ReSilence

Retune the Soundscape of future cities through art and science collaboration

HORIZON- 101070278

D5.1

Roadmap towards the implementation of ReSilence toolkit

Dissemination level:	Public
Contractual date of delivery:	Month 10, 30/09/2023
Actual date of delivery:	Month 12, 7/11/2023
Work package:	WP5 Art-driven experimentation toolkit development
Task:	T5.1 Technical requirements and tools architecture T5.2 Integration of services at component level
Type:	Report
Approval Status:	Approved
Version:	0.5
Number of pages:	38
Filename:	d5.1_Roadmap towards the implementation of ReSilence toolkit _31102023_v0.5.docx

Abstract

This document is linked to project milestone MS1, signifying the successful project setup and

the creation of a roadmap for ReSilence's toolkit implementation. It provides an overview of its contents, outlines the toolkit's architecture, describes development modules, discusses integration strategies, presents pilot use cases, and offers a project timeline. This roadmap serves as a technical agreement among partners, guiding the development and integration of ReSilence components.

co-funded by the European Union

History

Version	Date	Reason	Revised by
V0.1	30/07/2023	Creation of ToC and initial content - Distribution to the ReSilence consortium	Nefeli Georgakopoulou
V0.2	29/09/2023	Submitted for internal review	Aleksandr Shvets
V0.3	4/10/2023	Revised after internal review by UPF	Aleksandr Shvets
V0.4	18/10/2023	Minor changes, formatting figures, tables and adding footnotes	All authors
V0.5	31/10/2023	Final Version completion	Nefeli Georgakopoulou

Author list

Organization	Name	Contact Information
CERTH	Sotiris Diplaris	diplaris@iti.gr
CERTH	Nefeli Georgakopoulou	nefeli.valeria@iti.gr
UNIGE	Antonio Camurri	antonio.camurri@unige.it
UPF	Aleksandr Shvets	alexander.shvets@upf.edu
MU	Beatrice de Gelder	b.degelder@maastrichtuniversity.nl
MPIEA	Melanie Wald-Fuhrmann	melanie.wald-fuhrmann@ae.mpg.de

Executive Summary

This deliverable is connected to the project milestone MS1, “Project set up and toolkit implementation roadmap”. This milestone marks the successful set up of the project, identifying the design needs and challenges and building the roadmap for the implementation of ReSilence’s toolkit.

Initially, an overview of the document's contents is provided. Subsequently, a high-level view into the prospective architecture of the ReSilence toolkit is provided. An in-depth description of the modules designated for development within the ReSilence framework follows. These insights encompass the functionality and specifications for each system component, along with a comprehensive breakdown of the necessary resources and corresponding development milestones.

To ensure the smooth integration of ReSilence services, a description of the available tools, practices, and integration patterns is provided. Additionally, a comprehensive contextualization of the ReSilence system is included, offering a well-rounded perspective. Furthermore, the document includes a perspective on the proposed pilot use cases, delving into the motivations behind each scenario and offering detailed descriptions. Lastly, the document presents a comprehensive project timeline that delineates scheduled iterations of the ReSilence toolkit and the expected levels of functionality at significant project milestones.

This roadmap signifies a shared understanding and technical agreement among the partners responsible for the development and deployment of services and components within the ReSilence toolkit. It establishes the framework for implementing the toolkit, addresses critical considerations related to this type of development. Above all, it serves as a guiding compass for the development and integration of any component within the ReSilence ecosystem.

Abbreviations and Acronyms

AI	Artificial Intelligence
API(s)	Application Programming Interface(s)
CSV	Comma Separated Values
DB	Data Base
EEG	Electroencephalography
JSON	JavaScript Object Notation
OC	Open Call
PUC	Pilot Use Case
SDK	Software Development Kit
VR	Virtual Reality
WP	Work Package
XR	Extended Reality
XML	eXtensible Markup Language

Table of Contents

1 INTRODUCTION

This document outlines the roadmap and details the strategy for the development of the ReSilence toolkit. The primary objective of this deliverable is to specify in detail:

(i) The evolving functionalities of ReSilence modules and activities (in particular at the time points of the milestones), (ii) the integration process required to ensure seamless module functionality (iii) the resource requirements essential for achieving these functionalities (iv) and the technical and research methodologies to be employed.

To accomplish this, an iterative development model was adopted, and the roadmap encompasses the following key components:

I. A high-level conceptual architecture, providing an overarching view of the system's structure.

II. A detailed description of the functionality of each technical module, including: a summarised scientific and technical objective, along with the supporting technologies. The progressive evolution of functionality over time and a timeline featuring schedules and resource allocation.

III. The strategy for system integration.

IV. A comprehensive description of the proposed pilot use cases, including their motivations and scenarios per use case.

V. An overall project timeline outlining the planned iterations of the ReSilence toolkit and the anticipated functionality levels at project milestones.

To maintain clarity and structure, this deliverable is organised as follows: Section 2 provides an overview of the high-level architecture, while Section 3 offers functional descriptions of the modules, aligning with the project work packages (WPs). Section 4 addresses the intricacies of system integration. Section 5 presents the development cycle and introduces the pilot use cases. Finally, Section 6 presents the overall project timeline, highlighting significant milestones while Section 7 offers a concise conclusion to this deliverable.

2 RESILENCE ARCHITECTURE

The ReSilence toolkit is a collaborative system that unites various concerns from different fields and practices. Its purpose is to assist artists in planning their interventions in both private and public spaces. To achieve this goal, the toolkit gathers data from diverse sources, establishing connections and obtaining knowledge about the discourses related to the project's concerns. In essence, it aims to investigate the boundary between noise and music within a changing world, fostering an acoustic awareness in urban environments. This involves not only reducing noise intensity but also recognizing it as an energy source and creating positive sounds. The intention is to preserve and multiply sounds while generating novel sonic experiences in cities. These experiences expand opportunities for accessibility, active participation, sustainability, and establish trust in Artificial Intelligence (AI) and Extended Reality (XR) technologies. For instance, the toolkit utilises Electroencephalography (EEG) analysis to examine emotions evoked by specific soundscapes, employs machine learning algorithms and computer vision techniques to translate sounds and images into visual representations, and features web-based applications accessible on smartphones and tablets, which enable individuals experiencing sound in specific moments to report their emotional responses, among other functionalities (see Figure 1).

This ecosystem of concerns that the platform brings together is represented in Figure 1. In this chapter, we introduce the preliminary design of the ReSilience architecture. This design brings together different technologies and practices and connects them in a coherent manner that corresponds to the expected role and functionality of advanced technologies in supporting the design of artistic interventions that address a variety of challenges allocated in specific Pilot Use Cases (PUCs).

In order to create this architecture model, a bottom-up approach was followed, starting from identifying and describing the elemental modules that form the basic services of a service-oriented-architecture. Following this, a toolbox facilitates the separation of concerns in order to consolidate each main aspect of the architecture. The technologies were analysed from the viewpoint of deployment, or in other words, under the envisioned usage scenarios. The knowledge and information gathered facilitated the conceptualization of the first version of the architecture, which was consequently described through its main pipelines in this chapter.

