Time Machine

A Large-Scale Research Initiative on Building the Big Data of the Past for the Future of Europe

Document for a public consultation

Prepared by: The Time Machine Consortium
Date: June 2019
Contact: contact@timemachine.eu

Disclaimer

1. Although elaborated with EC funding support the content presented in this document has not be reviewed yet by the EC services.
2. The information and views set out in this document are those of the authors (Time Machine Consortium) and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

How to cite this document:

Time Machine Consortium (2019) Time Machine Document for a public consultation, available at: https://www.timemachine.eu/time-machine-roadmaps-mapping-the-future

---

CONTENTS

1 INTRODUCTION

Time Machine (TM) is a Large-Scale Research Initiative (LRSI), pushing the frontiers of scientific research in Information and Communication Technologies (ICT), Artificial Intelligence (AI) and the Social Sciences and Humanities (SSH).

TM is built around the vision to develop the Big Data of the Past, a huge distributed digital information system mapping the European social, cultural and geographical evolution. This large-scale digitisation and computing infrastructure will enable Europe to turn its long history, as well as its multilingualism and multiculturalism, into a living social and economic resource for co-creating a common future.

The proposed LRSI will use space and time as shared references across domains, disciplines and cultures, to understand and give value to constructions, artefacts, observations and data produced over centuries, enabling Europeans to better appropriate their heritage and strengthen the feeling of European belonging.

The key objective of the TM CSA project is to develop a full LSRI proposal around this TM vision. Detailed roadmaps will be prepared, organised around four pillars, namely: science and technology (pillar 1), TM operation (pillar 2), exploitation avenues (pillar 3) and addressing/enhancing the framework conditions for the development of the LSRI (pillar 4).

The methodology for building the roadmap of this initiative foresees the elaboration of draft roadmaps for each pillar by working groups composed of Consortium experts, followed by a round of consultations with the largest possible number of relevant external stakeholders. The consultations will enable the Consortium to finalise the pillar roadmaps in a way that reflects the needs and expectations of a pan-European ecosystem that has been built around Time Machine and is currently expanding at fast rate.

This document is prepared having in mind these public consultations. It gives an account of the stage of the LSRI design on Month 4 of the TM CSA (June 19), organised in the following sections:

- An overview of the TM LSRI, presenting the rationale, vision and proposed structure of Time Machine, as well as the key concepts for the design of the roadmaps (section 2)
- The main aspects in the current state of development for the TM pillars (sections 3 – Science and Technology, 4 – TM Operation and 5 – Exploitation Avenues)
- The expected outcomes and impact (section 6)
- The Time Machine community and its evolution to a dense ecosystem of research and innovation actors, under the governance of a purpose-built association, the Time Machine Organisation that was officially launched in June 2019

Definitions of terms and abbreviations used are provided at the end of the document.

2 OVERVIEW OF THE INITIATIVE

2.1 Rationale

Over the centuries, the national, regional and local identities of Europe have evolved in relation to one another through large swathes of transnational mobility and through dense exchanges that have shaped European languages, traditions, arts and many other aspects of human activity. These processes have largely contributed to the creation of a European culture characterised by diverse historical memories, which have laid the foundations to values and ideas harmonised by pluralistic and democratic dialogue. To-date, however, increased globalisation, changing demographics and their threat against the idea of a shared past, as well as the resurgence of unresolved conflicts deep-seated in European memory are key drivers of a 'localisation backlash' that places local and personal interests above any other. These growing trends present a clear threat to the cohesiveness of European cultural identity and sense of belonging.

Pluralistic and democratic dialogue in Europe has traditionally been facilitated by important intermediaries, such as cultural media and institutions acting as cornerstones of our shared values, principles and memories. Today, the dialogue between different actors and the historical visions they embody is complicated by the rise of private digital platforms that have created a new space of opinion-leadership, as well as new forms of political expression and participation. Managed by proprietary algorithms, such platforms may prioritise popularity and personal agendas over historical and cultural data, opening the way to fake news. In the resulting crisis of authority that affects journalism, academia and politics, many people do not trust anymore the information received from these institutions.

