------------------|---|
| Corresponding Work Package | WP1 Project Management, Organization, and |
| | Dissemination |
| Related Tasks | T1.1 Project Management |
| Dissemination Level | Public |
| Due Submission Date | 01.06.2021 |
| Actual Submission Date | 01.12.2021 |
| Responsible Partner | University of Vienna |

| Version | Date | Author | Comments |
|---------|------------|-------------|------------------------|
| 0.1 | 28.05.2021 | Atakan Aral | First draft |
| 1.0 | 30.11.2021 | Atakan Aral | First complete version |
| | | | |
| | | | |
| | | | |







1. TABLE OF CONTENTS

| Docum | ent Record1 |
|---|--|
| 2. Exe | ecutive Summary |
| 3. Inti | oduction4 |
| 4. Pro | ject Overview5 |
| 4.1. | Project Objectives5 |
| 4.2. | Project Milestones5 |
| 5. Ma | nagement Structure |
| 5.1. | Executive Board6 |
| 5.2. | Advisory Board6 |
| 5.3. | Project Coordinator6 |
| 5.4. | Work Package Leaders7 |
| 6. Co | nmunication8 |
| 6.1. | Project Meetings8 |
| 6.2. | Correspondence8 |
| 63 | File Sharing |
| 0.0. | The Gharing |
| 7. Imp | blementation |
| 7. <i>Imp</i> 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. | Dementation 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 |
| 7. <i>Imp</i> 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. | Dementation 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 |
| 7. Imp 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. 7.3. | Demonstration 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 Deliverables 16 |
| 7. Imp 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. 7.3. 8. Do | Demonstration 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 Deliverables 16 Cumentation 17 |
| 7. Imp 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. 7.3. 8. Do 8.1. | olementation 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 Deliverables 16 cumentation 17 Deliverables 17 |
| 7. Imp 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. 7.3. 8. Dow 8.1. 8.2. | Dementation 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 Deliverables 16 cumentation 17 Deliverables 17 Meeting Minutes and Agenda 17 |
| 7. Imp 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. 7.3. 8. Dow 8.1. 8.2. ANNEX | Dementation 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 Deliverables 16 cumentation 17 Deliverables 17 Meeting Minutes and Agenda 17 A: Deliverable Report Template 18 |
| 7. Imp 7.1. 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. 7.3. 8. Dow 8.1. 8.2. ANNEX ANNEX | Deferentation 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 Deliverables 16 cumentation 17 Deliverables 17 Meeting Minutes and Agenda 17 B: Meeting Agenda Template 19 |
| 7. Imp 7.1. 7.1. 7.1. 7.1. 7.1. 7.2. 7.3. 8. Doi 8.1. 8.2. ANNEX ANNEX ANNEX | Delementation 10 Work Packages 10 1. WP1 – Project Management, Organization, and Dissemination 10 2. WP2 – Micropollutant Transport Model 10 3. WP3 – Deep Learning Framework for Anomaly Detection and Pollution Predictions 10 4. WP4 – Sustainable IoT Infrastructure 10 5. WP5 – Pilot Implementation and Validation 10 Tasks 10 Deliverables 16 cumentation 17 Deliverables 17 Meeting Minutes and Agenda 17 B: Meeting Agenda Template 19 C: Meeting Minutes Template 20 |







2. EXECUTIVE SUMMARY

This document is the deliverable D1.1 Project Management Plan of the CHIST-ERA-19-CES-005 project SWAIN - Sustainable Watershed Management Through Internet of Things Driven Artificial Intelligence (referred to as SWAIN in the rest of the document). The purpose of the SWAIN Project Management Plan is to provide guidance regarding the management, monitoring, and controlling processes throughout the project. Therefore, it includes:

- Roles and responsibilities of each partner
- Boards and their responsibilities
- Procedures for meetings and communication
- Deliverables of SWAIN
- Reporting and documentation

This document is intended for the SWAIN consortium partner. It complements the previously signed consortium agreement.







3. INTRODUCTION

The project management plan is a deliverable in the scope of Work Package 1 of the SWAIN project. It aims to facilitate project operation by defining roles and responsabilities of partners and boards. It also defines the project communication strategy through meetings, mailing lists, repositories, and reports. The ultimate goal of this plan is to ensure that the project is delivered according to the promises in the full project proposal that is selected for funding.