Images Source: https://www.freepik.com/

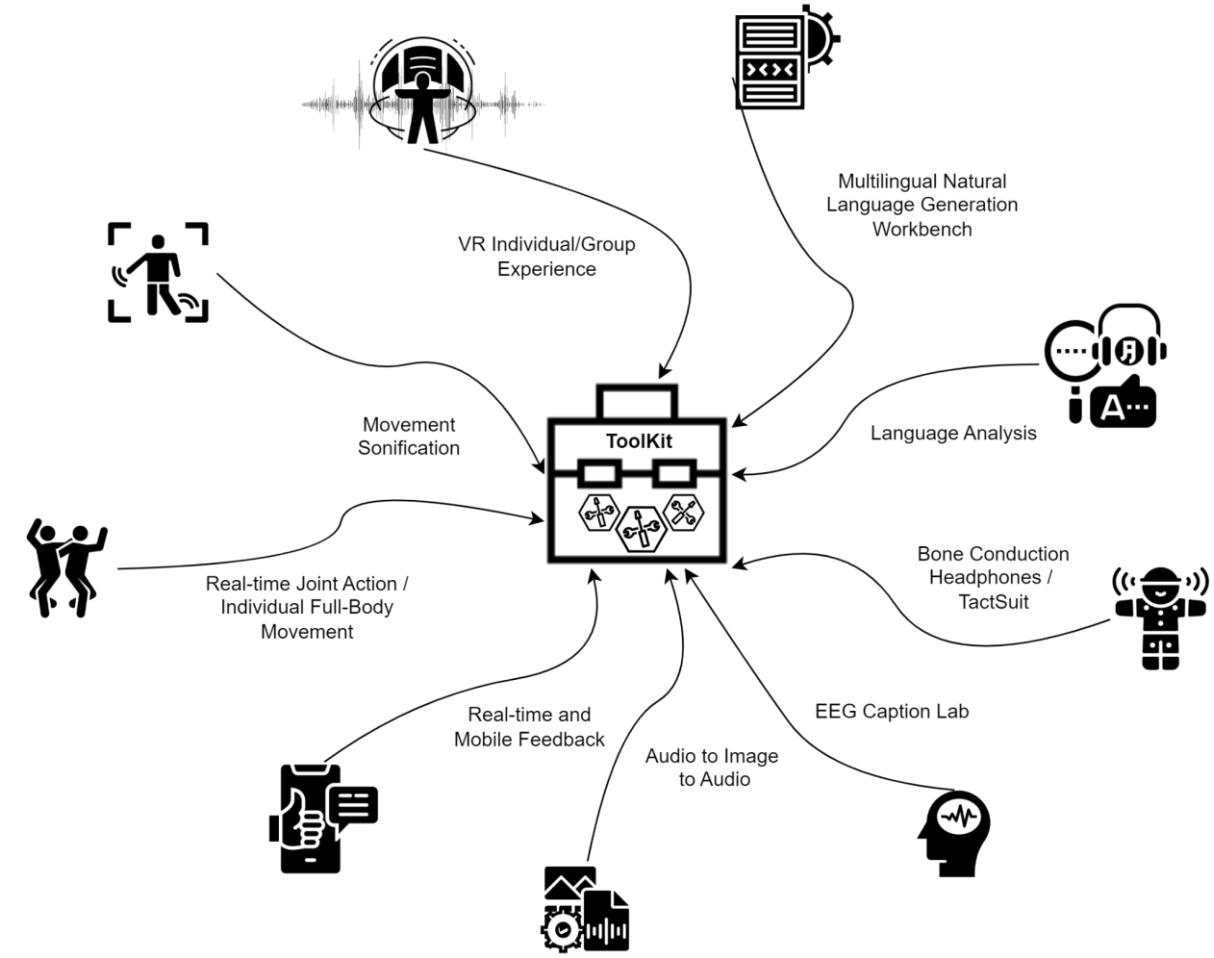

Figure 1: ReSilence Toolkit Overview

2.1 **Methodology for designing the architecture**

In order to create the architecture for ReSilence, we have developed a bottom-up methodology. This method entails first defining the system's individual elements and subsequently establishing connections and linkages among them to form a unified and integrated framework.

A detailed depiction of the ReSilence toolkit architecture is shown in Figure 2. The architecture design of ReSilience is centred around providing a robust and accessible platform for users and artists. Each service within the project is hosted on a server, which also includes its own database if necessary. The key focus is to ensure that the services are easily accessible through API endpoints, providing a seamless interface for users and artists to interact with.

By following this architecture, Resilience aims to establish a reliable infrastructure that allows users and artists to leverage the project's services efficiently. The server-based approach ensures that each service operates independently, promoting scalability and fault tolerance. Additionally, the provision of API endpoints enables integration with external systems or applications, allowing for expanded functionality and collaboration.

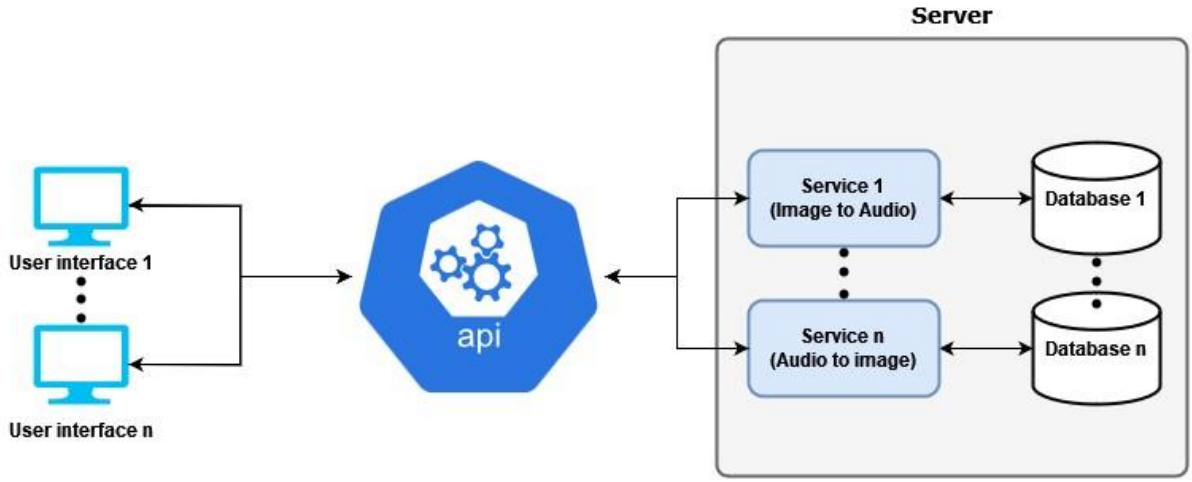

Figure 2: ReSilence Architecture

3 FUNCTIONAL DESCRIPTION OF MODULES

All the individual modules, each based on the design needs from deliverable D2.1 “Design needs and challenges in orchestrating the future sounds and experiences of cities”, are present in this chapter. The requirements of the modules are presented in detail in deliverable D6.1 “Pilot use cases and initial requirements and challenges”.

3.1 Crawling and Scraping service

The social media and web data crawling/scraping service has been tasked with the extraction of publicly accessible textual, visual and audio content from a variety of open web resources, including social media platforms. A range of mechanisms will be deployed to accommodate different source categories. Specifically, APIs from social media platforms like YouTube can be integrated into the service. Meanwhile, for online media sources such as blogs and web pages, a combination of crawling and scraping techniques can be applied. Importantly all the generated content will be available to other modules providing them with relevant input data. This service is undertaken under the task T4.2.

The properties of the crawling and scraping service are included in the following Table.

INPUT(S)	Text and multimedia content
OUTPUT(S)	Json serialisation of harvested items
PROGRAMMING LANGUAGES/TOOLS	Java, Python
INTEGRATION	This service will communicate with a data storage service responsible for saving aggregated online content.
DEPENDENCIES	Selection of sources to be processed by the crawlers
CRITICAL FACTORS	Source selection: selection of sources for multimedia should be relevant to the project goals and be updated regularly. Outdated, incorrect or irrelevant data can compromise the quality of the results.

Table 1: Crawling and Scraping properties table

The schematic representation of the crawler service's logical structure is illustrated in Figure 3. This service comprises a constellation of distinct components, each endowed with diverse data input and output capabilities, all of which contribute to the execution of multiple functions. The principal role of the web crawling component resides in the

identification of nodes designated for extraction, thus serving as input web entry points. The web and social media search component primarily relies on the utilisation of available APIs, enabling the scrapping procedure. This scrapping process covers two distinct types of web input, specifically websites and social media platforms. When handling web content, this component receives web resource URLs and proceeds with the extraction of content, in textual or visual form. In parallel, regarding social media, the corresponding component sifts through content sourced from hashtags, user accounts, and other related channels, subsequently delivering social media posts. Ultimately, the resource filtering component assumes the role of implementing classifiers to evaluate the appropriateness of resources for the project's goals. It is charged with the duty of receiving and processing all the aggregated web information, resulting in the provision of a polished and relevant subset to the content database.

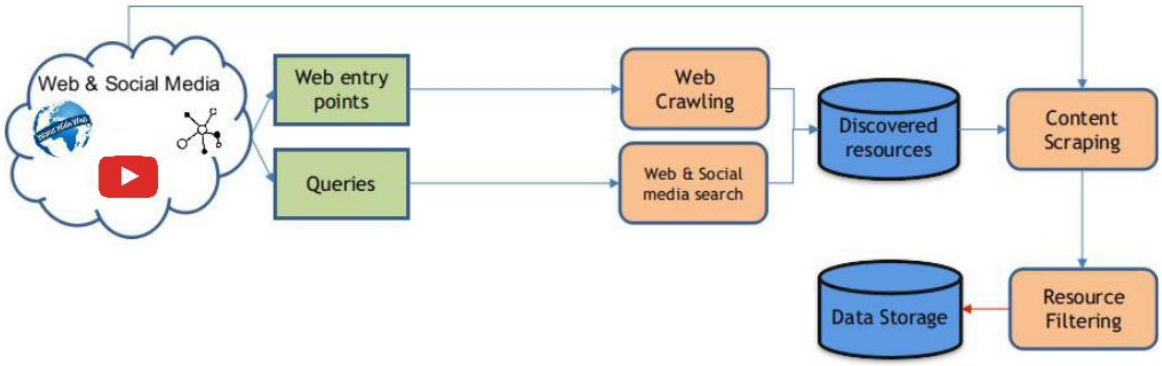

Figure 3: The logical design of social media and web data crawling/scraping service

3.1.1 Timeline and dependency from other modules

Development milestones: This service is currently under development, as the respective task started in M3, with expected delivery according to the following milestones:

[M16]: The 1st version of the crawling and scraping service will be delivered.

[M33]: The final version of the crawling and scraping service will be delivered.

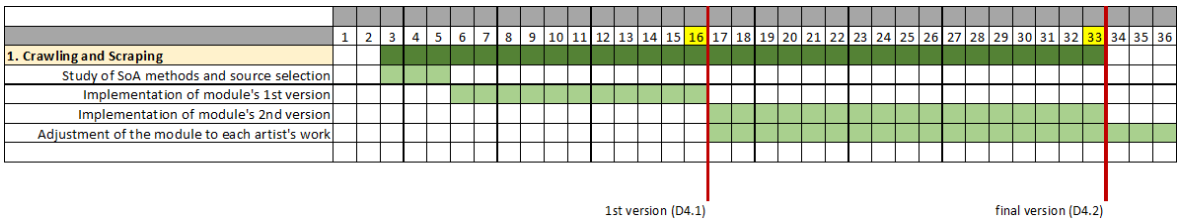

Figure 4: Timeline for the crawling and scraping service

3.2 Audio to Image techniques

This service will be developed in the context of task T3.3. Its objective is to produce image representations of captured sound signals by exploiting AI techniques. The audio input can be either speech, music, or other sounds (e.g., environmental sounds). The output images can depict an artistic visualisation of the audio features (e.g., spectrogram), an image that illustrates the source of the input sound (e.g., given the sound of a plane, an image of a plane is generated), or a related scene.