These unprecedented transformations create a vital need for Europe to restore and intensify its engagement with its past as a means of facilitating an evidence-based dialogue between diverse historical memories, their values and mutual interdependencies and building a common path across generations. Time Machine responds to this need by building the required infrastructure, and an operational environment for developing the "Big Data of the Past" that will transform history and culture across Europe, opening the way for scientific and technological progress to become a powerful ally to safeguarding European identity and democratic values.

For Time Machine, digitisation is only the first step of a long series of extraction processes, including document segmentation and understanding, alignment of named entities and simulation of hypothetical spatiotemporal 4D reconstructions. The hypothesis pursed by Time Machine is that such computational models with an extended temporal horizon are key resources for developing new approaches to policy making and to offering services to European citizens and consumers.

Still, there is one more crucial reason supporting the cause of Time Machine. After the creation of the web that digitised information and knowledge and the social media that digitised people and characteristics of human behaviour, a third technology platform is being created, digitising all other aspects of our world, giving birth to a digital information "overlay" over the physical world, a "mirror-world". The mirror-world will aim to be an up-to-date model of the world as it is, as it was and as it will be. All objects (including representations of landscapes) of the mirror-world will be machine-readable, and, therefore, searchable, traceable and subject to be part of simulations by powerful algorithms. In the mirror-world, time will be a fourth dimension, as it will be very easy to go back to the past, at any location, reverting to a previous version kept in the log. One may also travel in the other direction, as future versions of a place can be artificially created based on all information that can be anticipated about the predictable future. Such time-trips will have an increased sense of reality, as they will be based on a full-scale representation of the present world. Time Machine is today the most advanced concrete proposal to build the first version of a European mirror-world.

Like the other two platforms, the mirror-world will disrupt most forms of human activity, as we know them today, giving birth to an unimaginable number of new ideas (and many problems) and creating new forms of prosperity from new forms of economic and social activity that will shape new behaviours and ecosystems. In this scenario that is currently unfolding, Time Machine will enable Europe to be one of the leading players, shaping the mirror-world according to its democratic values and fundamental ethics related, among others, to open standards and interoperability. With Time Machine, while it will have a powerful tool to strengthen its cohesion and sense of belonging, Europe has, moreover, an opportunity to impose its own terms against the multinational technology giants that will fight for dominating this new technology platform, just as those who now govern the first two platforms have done in the past.

2.2 The Vision

The vision behind Time Machine is structured around the concept of "Big Data of the Past". Figure 1, below, symbolically represents the digital information currently available for each period of our history.

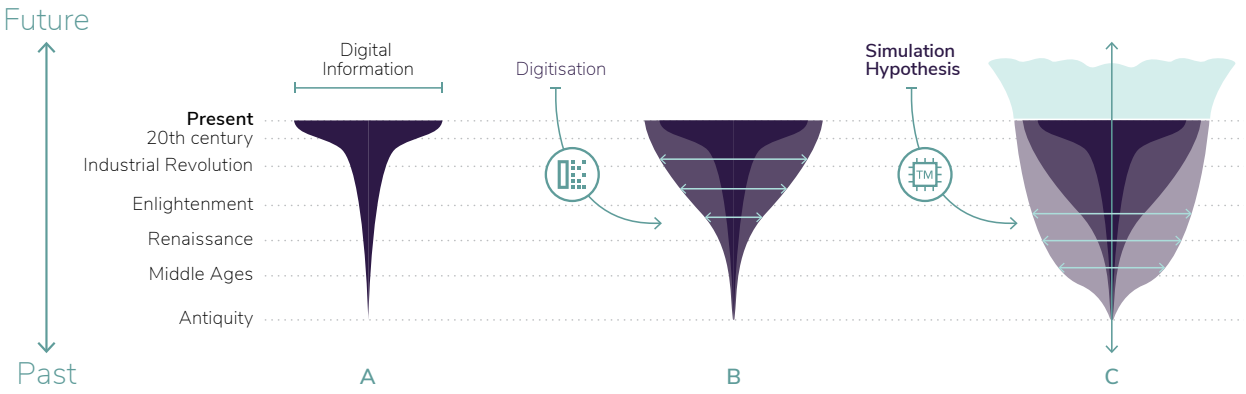

FIGURE 1: Creating the Big Data of the Past:
(A) Current situation. (B) Extension based on digitisation and processing of new sources. (C) Extension based on simulation.