This plan is prepared based on previous documents and discussions as listed below.

- SWAIN project full proposal, 02.07.2020
- SWAIN consortium agreement, 11.03.2021
- SWAIN project kick-off meeting, 01-02.03.2021
- SWAIN project workshop, 15.04.2021

The project management plan is intended as a live document. It will be updated and revised as needed. The new version as well as the changes they bring will be clearly listed on the document report on Page 1. The project coordinator will be responsible for coordinating and implementing such revisions.

The rest of the project management plan presents an overview of the project objectives and milestones (Section 3), the management structure including the management boards and roles (Section 4), communication strategy in terms of project meetings, correspondence, and file sharing (Section 5), implementation details of work packages, tasks, and deliverables (Section 6), and the documentation of deliverables and meetings (Section 7) along with the corresponding templates (Annexes).









4. PROJECT OVERVIEW

4.1. PROJECT OBJECTIVES

Towards our general goal of European-wide deployment of AI-enhanced micropollutant detection systems, we identify the following objectives.

OBJ1. We will geo-locate the micropollutant sources and estimate the discharge amount with unprecedented accuracy.

OBJ2. We will achieve continuous adaptation to environmental changes using timely watershed data from previously excluded remote areas.

OBJ3. We will design and demonstrate the first AI-based early warning and prediction system for domestic, industrial, and agricultural pollution in European watersheds.

4.2. PROJECT MILESTONES

To realize the project objectives, the following milestones are defined.

M1. Micropollutant detection model – This milestone is aimed to be achieved by month 18 based on the contributions from work packages 2 and 3.

M2. IoT infrastructure deployment plan – This milestone is aimed to be achieved by month 28 based on the contributions from work package 4.

M3. Micropollutant decision support software – This milestone is aimed to be achieved by month 36 based on the contributions from work packages 2 and 5.







5. MANAGEMENT STRUCTURE

5.1. EXECUTIVE BOARD

The executive board (EB, also called The General Assembly in the consortium aggrement) is the decision-making body of the project and consists of the coordinator and the national PIs. The chairperson is the coordinator. This body decides about the progress of the project, the scientific work plan, scientific outcomings as well as all organizational and financial matters. Regular virtual meetings of the EB will be organized at least monthly.

Every member of the EB should be present or represented at any meeting; may appoint a substitute or a proxy to attend and vote at any meeting; and shall participate in a cooperative manner in the meetings.

The EB shall not deliberate and decide validly unless two-thirds (2/3) of its Members are present or represented (quorum). Decisions shall be taken by a majority of two-thirds (2/3) of the votes cast. The chairperson shall produce written minutes of each meeting which shall be the formal record of all decisions taken.

5.2. ADVISORY BOARD

The Advisory Board (AB) consists of at most three outstanding, internationally renowned scientists. The AB meets at least once a year in the course of one EB meeting. The tasks of this body are scientific consulting, evaluation, and peer-reviewing of the scientific outcome. The EB has elected the following initial members of the advisory board after the project kick-off meeting.

- Prof. Ewa Deelman, University of Southern California
- Prof. Ching-Hua Huang, Georgia Institute of Technology
- Prof. Hans Vangheluwe, University of Antwerp / McGill University

5.3. PROJECT COORDINATOR

The Coordinator shall be the intermediary between the Parties and the Funding Authority and shall perform all tasks assigned to it as described in this Consortium Agreement. In particular, the Coordinator shall be responsible for:

- monitoring compliance by the Parties with their obligations,
- keeping the address list of Members and other contact persons updated and available,
- collecting, reviewing to verify consistency and submitting reports, other deliverables and specific requested documents to the Funding Authority,
- preparing the meetings, proposing decisions and preparing the agenda of General Assembly meetings, chairing the meetings, preparing the minutes of the meetings and monitoring the implementation of decisions taken at meetings,







- transmitting promptly documents and information connected with the Project to any other Party concerned,
- providing, upon request, the Parties with official copies or originals of documents that are in the sole possession of the Coordinator when such copies or originals are necessary for the Parties to present claims.

5.4. WORK PACKAGE LEADERS

Work package leaders are responsible for the coordination of the scientific work described in the project proposal. The following persons are the designated work package leaders.