The audio-to-image techniques that will be deployed for the development of this service will be based on Artificial Neural Networks (ANNs) and specifically Generative Adversarial Networks (GANs), similarly to Chia-Hung et al.[1]. The goal is that the algorithm will learn the relationship between sounds and images in order to generate meaningful visual representations of the sound input. For the training of the audio-to-image model, we will use proper annotated datasets of paired sounds and images like SoundNet[2]. A pre-processing step, such as audio classification or speech recognition, will be deployed, depending on the needs of the ReSilence use cases.

An API will also be created so that the service can receive input requests and send its output to users.

[1] Wan, Chia-Hung, Shun-Po Chuang, and Hung-Yi Lee. "Towards audio to scene image synthesis using generative adversarial network." ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019.

[2] http://soundnet.csail.mit.edu/

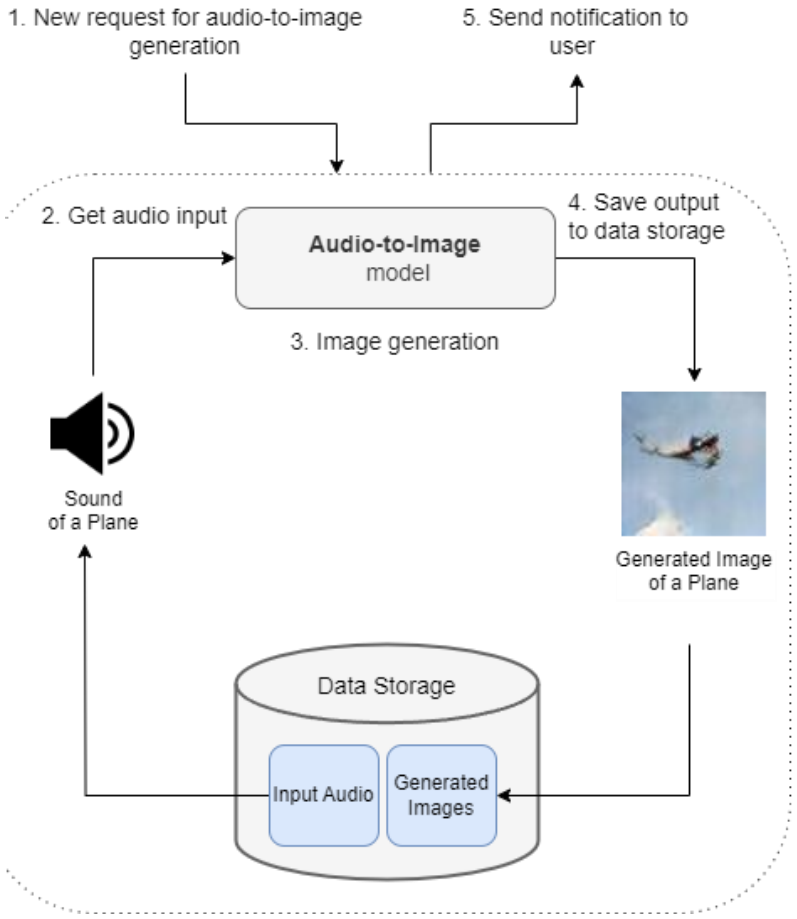

Figure 5: The logical diagram of the audio-to-image service with example input and output

The properties of the audio-to-image service are included in the following Table.

INPUT(S)	Audio (.mp3, .wav, .ogg, .mat files)
OUTPUT(S)	Images (.png, .jpg, .jpeg)
PROGRAMMING LANGUAGES/TOOLS	Python / OpenCV, Tensorflow, PyTorch, Flask
INTEGRATION	This service will communicate with a data storage service responsible for providing audio inputs to the module and storing the generated images. Depending on the users' requirements, it will either be an offline or online service.
DEPENDENCIES	Crawling and Scraping service to provide audio input

CRITICAL FACTORS	The performance and application of the service will depend on the context and quantity of annotated data used for the training of the audio-to-image model.

Table 2: Audio-to-Image properties table

3.2.1 Timeline and dependency from other modules

Development milestones: This service is currently under development, as the respective task started in M3, with expected delivery according to the following milestones:

[M17]: The first version of the audio-to-image service will be delivered.

[M22]: The final version of the audio-to-image service will be delivered.

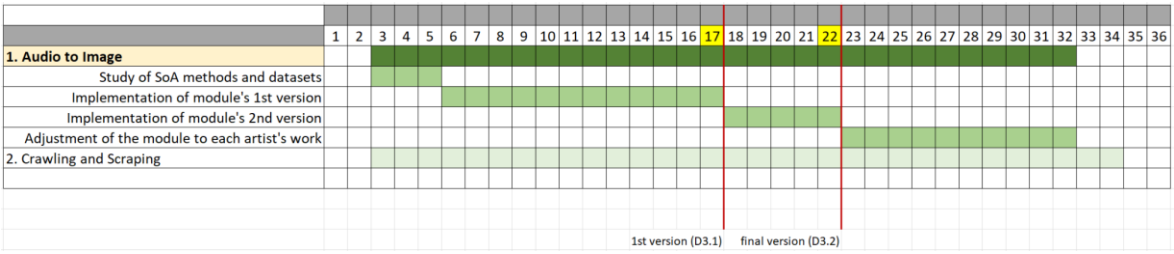

Figure 6: Timeline for the Audio to Image service

3.3 Image to Audio techniques

This service is undertaken under task T3.3, and its purpose is to synthesise audio from visual input using AI techniques. Input images can depict urban scenes, people, objects, text, paintings, or even be abstract. Output audio files can contain urban sounds, sounds of objects, speech, or music, depending on the given visual input.

Similarly to the audio-to-image service (see 3.2), image-to-audio techniques are also based on Artificial Neural Networks (ANNs) and Generative Adversarial Networks (GANs), similarly to Sheffer and Adi[3].

For the training of the image-to-audio model, we will use proper annotated datasets of paired sounds and images, such as VGGSound[4]. A pre-processing step, such as image classification or semantic segmentation, will be deployed, depending on the needs of the ReSilence use cases.

An API will also be created so that the service can receive input requests and send its output

[3] Sheffer, R., & Adi, Y. (2023, June). I hear your true colours: Image guided audio generation. In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-5). IEEE.

[4] https://www.robots.ox.ac.uk/~vgg/data/vggsound/

to users.

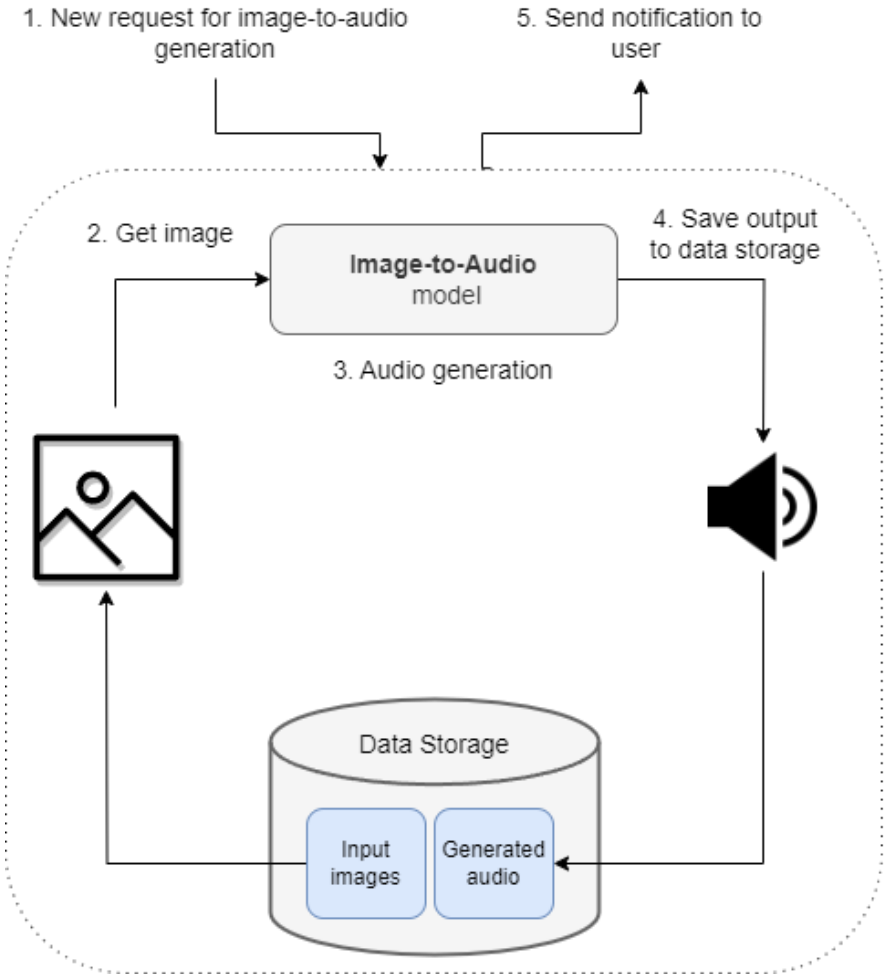

Figure 7: The logical diagram of the image-to-audio service

The properties of the image-to-audio component are included in the following Table.