By plotting the amount of digital information available today (horizontal axis) against time (vertical axis), we should expect to see a funnel-shaped figure. Information about the most recent years is abundant, forming the large plateau representing the funnel mouth: The Big Data of the Present. The curve shrinks rapidly as one moves down the graph and back in time (Figure 1A).

TM aims at enlarging the stem of this funnel, firstly, by developing the technology and infrastructure for conducting massive digitisation and processing of cultural heritage sources (Figure 1B). Secondly, this enlarged dataset will be the basis for simulating possible pasts in order to reach an unprecedented density of information: the Big Data of the Past (light grey area in Figure 1C); this enormous volume of data will also boost modelling capacity, enabling us to make evidence-based predictions for the future (light blue area in Figure 1C).

2.3 Proposed Structure

The Time Machine LSRI is articulated around four pillars, each defining a specific objective of the initiative and comprising thematic areas, as shown in Figure 2 and discussed below.

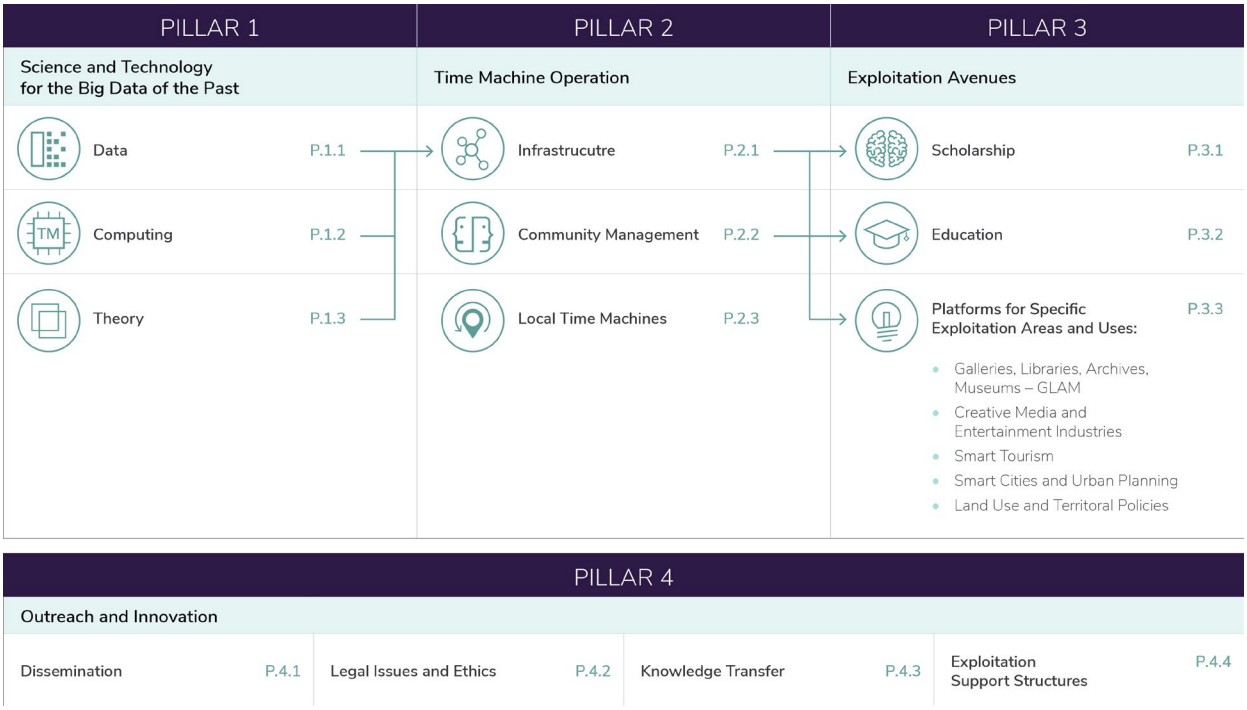

FIGURE 2: The Time Machine Pillars and Thematic Areas and their interrelations

Pillar 1 – Science and Technology for the Big Data of the Past

Addressing the scientific and technological challenges in AI, Robotics and ICT, for developing the Big Data of the Past, while boosting these key enabling technologies in Europe. Pillar 1 adopts a modular, layered structure of interdependent modules, in three directions:

- Data, enabling persistent digital access to millennia of linked historical data
- Computing, developing AI methods to explore, connect, and simulate historical information.
- Theory, focusing on SSH models of historical evidence that lead to new, plausible narratives, radically transforming the manner in which SSH engages with and interfaces with the past

Pillar 2 – Time Machine Operation

Designing the operational infrastructure and the sustainable management model for the creation and extensive use of Time Machine, with particular focus on:

- Building the TM infrastructure for digitisation, processing and simulation
- Drafting the community management systems
- Setting out the principles and processes for a network of Local Time Machines, defined as geographical zones with higher density of “rebuilding-the-past activities”

Pillar 3 – Exploitation Avenues

Creating innovation platforms in promising application areas, by bringing together developers and users to exploit scientific and technological achievements, therefore leveraging the cultural, societal and economic impact of Time Machine. The main areas explored cover:

- Scholarship
- Education
- Specific exploitation areas and uses in key economic sectors, including GLAM, Creative Industries, Smart Tourism, Smart cities & urban planning, and Land Use and Territorial policies

Pillar 4 – Outreach and Innovation

Developing favourable framework conditions for the outreach to all critical target groups, and for guiding and facilitating the uptake of research outcomes produced in the course of the LRSI. The main areas explored cover:

- Dissemination
- Policy, legal issues & ethics
- Knowledge transfer
- Exploitation support structures

This pillar will address requirements related to the development of the other 3 pillars, so its elaboration is planned for Months 5 to 8. For this reason, pillar 4 is not considered in this document.

An overview of progress to be made in each TM area is presented in Table 1 below, while a more detailed discussion of targeted achievements can be found in sections 3, 4 and 5 for the respective pillar.

Thematic Area	State of the Art	With Time Machine
P.1.1 Data	Fragmented datasets only sparsely covering European Cultural Heritage.	Large sets of aligned and standardised Cultural Heritage data.
P.1.2 Computing	Disorganised sets of academic tools most of the time developed for specific projects based on state-of-the-art technology. Very few systems considering the temporal dimension.	Development of a unique computing infrastructure dedicated to the massive extraction of knowledge in Cultural Heritage sources, probably the most advanced artificial intelligence system ever built.
P.1.3 Theory	Digital Humanities analysis of current practices on the role of computing for the Humanities and Social Sciences.	Largest effort ever undertaken to build a critical theory and sound epistemological concept on digital Cultural Heritage and multi-level historical simulation.
P.2.1 Infrastructure	A number of European coordination initiatives of Cultural Heritage with consortia, infrastructures and networks like EUROPEANA, Europa Nostra, DARIAH-ERIC, CLARIN-ERIC, CERL and ICARUS.	Reinforcement and extension of these initiatives with shared computing infrastructure for processing and transforming digital Cultural Heritage at unprecedented scale. Cooperation with initiatives for optimising computing infrastructure, e.g. EuroHPC JU.
P.2.2 Community Management	Many communities dedicated to specific topics – limited exchange due to lack of shared data or concepts.	A large community of communities, sharing a standardised platform, with more empowering tools.
P.2.3 Local Time Machines	Uncoordinated efforts to recreate the past of several cities in Europe.	A franchise-based system enabling each initiative to benefit from the highest level of technology and rapidly develop sustainable models of development.
P.3.1 Scholarship platform	Global commercial search engines are the mediators of our access to knowledge and culture.	Time Machine's open and transparent interfaces, not only analysing the pulsations of the present, but embracing wider geographical and temporal horizons, transforming the way we study, visualise and narrate the past and the future.
P.3.2 Education platform	Books and video about history and culture. Some online lectures.	Massive Open Online Courses, immersive and interactive experiences, engaging material for students and continuous life-long learning. Development of a dynamic new industry for the production of educative digital material, based on aligned massive datasets.
P.3.3 Specific Exploitation and Uses platforms	Cultural Heritage seen as a cost more than a source of innovation, with remaining silos and difficulties to manage solutions truly valorising all available data and across time. Citizens regard policies as a European burden.	Fast development of market-driven platforms of Time Machine to develop relevant services for GLAM[5], the creative industries, smart tourism, smart cities, and land use and territorial policies. Users regard Time Machine as a means to interact with authorities and contribute to policy making.