- WP1: Project Management, Organization, and Dissemination Atakan Aral (UNIVIE)
- WP2: Micropollutant transport model Janne Juntunen (SYKE)
- WP3: Deep Learning Framework for Anomaly Detection and Pollution Predictions – Slobodan Lukovic (USI)
- WP4: Sustainable IoT infrastructure Vincenzo De Maio (TUWIEN)
- WP5: Pilot implementation and validation Mehmet Tahir Sandikkaya (ITU)







6. COMMUNICATION

6.1. PROJECT MEETINGS

In order to enable collaboration and communication between the project partners, monthly meetings are organized. National PI's are expected to attend these meetings. Additionally, project personnel from all partners are invited to participate the project meetings. Younger researchers such as master's and doctoral students are particularly encouraged. The following monthly meetings are organized in 2021. The meetings marked with and asterisk (*) are co-located with other project-related events.

- 1. March SWAIN Partner Meeting*, 01.03.2021 02.03.2021
- 2. April SWAIN Partner Meeting*, 15.04.2021
- 3. May SWAIN Partner Meeting, 07.05.2021
- 4. June-July SWAIN Partner Meeting, 10.06.2021
- 5. August-September SWAIN Partner Meeting, 14.09.2021
- 6. October SWAIN Partner Meeting*, 18.10.2021 19.10.2021
- 7. November-December SWAIN Partner Meeting, 02.12.2021

The following three events were held co-located with the March, April, and October project meetings.

- 1. SWAIN Kick-Off Meeting, 01.03.2021 02.03.2021
- 2. SWAIN Project Workshop, 15.04.2021
- 3. SWAIN Yearly Face-to-Face Meeting Vienna, 18.10.2021 19.10.2021

Two task forces were identified in the SWAIN Project Workshop. These task forces had their internal meetings more frequently depending on work progress. Due to the global pandemic conditions, all monthly meetings except the October meeting are held in the virtual environment via teleconferencing.

6.2. CORRESPONDENCE

E-mail is the main form of correspondence during the SWAIN project. For discussions concerning the whole consortium, a mailing list swain@lists.univie.ac.at has been formed to that end. E-mails sent to this list is delivered to all members (currently 26) and the tag [SWAIN] is automatically added to the subject field.

For bilateral or multilateral conversations that are not concerning the whole consortium such as discussions within task forces, regular e-mails are used. The manual addition of [SWAIN] tag is highly recommended for such e-mails so that recipients can label or prioritize them according to their needs.

6.3. FILE SHARING

The following services are used for file sharing. More details are given in the D1.6 Data Management Plan.





| Source code | GitHub (https://github.com/SWAIN-Project) |
|-------------------------------|---|
| Data sets (In preparation) | Google Drive (https://drive.google.com/drive/fold- |
| | ers/1_07bdcec-ec0tWYrHeLfkq44v_uf4U7B) |
| Data sets (Ready, public) | Zenodo |
| Publications (In preparation) | Overleaf |
| Publications (Published) | Zenodo, Project Website (<u>http://swain-project.eu/</u>) |







7. IMPLEMENTATION

7.1. WORK PACKAGES

7.1.1. WP1 – PROJECT MANAGEMENT, ORGANIZATION, AND DISSEMINATION

The purpose of this WP is to ensure the high-quality results in SWAIN, through the continuous monitoring of the execution of the project activities, milestones, and deliverables, safeguarding their proper execution according to the project's work-plan, while optimizing the collaboration among the partners. This WP also ensures the project's success during and beyond its implementation in terms of dissemination, standardization, exploitation, and sustainability by communicating results, representing the project, and eliciting requirements and input from stakeholders.

7.1.2. WP2 – MICROPOLLUTANT TRANSPORT MODEL

In WP2, we will develop a hybrid contaminant transport model for a watershed, which simulates the concentration distribution of source-specific (fingerprint) micropollutants along the river. The model does not rely on continuous execution and works with a minimum amount of variables, selected via AI.

7.1.3. WP3 – DEEP LEARNING FRAMEWORK FOR ANOMALY DETECTION AND POLLUTION PREDICTIONS

In this WP we look to build efficient architectures to train and learn needed knowledge on previously structured datasets. We look into changes detections as well as into predictions of future trends in short and long terms. The lifelong adaptation of the models will be investigated and proposed.