INPUT(S)	Images (.png, .jpg, .jpeg)
OUTPUT(S)	Audio (.mp3, .wav, .ogg, .mat files)
PROGRAMMING LANGUAGES/TOOLS	Python / OpenCV, Tensorflow, PyTorch, Flask
INTEGRATION	This service will communicate with a data storage service responsible for providing images as inputs to the module and storing the generated audio files. Depending on the users' requirements, it will be either an offline or online service.

DEPENDENCIES	Crawling and Scraping service to provide visual input (images, videos)
CRITICAL FACTORS	The performance and application of the service will depend on the context and quantity of annotated data used for the training of the image-to-audio model.

Table 3: Image-to-Audio properties table

3.3.1 **Timeline and dependency from other modules**

Development milestones: This service is currently under development, as the respective task started in M3, with expected delivery according to the following milestones:

[M17]: The 1st version of the image-to-audio service will be delivered.

[M22]: The final version of the image-to-audio service will be delivered.

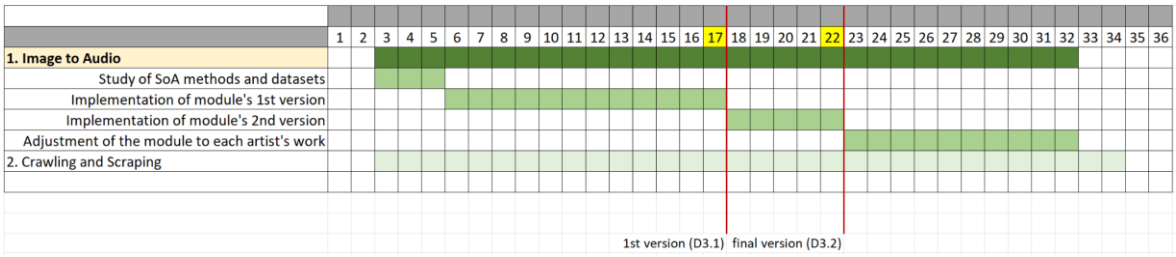

Figure 8: Timeline for the Image to Audio service

3.4 **Tactile Sound**

A wearable device (Tactsuit40), being explored and used in Task 5.4, is a vest equipped with haptic (touch or tactile) feedback technology. It works based on the process of audio to haptics which uses audio tracks and transforms the audio waveforms into haptic waveforms. We are going to integrate Tactsuit40 into game engines by using pre-existing SDKs.

INPUT(S)	Audio (.mp3, .wav, .ogg, .mat files)
OUTPUT(S)	Haptic vibrations (.haps files) written in a JSON-based format.
PROGRAMMING LANGUAGES/TOOLS	C#

INTEGRATION	This service will probably communicate with a data storage service responsible for providing sound. Depending on the users' requirements, it will be either an offline or online service. Test it in real-time through a game engine.
DEPENDENCIES	Image to audio service (audio).
CRITICAL FACTORS	The performance and application will depend on the quality of sound.

Table 4:Tactile Sound properties table

3.4.1 **Timeline and dependency from other modules**

Development milestones: The integration is currently under development, as the respective task started in M6, with expected delivery according to the following milestones:

[M19]: The 1st version of the image-to-audio service will be delivered.

[M34]: The final version of the image-to-audio service will be delivered.

3.5 **Multilingual Language Analysis Pipeline**

Multilingual language analysis pipeline, being developed within task T4.2, is a library of interconnected software modules designed to derive semantic structures from given written material in a number of languages. It includes software modules for identification of entities, i.e. concepts, in textual material and their linking with entries from open source lexical and semantic resources for the purpose of semantic disambiguation. It also includes a realisation of techniques for entity-oriented emotion analysis. The outline of the pipeline is shown in Figure 9.

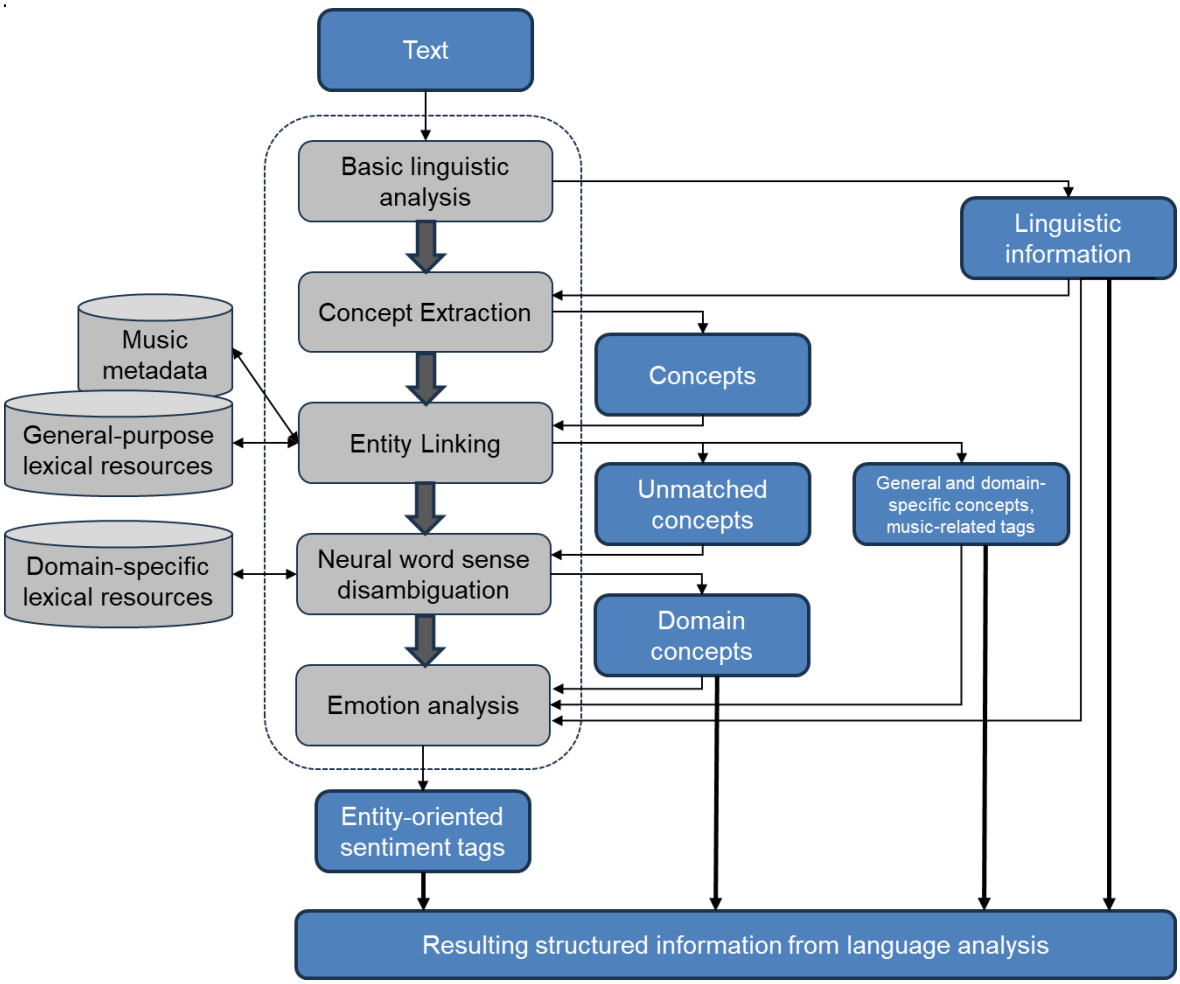

Figure 9: The logical diagram of the multilingual language analysis pipeline

3.5.1 Concept extraction and disambiguation modules

Concept extraction and disambiguation are core modules in the ReSilence language analysis pipelines. Concept extraction will be based on multilingual neural transformer models that will be implemented using Torch[5] architecture. Within the ongoing development, we have designed and evaluated an approach to extract core domain-specific keywords using generative transformer T5[6]. Comparison to UPF's pointer-generator network was carried out. The next steps include training concept extraction models in the music domain using relevant Wikipedia articles and specialised publicly available datasets. We will also automatically annotate a collection of crawled YouTube comments using labels from music-related knowledge bases and datasets depending on the user's requirements and fine-tune the best performing models.

[5] https://pytorch.org/

[6] Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W. and Liu, P.J., 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. The Journal of Machine Learning Research, 21(1), pp.5485-5551.

Disambiguation is envisioned as a two-stage process as shown in Figure 9. First, we will develop techniques to link concepts to various in-domain knowledge bases and general lexical resources. Second, for out-of-vocabulary unmatched entities, domain-oriented neural word sense disambiguation models will be trained. In the ongoing effort, UPF's linking to BabelNet and DBpedia has been deployed as a baseline for disambiguation. We also developed a basic neural sense-matching using contextual word-embedding models.

3.5.2 Entity-oriented emotion analysis

Entity-oriented emotion analysis is carried out at the last stage of the analysis pipeline. The ongoing development consists in adjusting existing advanced neural models for sentiment analysis to comply with ReSilence domain and concerned languages. Currently, we experimented with an application of large language models like GPT[7] to characterise open-class sentiments and experiences shared by concert spectators within the same topic. We also started exploring approaches to generalise resulting topics and sentiments into broader categories practical to the ReSilence users.

3.5.3 Timeline and dependency from other modules

The properties and dependencies of the multilingual language analysis component are included in the following Table.

INPUT(S)	Text and metadata (.csv, .json)
OUTPUT(S)	Key-value data structure (.json)
PROGRAMMING LANGUAGES/TOOLS	Java, Python
INTEGRATION	This service will communicate with a crawling and scraping service responsible for providing texts as inputs to the module. It will store results at a temporary data storage that can be accessed by components for multimodal data analysis and self-report generation. Users will be given access to the results for statistical and explorative analysis.
DEPENDENCIES	Crawling and Scraping service provides input; Multilingual Natural Language Generation Workbench

[7] Radford, A., Narasimhan, K., Salimans, T. and Sutskever, I., 2018. Improving language understanding by generative pre-training.

	utilises results.
CRITICAL FACTORS	The performance and application of the service will depend on the quality of input from social media, completeness of multilingual lexical resources, actuality of knowledge bases in the music sector, and performance of the individual modules within the pipelined analysis that can cause error propagation.

Table 5: Multilingual language analysis component properties table

Development milestones: This service is currently under development, as the respective task started in M3, with expected delivery according to the following milestones:

[M16]: The 1st version of the multilingual language analysis service will be delivered.

[M33]: The final version of the multilingual language analysis service will be delivered.

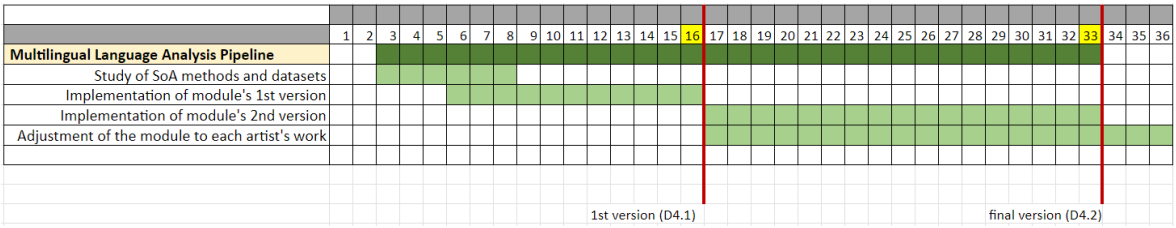

Figure 10: Timeline for the Multilingual Language Analysis Pipeline

3.6 Multilingual Natural Language Generation Workbench

Multilingual natural language generation workbench consists of a series of editors, debuggers, and graph transductors for the development of grammar-driven natural language generators. Primarily, the generator will serve for the creation of self-reports within task T4.4. It will be drawn upon FORGe system[8], which is designed to verbalise multimodal ontologies via a sequence of graph-transduction grammars, and will rely on project-specific rules of the semantic grammars and language-specific rules of the syntactic and morphological grammars. The generation workbench will include techniques for content selection to match individual user preferences. It will also enclose a collection of machine learning models for different natural language processing applications including neural modules to improve the fluency and wording style adequacy of texts produced by grammar-based systems.

[8] Mille, S., et al., 2019. Teaching FORGe to verbalize DBpedia properties in Spanish. In Proceedings of INLG, pp. 473-483.

3.6.1 Multilingual Lexical and Grammatical Resources

To adapt the grammar-based generator to the project domain, we will craft the semantic structures needed to cover the integrated outcome of the analysis in various modalities, craft the project-specific rules of the semantic grammars and the missing language-specific rules of the syntactic and morphological grammars, and craft descriptions of the lexical units in different languages. Within an ongoing development, we have been experimenting with automatic extracting generic abstract dependency templates to address the coverage issue faced by rule-based text generators. The further effort will be focused on obtaining semantic structures, rules and multilingual descriptions specific to the domain of ReSilence.

3.6.2 Suite of Machine Learning Models for NLP

This part of the workbench, primarily, will encompass pluggable neural NLG modules for improving the fluency of grammar-based generation. For this, off-the-shelf language models will be fine-tuned on project-specific and general domain textual data. The suite will also include specialised models to address particular needs within individual sub-projects in ReSilence (e.g., driven by open-call artists). Some of these models will be devoted to dynamic planning the composition of contents for the personalised reports to increase the relevance of the text for each user. At this stage, we are in ongoing development of a neural NLG module that combines outcomes of knowledge-to-text and text-to-text models for producing fluent and coherent integrated reports.

3.6.3 Timeline and dependency from other modules

The properties and dependencies of the self-report generation component are included in the following Table.

INPUT(S)	Analysis outcomes (.json)
OUTPUT(S)	Plain text (.csv, .json)
PROGRAMMING LANGUAGES/TOOLS	Java, Python, MATE text generation framework
INTEGRATION	This service will communicate with services performing various kinds of analyses for soundscape experience assessment. It will receive structured analysis outcomes and deliver personalised self-reports.
DEPENDENCIES	Emotional and cognitive sensing services, multilingual language analysis, and multimodal analysis of sound-related behaviour will provide inputs.

CRITICAL FACTORS	The performance and application of the service will depend on the availability of training material and coverage of the annotated resources in various languages.

Table 6: Self-report generation component properties table

Development milestones: This service is currently under development, as the respective task started in M3, with expected delivery according to the following milestones:

[M16]: The 1st version of the self-report generation service will be delivered.

[M33]: The final version of the self-report generation service will be delivered.

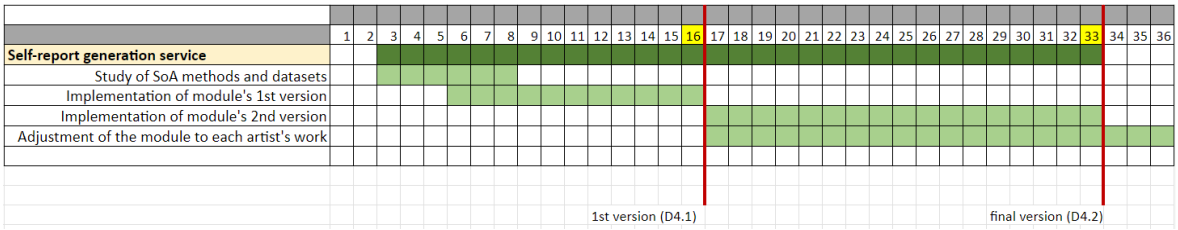

Figure 11: Timeline for the Multilingual Natural Language Generation Workbench

3.7 EyesWeb platform and libraries

The EyesWeb open software platform was developed by UNIGE to support the embodied interaction design process of art-inspired multimodal real-time applications (Camurri and Volpe 2016[9]; Volpe and Camurri 2011[10]; Piana et al 2016[11]) and for supporting research activities in non-verbal full-body expressive gesture, emotion, and social signals (e.g., Vaessen et al 2019[12]). It includes software libraries for the analysis of full-body mid-level expressive gestures. Current developments within T3.1.1 and T3.1.2 include (i) (T3.1.1) the extension of some of the library modules for individual expressive movement analysis, including the module on the automated measure of the Origin of Movement (Kolykhalova et al 2019[13]), (ii) (T3.1.2) the extension and development of modules for the analysis of non

[9] Camurri, A., & Volpe, G. (2016). The intersection of art and technology. IEEE MultiMedia, 23(1), 10-17.

[10] Volpe, G., & Camurri, A. (2011). A system for embodied social active listening to sound and music content. Journal on Computing and Cultural Heritage (JOCCH), 4(1), 1-23.

[11] Piana, S., Staglianò, A., Odone, F., & Camurri, A. (2016). Adaptive body gesture representation for automatic emotion recognition. ACM Transactions on Interactive Intelligent Systems (TiiS), 6(1), 1-31.

[12] Vaessen, M. J., Abassi, E., Mancini, M., Camurri, A., & de Gelder, B. (2019). Computational feature analysis of body movements reveals hierarchical brain organization. Cerebral Cortex, 29(8), 3551-3560.

[13] Kolykhalova, K., Gnecco, G., Sanguineti, M., Volpe, G., & Camurri, A. (2020). Automated analysis of the origin

verbal social signals in dyads and in small groups. Finally, a subset of modules from these two libraries (individual and social), originally developed for the EyesWeb software platform, will be subject to porting as external stand-alone Python modules.

INPUT(S)	movement data from sensors, video cameras, motion capture system
OUTPUT(S)	Plain text, OSC
PROGRAMMING LANGUAGES/TOOLS	C++, Python
INTEGRATION	Software modules can communicate in EyesWeb via OSC and with other network protocols.
DEPENDENCIES	Run as modules in the EyesWeb platform
CRITICAL FACTORS	Modules are designed for different types of sensor systems, real-time performance and flexibility in real world applications need a careful interaction design development process.

Table 7: EyesWeb platform properties table

3.7.1 **Timeline and dependency from other modules**

Development milestones: expected delivery according to the following milestones:

[M17]: Extensions of EyesWeb library modules.

[M22]: Stand-alone Python modules for the selected library modules.

3.8 **Interactive movement sonification**

The activities will focus on the development of sonification modules for the automated measure of cues on intrusiveness of sounds in city soundscapes, with a particular attention to the specific sound features that contribute to explain the phenomenon. This work will be developed in T5.5 in strict collaboration with the artist Andrea Cera in his project selected in the First Call of ReSilence. In particular, an in-depth analysis of the state of the

of movement: An approach based on cooperative games on graphs. IEEE Transactions on Human-Machine Systems, 50(6), 550-560.

art in sound features responsible to explain intrusiveness and annoyance in soundscapes is ongoing (M17). This will result in the selection of a reduced number of sound features, and the development of a software library for the real-time measure of such features and quantitative analysis of sound annoyance and sound intrusiveness (both off-line and real-time) (M19, M34).