7.1.4. WP4 – SUSTAINABLE IOT INFRASTRUCTURE

The aim of this work package is the design of a fault-tolerant and energy-efficient measurement infrastructure.

7.1.5. WP5 – PILOT IMPLEMENTATION AND VALIDATION

A software, which integrates the components developed under previous work packages, such as contaminant transport model or AI model, will be designed and tested for in situ data acquisition, computation, and visualization to simulate the early warning system.

7.2. TASKS

| T1 1 | Project Management (M1-M36: responsible: 1) |
|------|--|
| | |
| | |
| | This task deals with all necessary mechanisms and structures for the management and |
| | administrative coordination of the project. It incorporates Administration Management |
| | activities, including procedures and guidelines for activity planning and update, cost and |
| | time management, submission of periodic progress reports and cost statements, prep- |
| | aration of annual review reports and review presentations, and submission of delivera- |
| | bles to the CHIST-ERA and national funding agencies. |
| 1 | |







| T1.2 | Quality Assurance & Risk Assessment (M1-M36; responsible 1) |
|------|---|
| | The task deals with defining and specifying the appropriate mechanisms and processes that will be established in order to maintain a certain quality level in the whole project structure and outcomes. Unexpected compromises to the quality of the results are also guaranteed through the identification, monitoring, and management of potential project risks, together with the definition and application of contingency plans. |
| T1.3 | Dissemination and Exploitation (M1-M36; responsible: 1; involved: 2,3,4) |
| | This task deals with the planning and execution of the activities described in Section 2.2. One partner per country is responsible for national dissemination and exploitation activities whereas TUW is responsible for the overall coordination of them. |
| T1.4 | Reporting (M1-M36; responsible: 1) |
| | The project coordinator will act as the point of contact for partners in communications with the CHIST-ERA. The coordinator will ensure annual reporting, semi-annual technical internal reporting, milestone review, midterm review after 12 months. |
| T2.1 | Analysis of the factors relating to micropollutant transport and presence in watersheds (M1–M9; responsible:4; involved: 2,3) |
| | The factors affecting the micropollutant transport could vary for different watersheds. In order to provide a model that is generally applicable to European rivers, it is important to analyze the factors from different environments. We will first extract such factors from open data sets by various European countries (e.g. NORMAN data set repository). Then, we will combine these with the findings in the current state-of-the-art to find out existing relations between the occurrence of MPs and other water quality variables and characteristics of catchments. Usually, it is accepted that correlation itself is enough to prove existing relations but we strengthen requirements by demanding process-based description for observed relations. Finally, we will measure the extent of the relations on the data collected from our use case watersheds. We also take into account the properties of different kinds of micropollutants. We will analyze the data provided by IST and SYKE in the first iteration. In the second iteration instead, the more granular data from WP3 will be utilized to improve the model. |
| | This task also reduces the search space AI-based source identification by choosing the most relevant factors. Therefore, it renders near-real-time decision making possible in the operative system. |
| T2.2 | Refinement and update of available data and consolidation of a single geodatabase (M1–M9; responsible: 3; involved: 2,3,4) |
| | Our main study areas will be the Ergene River and River Kokemäenjoki. Ergene River is a 283 km long river spanning the European part of Turkey from northeast to southwest. The complete river network including the mainstream and the tributaries reaches 6188 km of length. We have all geographical, hydrological, and meteorological data of the watershed between 2012-2016. River cross-sectional profile, flow velocity distribution and flow rate at 40 points on the river are available. We also collated substantial |







data on industrial facilities on the watershed and their activities; animal farms and their livestock capacities; agricultural fields, their crop spectrum, and the pesticides used for each crop. We have also measured 222 micropollutants and conventional water pollution parameters over a year at 75 different locations on the Ergene River. Under this task, all necessary data found critical in task 2.1 and available on the watershed will be updated and consolidated in a single geodatabase.

We will separate the collected data in several axes; (1) spatial, (2) temporal, and (3) spatio-temporal. The collected data will be independently divided into chunks to represent each axis. An arbitrary chunk in the selected axis is used for reasoning and training of the AI-models. This leads to a set of forecasts on each axis with several training data. Each of the forecasts will be analyzed with the actual data to observe the correlation in between.

In addition, the micropollutant measurement dataset may be enhanced by additional non-target analysis of stored samples using liquid chromatography and high-resolution mass spectroscopy (LC-HRMS) to expand the number of micropollutants in the samples. The expanded set of micropollutants may enhance the precision of attaining more specific micropollutants to the potential sources using AI.

T2.3 Design and implementation of Al-assisted micropollutant transport model (M6 – M18; responsible: 2; involved: 1,3,4)

We find optimal ways to obtain flows for the whole river with high spatial and temporal accuracy. We use information gathered in task 2.2 to implement a water quality module (WQM) that is satisfactory for our purposes. This WQM includes methods to derive preliminary possible sources for given micropollutants that create restrictions on optimization problems encountered when finding sources of pollution.

Due to the high nonlinearity and complexity of micropollutants spread, we introduce spatio-temporal graph networks. We utilize this kind of based on graphs to better capture interdependencies among various factors to effectively capture temporal trends. We will develop a novel adaptive dependency matrix and learn it through node embeddings. These models will be an input for graph neural networks that will be built to efficiently detect changes as well as predict pollution spread.

We aim to combine purely data-driven approaches to deterministic models to obtain a hybrid modeling scheme. To obtain this we will investigate the feasibility of novel graphbased representation and learning. Starting from such functional graph representation, AI techniques can more efficiently solve regression and prediction tasks with respect to micropollutant spread.

T3.1 Anomaly detection (M12 - M24; responsible: 2; involved: 1)

On top of the graph-based models, we build a machine learning framework. We will investigate different deep architectures as also proposed in the literature (i.e., those based on recurrent networks, convolutional architectures, and hybrid solutions). As the first step fingerprinting of the micropollutants will be done based on reference models







| | developed in WP2. We aim at effectively capturing comprehensive spatio-temporal cor- relations so that changes in micropollutant spread are efficiently detected. |
|------|---|
| T3.2 | Model Hybridization (M24 - M30; responsible: 1; involved: 3,4) |
| | One of the major breakthroughs of SWAIN lies in that it does not solely rely on expert- and data-driven modeling, both of which have their shortcomings. Instead, it integrates the expert-driven physical environment knowledge and data-driven, evidence-based AI. To that end, we will utilize transfer learning techniques to store knowledge gained from environmental experts and apply it to improve the data-driven model. Additionally, we will develop a feedback mechanism through expert elicitation techniques to label col- lected data with minimum human intervention. The synthesis of evidence-based data and knowledge from environment experts can be used for training or updating the hybrid model. |
| T3.3 | Model Adaptation (M24 - M30; responsible: 2; involved: 1,4) |
| | We will consider the non-stationarity of the problem by developing an architecture that allows the mapping from field measurement devices onto nodes to adapt over time. We will study existing spatio-temporal graph neural network architecture in the context of pollution spread forecasting and assess whether the graph representation leads to im- proved predictions in the context of short-term pollution forecasting. In particular, we will study predictors that use a predefined similarity measure to model the relationships be- tween nodes, as well as models where the relationships, i.e., the edge attributes, are learned end-to-end from data. |
| T3.4 | Near-real-time data quality control (M30 – M36; responsible: 4; involved: 1,2) |
| | This task sets common strategies for near-real-time data quality control algorithms for the (sensor) data delivered in the project. Overall data quality plan will be developed. Several automatic quality control tests which identify (and flag/remove) erroneous or suspicious measurements will be implemented. Information on noise, exceeding pre- determined parameter thresholds, missing data, and lack of variation will be calculated using data-driven methods and machine learning algorithms. The main purpose of im- plementing near-real-time data quality control algorithms is to ensure good quality of the data that will be used in model calibration and validation. |
| T4.1 | Design of placement algorithm (M3 – M9; responsible: 1; involved: 2,3,4) |
| | Starting from TUW work on edge provisioning, we will design an algorithm for the place- ment of data aggregation points among the available measurement stations. We will select different metrics for resilience, latency, and energy efficiency existing in literature, identifying the most suited to our purpose. Since the goal is to identify a trade-off be- tween these three metrics, we will investigate different optimization methods, ranging from multi-objective optimization to constrained optimization. The algorithm will be eval- uated using data collected by USI, IST, and SYKE, in order to assess its performance in different scenarios, as well as the ontology developed during FEnCY project. |