3.8.1 **Timeline and dependency from other modules**

Development milestones: expected delivery according to the following milestones:

[M17]: Delivery of a comprehensive state of the art on techniques to model and measure intrusiveness in sound.

[M19 and M34]: Software modules for the automated measure of sound features (T5.5).

3.9 **Self and other whole-body perception and interaction**

Overall goal is to develop methods for participant experience at the individual as well as at the small group level. The focus is on nonverbal indicators of individual and group based artistic experience.

Collaborations are being prepared a.o. with UNIGE and MPG and include current OC artists.

Our goal is to explore the neural and behaviour mechanism underlying viewing personalised (self) and generic (non-self) avatars. This is followed by other behavioural and EEG studies of interactions between agents and studies that include physiological measures.

During the first behavioural experiment, participants first did the embodiment VR task. They saw one of the avatars in the mirror, which included self-standard body shape, self with modified body shape, non-self-standard body shape and non-self-modified body shape. They touched the bubbles, catched crystals in front of the mirror, as well as lightly lifted their arms up. After the VR embodiment task, they passively saw the embodiment avatars in the computer task without VR headset. We used three colours of T-shirts of avatars (blue, purple and green), and three orientation degrees of avatars (0°, -90°, +90°).

In parallel studies we investigate whether body perception depends on full awareness or can also take place under limited awareness conditions.

3.9.1 **Timeline and dependency from other modules**

[M16]: The 1st version of the EEG analysis will be delivered.

We recruited 27 healthy participants and collected their Electroencephalography (EEG) data, as well as behaviour data. All the participants are from Maastricht university.

[M33]: The final version of the EEG analysis service will be delivered.

Data analysis is currently in progress. We found the Event-related potentials N170 and VPP are driven by not only face, but also body expressions from avatars and inside a VR environment as in Lu et al. 2023.

Before this methodology can be exported and used by third parties, we need to fully understand the underlying neurophysiology. When that is clear a simplified version of the set up can be developed and would be available for use by OC artists. We expect that

project development will be on an individual artist basis.

3.10 Integrated modular system for the mobile measurement of participant behaviour and experiences

Measuring how people respond to the artistic interventions and performances that will result from the OC is of crucial importance for the aims of ReSilence. Such responses can be both observable behaviour and subjective experiences. Based on an existing stationary solution at the MPIEA, we will develop a modular and mobile measurement system that provides all teams with the necessary means to measure those participant responses that are relevant for the respective PUC and artistic project. These modules will consist of 1) experience questionnaires that can be administered on people's own mobile devices (such as smartphones), 2) experience measurement via wearable physiological sensors, 3) behaviour and movement tracking via cameras. Modules can be used individually or in combination.

The crucial part will be to develop a remote-controllable interface that interconnects and synchronises the three systems and links all incoming data to the respective participant.

3.10.1 Timeline and dependency from other modules

A prototype will be tested in M17 and M18 in the context of PUC1. Further development work will be conducted in M19–M22. There are no dependencies from other modules.

4 SYSTEM INTEGRATION

To design the architecture of ReSilence, a bottom-up methodology has been crafted. This approach involves defining the individual elements of the system and then linking and connecting them to establish a cohesive and integrated framework.

This approach consists of the following five steps:

- Technology mapping: using a questionnaire-based survey, which was circulated within the technology consortium partners, we gathered their feedback in order to identify the technologies that will form part of the toolkit. We also inquired about their use, gave feedback about their technical and system requirements, as well as their data management aspects, internal architecture and other concerns.
- Technology classification: once the technologies were identified and described clearly in the preceding step and classified according to technological and system criteria that tried to locate the “region” in the architecture relevant to each technology. As if a puzzle piece is placed in the general area where it belongs, other pieces and service combinations were also examined.
- Technology allocation: during this stage of technology allocation, the focus lies on grouping components and technologies that are intended to be utilised by a specific artist and address the challenges of a particular PUC.
- ReSilence toolkit: the toolkit will allow artists to create different audiovisual scenarios allowing the multisensory experience of outdoor and indoor spaces. Artists will be using the collected audio-visual data to assist and empower specific user groups by means of multisensory technology. This toolkit, for instance, through an audio-to-image and image-to-audio synthesis model, will facilitate the virtual design of soundscapes and vice versa.
- Architecture design: the final step of this methodology consists of a consolidation of the architecture design obtained in previous steps, first by revisiting the requirements and making sure that critical requirements were met, and important functionalities were streamlined. In addition, flexibility in the design was introduced by allowing the architecture to evolve iteratively, and morph into a paradigm that will meet the project objectives.

The following diagram represents this methodology.

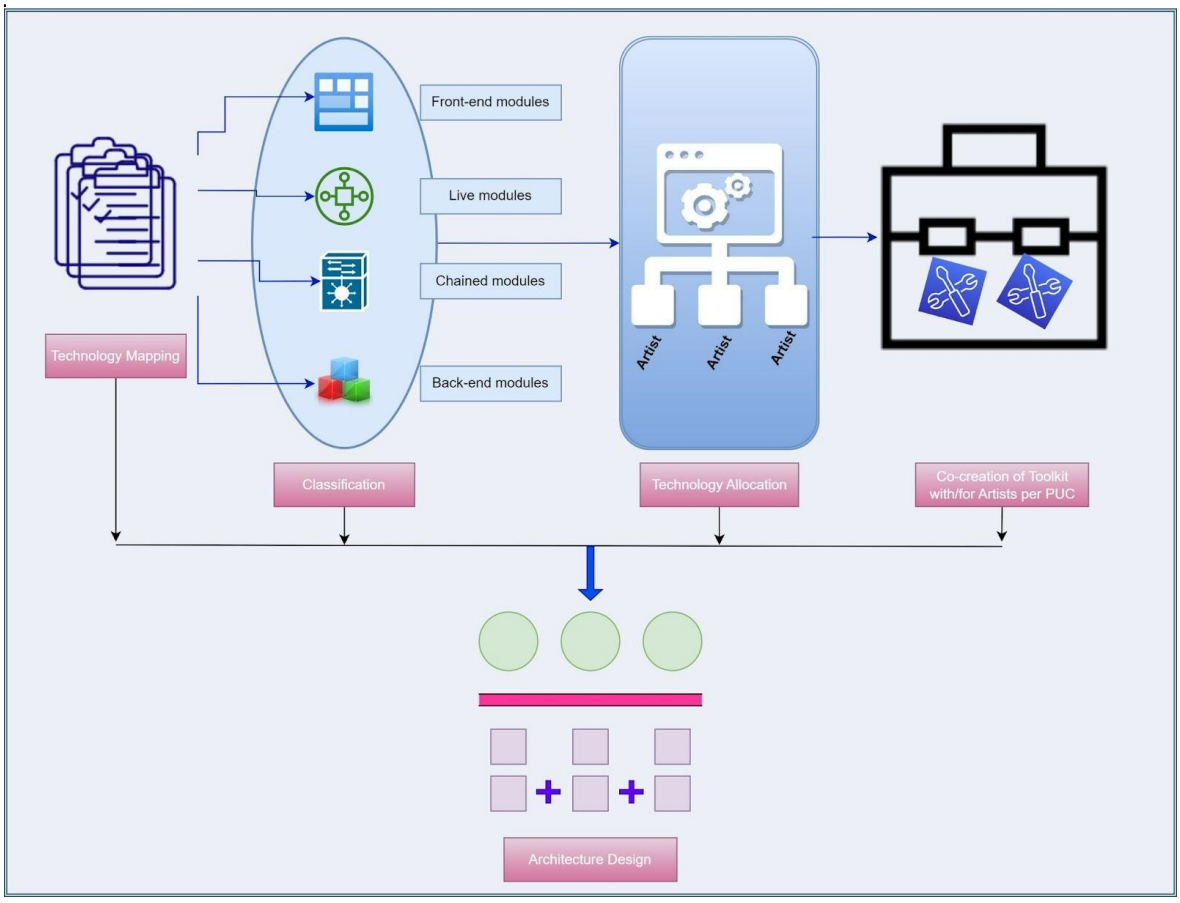

Figure 12: ReSilence Architecture Design Methodology

System integration is the backbone of the project, serving as the bridge that connects a diverse range of tools and services into a unified ecosystem. As outlined in previous chapters, the project boasts an impressive array of tools and services, each with its own specialised function. They act as the building blocks of the system, each bringing its unique capabilities, creating a cohesive and synergistic system.

One of the main integration challenges is the compatibility and connection of the various modules developed in each work package (WP). To address this challenge, a strategy involving the development of dedicated APIs (Application Programming Interfaces) has been implemented for selected modules. Each of these tools and services in the ecosystem is equipped with its own clearly defined API, allowing communication between compatible modules and intuitive interaction with users. Below we present how this API-driven approach will benefit the project, when used:

- **Unified Access Point:** Users and artists can access a wide range of functionalities through a single-entry point - the APIs approach, when present. This simplifies the user experience and streamlines access to the various tools.
- **Cross-Tool Collaboration:** The APIs enable cross-tool collaboration for compatible services. For instance, data generated by one tool, such as the crawling and scraping

service, can be utilised by another, opening up new possibilities for creativity and data synergy.

- **Real-Time Collaboration:** Artists can collaborate in real time, where provided, harnessing multiple tools simultaneously to create dynamic and multi-dimensional artworks.
- **Scalability:** The modular nature of APIs allows for easy scalability. New tools and services can be integrated into the system without disrupting existing functionality.
- **Data Flow Optimization:** Data generated or processed by one tool can flow seamlessly to others, optimising resource utilisation and minimising redundancy.

As a part of our system integration strategy, we have established a dedicated GitHub repository accessible at https://github.com/resilencerepo/resilence_main. This repository serves as a collaborative platform where each partner and developer can efficiently upload, manage, and track the progress of the modules hosted, ensuring efficient version control.

This integration not only enhances user experiences but also empowers artists to push the boundaries of their creativity. As moving forward, the possibilities within the unified ecosystem are boundless, bringing forth a new era of technological and artistic synergy.

5 DEVELOPMENT CYCLE AND PILOT USE CASES

The overall plan for the development of the ReSilence toolkit is to start supporting basic scenarios and proceed stepwise by adding more features towards the final system. The detailed timelines of the individual modules, which are presented in previous sections will follow the general cycle: after the completion of each prototype version of ReSilence (timing defined by Milestones as discussed in the next section), the functioning of the prototype is assessed, the needs for further development are identified and agreed upon, a detailed work plan for the next prototype version is finalised and the development cycle towards the next prototype is initiated. This iterative procedure will be documented in D5.2, explaining the different development phases of ReSilence tools and its architecture. Each version will happen before each Milestone and will give extensive details about the 1st and 2nd prototypes of ReSilence tools.

The objectives of ReSilence are presented in the following Figure.

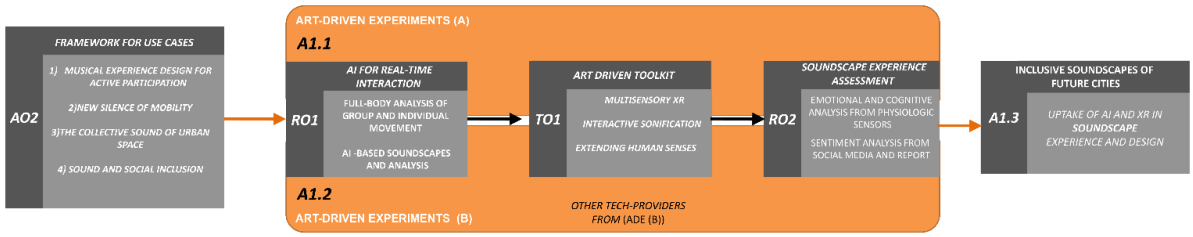

Figure 13. The objectives of ReSilence

5.1 Pilot Use Case 1: Musical Experience Design

Motivation: This use case explores how new digital technologies can redefine conventional roles in music and performance. Integrating audience interactivity and leveraging emerging technologies, artists have been invited to disrupt the dynamics of musical experiences. They should aim to seek to create innovative music performances, dance shows and participatory spaces that break free from traditional constraints. Collaboration between artists, technologists, scientists, architects and designers is crucial in order to bring to life innovative artistic expressions. Through this quest, the goal is to reimagine the relationship between performers, audiences and their surroundings, offering fresh perspectives in both physical and virtual realms. Related to this is the idea that a creative and aesthetically satisfying use of AI and XR technologies will positively affect people's general views of these technologies.

Story - Tim Otto Roth: The project "Theatre of Memory'' presents distinctive challenges in the examination of interplay and perception within an immersive sound environment. The primary objective is to unfold an innovative and dynamic acoustic network that reflects the fundamental neuronal processes associated with memory. In an effort to review established approaches, this project sets out to institute a new model for composition and interaction with musicians, using neurobiological feedback as a foundation. Key points of research include the translation of nervous activity into acoustic signals and the creation of interconnected networks of audioneurons. With a strong focus on microtonal relationships and individual

tone assignments, the project delves into the creation of complex and dynamic spatio-temporal tonal sequences. Additionally, it raises questions about the impact of sound runtime and the acoustic environment on the compositional process. The artist's motivation stems from the desire to design a novel musical experience that facilitates spatial participation and provides an intuitive understanding of neurobiological processes.

Test: The expected outcome of the residency is an installation that will feature a light augmented sound laboratory with an expanded number of audioneurons. Studying perception and experimenting with different auditory and visual experiences together with MPIEA, the project seeks to enhance the concept of immersion. The interactive mode will foster feedback from the audience or musicians, creating complex spatio-temporal tonal sequences. The principal ambition is to push the boundaries of music experiences and offer visitors a unique and engaging journey into neurobiological processes.

Developing a premiere version of "Theatre of Memory'' in Berlin, the project envisions creating a captivating narrative that unfolds within the domain of memory. The historic Tieranatomisches Theater, a captivating venue located in the heart of Berlin, has been confirmed as the first exhibition space. Built in 1789/1790, this anatomy hall at Humboldt University provides an exceptional setting for the project. The anticipated presentation in January 2024 will be a highlight in the accompanying program of the renowned brain exhibition at the Charité Medical History Museum,supported by the Charité Foundation. In addition to this, discussions with national and international neuroscientists are underway to explore further exhibition opportunities.

5.2 Pilot Use Case 2: The New Silence (Sound and Mobility)

Motivation: The rise of the industrial revolution ushered in a new era of machinery, marking a kinetic revolution, characterised by the movement of mechanics. In this transformative landscape, sound design emerges as a user-focused experience, where utilitarian sounds can evolve into sources of artistic innovation. Electric cars, for instance, offer an unexplored territory for composers to craft unheard auditory landscapes, while also presenting an opportunity for the vehicle itself to become a dynamic performance art installation, where the driver takes on the roles of composer and performer. Moreover, the reflective properties of surfaces play a significant role in reimagining urban spaces and determine the choice of construction materials. Architects and artists explore the domain of indoor acoustic design, investigating the resonance of sound and the way our bodies engage with space, resulting in the creation of distinct sounds. Expanding on this knowledge, it is essential to apply it to public spaces like restaurants, subway stations and train stations, treating them as musical instruments that can be enhanced through intentional design interventions. The sound absorption and reflection qualities in these environments will shape the total impact for individuals.

Story: Designing sounds for electric cars presents a dual challenge of balancing safety and functionality. It entails careful consideration of sound composition and placement within the

vehicle to ensure that the sounds effectively alert pedestrians and other drivers without being intrusive or distracting. Furthermore, collaborative efforts between sound designers, engineers and user experience experts can enhance the driving journey through well-integrated sound elements. Another aspect to address is creating distinctive auditory sensations that align with the brand identity and recognition of each electric car manufacturer. Developing sounds that reflect the unique characteristics of the brand, electric car manufacturers can establish a sonic signature. In addition, the assimilation of sound design into electric cars has a profound impact on the urban space. These soundscapes can contribute to the character of urban environments, creating a harmonious atmosphere for both drivers and pedestrians.

Test: A potential artwork or prototype resulting from addressing these challenges could be a sound installation that simulates the sonic experience of driving an electric car developed together with MOBEN. This would include a customizable system that responds to user interaction, providing a personalised experience. The installation would allow visitors to navigate through different auditory landscapes, showcasing the creative possibilities in sound design for electric vehicles. Additionally, it could incorporate elements of urban design, demonstrating how sound interventions can transform public spaces into engaging environments.

Dissemination strategies for designing sounds in electric cars could include workshops, conferences and online platforms. Publishing research papers and forming partnerships with automotive manufacturers and industry associations would contribute to knowledge sharing. Public demonstrations and exhibitions would showcase the impact of sound design.

5.3 Pilot Use Case 3: Sound of Urban Spaces

Motivation: Noise pollution presents a significant issue in urban areas, and consequently, incorporating soundscape design into city planning has gained attention in recent years. However, there is still a need for effective methods to shape the soundscapes of spaces, ensuring the well-being of the public. The core challenge is to develop tools and techniques that enable architects and urban designers to address noise pollution, with an emphasis on revolutionary ideas from sound artists. This process will involve simulations and prototypes implemented in smaller sections of the urban environment, initially involving a limited number of individuals and gradually expanding to larger areas of cities and communities.

5.3.1 Andrea Cerra

Story: This project centres around the development of an interactive environment driven by sound, which aims to investigate the connection between movement, soundscapes and the perception of intrusiveness. The main goal is to create a space where participants have the ability to shape a virtual city soundscape through their movements and raise awareness regarding the negative impact of polluted urban soundscapes and foster active engagement in mitigating these issues. A number of sub-challenges are given special attention by the

project, such as examining the degree of intrusiveness in urban soundscapes, exploring innovative techniques to counteract it and utilising movement qualities and physiological signals to analyse how soundscapes influence participants' behaviour. The artist concentrates on investigating expressive movement qualities at a mid-level and understanding their variations in relation to different levels of intrusiveness, in order to comprehend how sound impacts both movement and overall quality of life.

Test: As a result of the residency, an interactive installation is being developed. This installation will feature portable micro-installations and software modules, developed with UNIGE, that can be experienced individually or in various settings. It will provide opportunities for participants to engage with the soundscape and analyse counter-measures to intrusiveness. The objective is to create a collection of modules that can be used in different situations, such as walks in noisy streets, home environments, conferences, or educational workshops.

The interactive installation, developed as a result of the residency, is potentially planned to be showcased in Genova, specifically during the EU STARTS event, as part of the Festival della Scienza.

5.3.2 **Caroline Claus**

Story: The main emphasis of the project lies in the investigation of sonic space shifts that evoke physical and emotional responses in individuals, thereby expanding the scope of acoustic strategies beyond acoustic comfort. A key aspect involves developing a sonogeographical dérivé, a cartographical work on sonic space shifts that exposes the impact of regeneration processes on the sonic materiality of urban areas. Ethical considerations regarding audio logging in public spaces will be assessed, involving the implementation of experiments in the fields of architectural and urban design. Creating immersive AR/VR audio experiences through the synthesis of audio logging data and constructing a sonorous scale model, the artist seeks to transform our understanding of sonic environments and bring about innovative approaches.