| T4.2 | Deployment of heterogeneous communication infrastructure (M6 – M15; responsible: 1; involved: 2) |
|------|--|
| | The goal of this WP is to investigate different transmission technologies for the selected scenario to identify a trade-off between QoS and costs. Due to the unreliable connections available in the selected scenario, we will develop methods to predict both connectivity and different quality of service parameters (i.e. latency, bandwidth, throughput) which will allow us to select the most suitable transmission technology according to the system state and our budget. To simulate different scenarios, we will employ connectivity traces collected in different environments. To this end, we will adapt the methodology proposed by TUW and USI for evaluation of the energy efficiency of networking to measure QoS metrics. Based on the predictions obtained by using connectivity traces, we will design an adaptive method to select the most suitable communication channel according to the availability, cost QoS. |
| T4.3 | Design of viscoelastic adaptation algorithms (M18 – M28; responsible: 1; involved: 2) |
| | The goal of this WP is to develop different dynamic scheduling algorithms to dynamically adapt to the infrastructure state. Dynamic adaptation algorithms will consider different parameters, such as resilience, runtime, and load balancing. Due to the limited resources available at the edge, as well as the necessity of dynamically moving computation according to distance from data aggregation points and latency constraints, we will deploy the computational logic of our system in functions, adopting well-known Function as a Service (FaaS) frameworks such as OpenWhisk and OpenFaaS. Concerning the dynamic scheduling of functions, we will investigate methods based on control theory. |
| T5.1 | Short-term experimental verification of AI and WQM model outputs (M6 – M24: responsible: 3; involved: 4) |
| | Al will supply a list of fingerprints micropollutants for each potential pollution source using historical data that we have collected between Aug-2017 and May-2018 and con- solidated to a single database in T2.2. These pollutants will be verified by measuring those pollutants in the composite samples collected each day for a week from the pol- lution sources. The model will predict the flow rate and concentration of fingerprint mi- cropollutants along the river. Both flow rate and micropollutant concentrations will be measured in the composite samples collected each day for a week at a location close to the middle of the river (i.e. Luleburgaz). |
| T5.2 | Controlled experiments to demonstrate the feasibility of the proposed techniques (M6 – M18: responsible: 3; involved: 1,2) |
| | After the fingerprint pollutants are determined in T5.1, the effectiveness of these finger- prints will be verified by observations. The observations are three-fold. In the first step, the available environmental data will be verified. As illustrated in section 1.1, the direct fingerprint measurement might not be possible. The second step is executed to over- come this. Manual samples are collected from selected locations (see T4.1). The sam- ples are analyzed and checked if the results confirm the expected fingerprint presence and concentration. In the third step, the WQM from T2.3 will be utilized in the software |







| | implementation. This step leads to a set of pollution forecasts; therefore availability to locate pollution sources. This approach will have the additional benefit of executing of- fline simulations and what-if analysis. Note that near-real-time forecasting might require bulky equipment (e.g. LC-MS) to be located in close proximity to measurement stations. This might not be possible in practice (see section 1.1). Therefore, the samples could be transferred to the laboratory, analyzed, and analysis data could be compiled with the complementary data of the relevant time span to carry out the forecasts. As a result, validation of the pilot infrastructure will be carried out by collecting, computing, and transferring near-real-time data as well as simulated pollutant analysis. So that, the efficiency of the loT/edge infrastructure could be observed with the available near-real-time data together with enriched simulation data. The optimized sampling locations and determined fingerprints are expected to be provided by TUW and the WQM is expected to be provided by USI. |
|------|--|
| T5.3 | Validation of created approach (M30 – M36; responsible: 4; involved: 1,2,3) We will apply methods created in WP2-4 to a completely different geographical location that has a different climate and structure of watershed to validate the generality of the proposed approaches. SYKE will provide concentrations measured at 13 locations alongside river Kokemäenjoki and released pathogens and chemical compounds from 10 WTPs. We will also compare the proposed watershed model to SYKE's nationwide operational model of Kokemäenjoki and others if possible. To prove the applicability of the hybrid modeling approach we use only data found essential to source identification. In addition, we compare results to the existing SOBEk model for the river. |
| T5.4 | Pollutant visualization tool (M18 – M36; responsible: 3; involved: 1,2,4) This task aims to bring together the models and algorithms that are implemented throughout the project and present the results via an interactive web-based application. With the help of this application, users can perform the following key tasks for the pilot basins: Visualize the pollutant model built during the project Visualize the data in near-real-time, collected by the compatible data sources Select the historic data and selectively train AI models implemented during the project Tweak the following parameters and compare different outcomes side-by-side Parameters of the pollutant model such as the relevant micropollutants, placement of the sampling stations, and sampling frequency. Parameters of the AI model such as model type, training set size, objective values, and performance goals. Data sources and data schema used in the system such as measurement types, data source types, data reliability objectives, and historic data size. Predict and project the possible status of the basin in an adjustable way. This feature will present the user with an adjustable visual simulation of time travel over the river basin with selected parameters. |







The output software is aimed to be a web-based application for the ease of access via a web browser across different systems and devices.

7.3. DELIVERABLES

| Deliverable | Month of delivery | Title of deliverable | Type of Doc- umentation | |
|-------------|----------------------|--|----------------------------|--|
| D1.1 | 3 | Project Management Plan | Report | |
| D1.2-D1.4 | 12, 24, 36 | Annual Public Reports | Report | |
| D1.5 | 6 | Quality Assurance & Risk Assessment Plan | Report | |
| D1.6 | 6 | Data Management Plan | Report | |
| D1.7 | 3 | Project Website and Promotional Materials | Code | |
| D1.8 | 3 | Dissemination and Exploitation Plan | Report | |
| D2.1 | 8 | Hydrogeographical and chemical parameters for WQM | Report | |
| D2.2 | 9 | Data consolidated to a single database | Database | |
| D2.3 | 24 | Method of combining data-driven and deterministic | Code | |
| | | modeling, a beta version of the program | | |
| D3.1 | 24 | Deep learning-based change detection module | Code | |
| D3.2 | 30 | Model adaptation method | Code | |
| D3.3 | 36 | Data quality plan | Report | |
| D4.1 | 5 | Placement model | Report | |
| D4.2 | 9 | Offline placement algorithm | Code | |
| D4.3 | 10 | Models for communication technologies | Report | |
| D4.4 | 15 | Adaptive communication channel selection method | Code | |
| D4.5 | 20 | Dynamic scheduling algorithm | Code | |
| D5.1 | 18 | In situ verification of contaminant transport model Report predictions | | |
| D5.2 | 24 | Optimized field data acquisition and processing for Report model simulations | | |
| D5.3 | 36 | Web based decision support application Code | | |







8. DOCUMENTATION

8.1. DELIVERABLES

Deliverables are categorized into three categories. Databases are published in Zenodo and Codes are published in GitHub. Finally, reports can be delivered as technical reports or scientific publications. For technical reports a template is provided (Annex A). The partners have agreed in principle that scientific publications can replace the report-based deliverables under the following conditions:

- If the publication is not yet published at the deliverable deadline, we should provide a preprint (e.g., in Arxiv or Zenodo).
- After the article is published, it should be openly accessible from the project website. If available, we can upload an extended technical report version, too.
- Since all tasks are defined as joint efforts by several project partners, the publication should include these partners as co-authors.

8.2. MEETING MINUTES AND AGENDA

Templates for the meeting agenda and minutes are provided in Annexes B and C, respectively.







ANNEX A: DELIVERABLE REPORT TEMPLATE









ANNEX B: MEETING AGENDA TEMPLATE

| Ŧ | SWAII | N MEETII | | da templ | ATE | |
|---|----------------------------|------------------------------------|------------------------|-------------|-----------|--|
| | Date: Time: | (AT, CH) (FI, TR) | | | | |
| | | | | (e | chist-era | |
| | Agenda I | items (All time | es in CEST) | | | |
| | | | | | | |
| | Additional Calendar inv | information vites with the Zoon | n links will be sent s | separately. | | |







ANNEX C: MEETING MINUTES TEMPLATE

| chist-era |
|-------------------------|
| SWAIN Project X Meeting |
| Date |
| Attendees |
| 1. Name (Partner) |
| Decisions / Tasks |
| 1. |
| |







ANNEX D: PRESENTATION TEMPLATE