Test: The artist aspires to develop a sonogeographical dérivé together with AUTH, CERTH, UPF, which serves as an audio observatory and an audio paper. The observatory will facilitate live streaming, collective discussions, long-term observation and the production, synthesis, mixing and sharing of recordings. It will provide insights into the sonic materiality of urban spaces and contribute to the understanding of the affective aspects of sonic geographies. Additionally, an AR/VR experience will be developed to present the sonorous scale model online, allowing users to submerge themselves in the transformed sonic environment. At the final stage, a book launch and a seminar on sonic urbanism will be organised, functioning as a platform to celebrate and disseminate the synthesised sonogeographical dérivé among the public.

Public meetings are planned to be organised in the Territory North of Brussels, creating vibrant urban spaces that bring together diverse communities. These local engagement tactics are complemented by international dissemination that will occur through participation in

conferences, symposia and sound artistic events, striving to reach a wider audience interested in sonic and affective environmental aesthetics and promote exchanges within the context of sonic urban regeneration.

5.4 Pilot Use Case 4: Full-Body Sound Experience

Motivation: This use case emphasises the transformative power of vibration and tactile sensation in redefining our perception of sound and visual inputs. It highlights two key aspects:

a) The significance of a holistic approach to music and sound, involving the entire body rather than just the ears.

b) The conversion of physical objects into captivating soundscapes and vibrations, drawing inspiration from artists like Christine Sun Kim.

With specific attention directed towards individuals with hearing or visual impairments, the purpose of this exploration is to bring about a technological revolution by acting as a catalyst for enhancing our sensory capabilities. By bypassing the impaired areas (eyes, ears), we can redirect and distribute sound and visual information through alternative ways, enabling a more comprehensive and immersive spatial experience.

5.4.1 Gustavo Maggio

Story: Enveloping themselves in the resonance of sound and the interplay with the human body across various spatial contexts, the artists embark on a creative project. Their goal is to uncover the complex connections between acoustics, the intricacies of human anatomy and the transformative power of haptic full-body sound experiences. Through this study, they expect to unlock new dimensions of sensory engagement, where the boundaries between sound, physical sensations and immersive environments blur, in order to create profound and captivating artistic encounters. A key focus is on reimagining the soundscape of silence for marginalised groups, particularly people with disabilities who often face exclusion from conventional auditory experiences. Furthermore, the artists set sights on addressing the potential disruption of the ocean's soundscape due to sonar ocean exploration. They want to portray the narratives of marine life and other planetary inhabitants, highlighting the interconnectedness of all beings. Through their venture, the artists strive to create interactions that challenge societal norms and expand the understanding of the world around us.

Test: The result anticipated from this project is a multisensory installation that combines haptic soundscapes with virtual reality (VR) mediascapes. Participants will have the opportunity to wear sensory wearable suits, programmed in collaboration with CERTH and undergo the haptic-sensory compositions inspired by the scrutiny of sound and the ocean. The installation aims to be accessible to visitors of diverse abilities, with a focus on gathering feedback from persons with disabilities to enhance the inclusivity of the experience. Additionally, the artists envision the potential for the wearable technology to be adapted into concert theatre chairs, transforming the audience's passive listening process into an active,

full-bodied musical participation.

In their quest for a suitable venue for the final showcase, the team plans to leverage their extensive network of museums, galleries and performing theatres with whom they have previously collaborated. They plan to premiere the first installation in Europe, potentially in Berlin, and subsequently tour it in different cities. Drawing on their experience in curating immersive exhibitions, Gustavo, Wendy and Joyce are well-equipped to translate their work into tangible and interactive exhibition experiences for the public. With potential venue partners, such as the Red Dot Design Museums, the National Museum of Singapore, Esplanade theatres and the Humboldt Forum, they have a range of options to consider for staging their art piece. Additionally, being part of research clusters and networks like Matters of Activity and Game Lab Berlin provides them access to further exhibition opportunities and venues in Berlin, such as the Kunstgewerbemuseum and the Tieranatomisches Theater.

5.4.2 **Loukia Tsafoulia**

Story: The project “Soft” deals with several challenges associated with the development of adaptive and therapeutic environments. One key challenge is to design a dynamic space that can intelligently respond to an individual's sympathetic nervous system activation. This requires a deep understanding of the physiological and psychological states of occupants and how they can be influenced by sensory aspects, such as sound and light. Another concern is to incorporate distant-to-the-body technology, with the intention of adapting the sonic and light characteristics of the environment in real-time. This involves creating biofeedback-mediated relaxation techniques and utilising body pressure pockets to modulate pressure intensity in relation to sound and light. Additionally, the team of artists seeks to confront the challenge of neurodiversity by creating an inclusive space that caters to the needs of both neurotypical and neurodivergent individuals.

Test: The residency aims to achieve the desired outcome of creating a functional prototype for the "Soft" project. This prototype will be a deployable and adaptive therapeutic environment conceptualised as a spatial wearable. It will take the form of an encapsulated pod, where individuals can retreat to when overwhelmed or overstimulated. The prototype will incorporate various technologies and design elements to modify sensory aspects of the interior environment, with a particular emphasis on sound and light. It will utilise body-based biometrics and closed-loop biofeedback to intelligently modify the sonic and light characteristics in real-time. The purpose of these modifications is to create a soothing and calming environment that assists the occupants in self-regulating their emotional and physiological states. The design will cater to various contexts, such as sensory-loaded lobbies, concert halls, exhibition venues and hospitals. Through its deployment, the team intends to further examine the therapeutic benefits of the adaptive environment and its potential in fostering the well-being of individuals, both neurotypical and neurodivergent.

'Soft' is set to undergo testing and evaluation at Thessaloniki Concert Hall (TCH). The preliminary idea will be presented at the UIA 2023 CPH World Congress and published under Springer's Sustainable Development Goals series. The outcomes of the residency will be shared through publications, conferences and events, including the Venice Biennial and Ars

Electronica festival. The 'Soft' prototype will be deployed at venues such as the MusiXLab space and Thessaloniki Concert Hall. Collaboration with neurodivergent communities, both in Europe and the US, is planned. Online platforms will be used for project updates and public engagement.

6 PROJECT TIMELINE AND MILESTONES

The following figure depicts the project timeline and milestones.

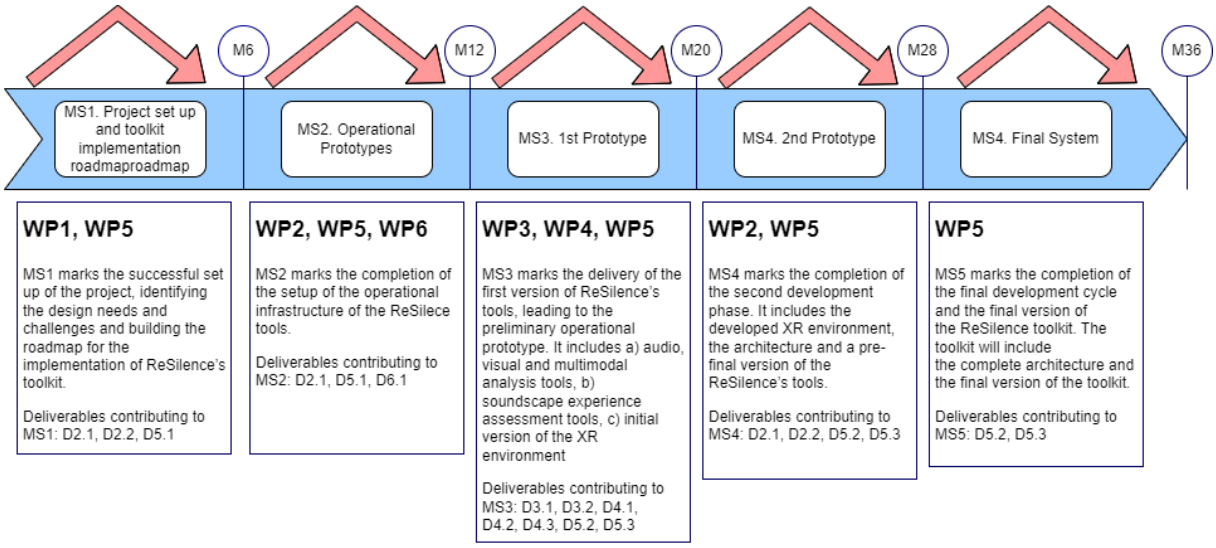

Figure 14: Project timeline and milestones

ReSilence toolkit will be built incrementally and iteratively. A project setup and toolkit development roadmap are included in the MS1. MS2 signifies the conclusion of establishing the operational infrastructure for the ReSilence tools. MS3 marks the delivery of the first operations prototype, which will be completed in MS4, through the 2nd prototype. Finally, MS5 will result in the completion of the development cycle.

7 CONCLUSIONS

This deliverable presents the technical roadmap of ReSilence toolkit. It includes the required information for the development of the project with a view to attaining the scientific and the technical objectives envisaged. However, since the project is following an iterative development procedure (use cases requirements-development-evaluation), it is expected that small adaptations and deviations on the initial technical specifications and the module functionalities foreseen by this document will be needed to satisfy the final user requirements. Potential changes/adaptations might be required at component or subcomponent level.