TABLE 1: Overview of Time Machine expected progress

2.4 Key Concepts

Requests for Comments

Reaching consensus on the technology options to follow in a programme as large as Time Machine is a complex issue. To ensure the open development and evaluation of work, a process inspired by the Request for Comments (RFC) that was used for the development of the Internet protocol[1] will be adapted to the needs of Time Machine. **Time Machine Requests for Comments** will be freely accessible publications, identified with a unique ID, constituting the main process for establishing rules, recommendations, core architectural choices for the Time Machine components. Their basic principles are presented in Box 1.

- **Accessibility** – TM RFC are freely accessible, free of charge.
- **Openness** – Anybody can write a TM RFC.
- **Identification** – Each TM RFC, once published, has a unique ID and no changes are allowed after publication. Any important changes result in a subsequent TM RFC. For this reason, some TM RFCs could be tagged as obsolete.
- **Incrementalism** – Each TM RFC should be useful for its own right and act as a building block to others. Each TM RFC must be aimed as a contribution, extension or revision of the TM infrastructure.
- **Standardisation and linguistic diversity** – TM RFCs should aim to make use of standardised terms to improve the clarity level of its recommendation but can be written in any language. Once published they should be translated in a maximum number of language.
- **Scope** – TM RFCs are designed contribution and implementation solutions solving practical problems. TM RFC are not research papers and may not necessarily contain experimental evidence.
- **Self-defining process** – Like for the development of the Internet, TM RFC could be the main process for establishing TM Rules, TM Recommendations, TM Standard Metrics but also the processes and roles for managing TM RFCs themselves

BOX 1: Basic features of the TM Requests for Comments

To bootstrap the publication process, the initial publication pipeline will follow the following stages:

- Submission of RFC text and figures
- Handling by RFC editors naming RFC reviewers
- Open Review Process
- Attribution of TIME RFC ID and a DOI publication with names of reviewers disclosed and possible additional comments by them
- Translation in several other languages

1 https://en.wikipedia.org/wiki/Request_for_Comments

Time Machine Organisation

The whole governance is conceived around the **Time Machine Organisation (TMO)** that sets the global rules for all actions and operations related to the initiative, including the entire set of processes, labelling system and related infrastructure. The organisational scheme and details of the TMO governance will be specified in the course of the CSA, based organisational requirements formulated in each pillar. For example, an RFC Editorial board will be responsible for the editing process of the TM Requests for Comments mentioned above.

Time Machine Digitisation and Processing Infrastructure

The Time Machine digitisation infrastructure will be composed of a network of digitisation hubs and will be organised on a European scale. A peer-to-peer platform will be in charge of managing and optimising digitisation strategies at European level, and will also be tasked with the development of generic solutions for archiving, directly documenting the digitisation processes, and swiftly putting the digitised documents online. The hubs will cover regional digitisation needs with standardised hardware for digitisation, storage, information exchanges and on-demand scanning, based on results of pillar 1 and existing metadata standards, like the one developed by Europeana.

The peer-to-peer platform will federate system integrators at European level, facilitating the deployment of this equipment. The effort will build upon existing EU Research Infrastructures (DARIAH, CLARIN) and infrastructures providing access to CH (Europeana, Archive Portal Europe, etc.). TM will introduce new processing pipelines for transforming and integrating CH data in such infrastructures.

Documents will be digitised using different kinds of acquisition machines and are treated separately depending on their nature (textual and audio-visual documents, iconographic elements, maps, 3D objects and environments). Information is extracted progressively, either manually or automatically, to produce elementary historical units, connected with one another. This progressive decomposition and refinement needs to be seen not as a mere automatic process but as a collective negotiation. Each intervention, either algorithmic or human, will be fully traceable and reversible. The results of the processing constitute the core dataset of the Big Data of the Past.

The **Time Machine processing infrastructure** is shown in Figure 3. It is composed of a digital content processor and three simulation engines: A 4D simulator, a large-scale inference engine and a universal representation engine:

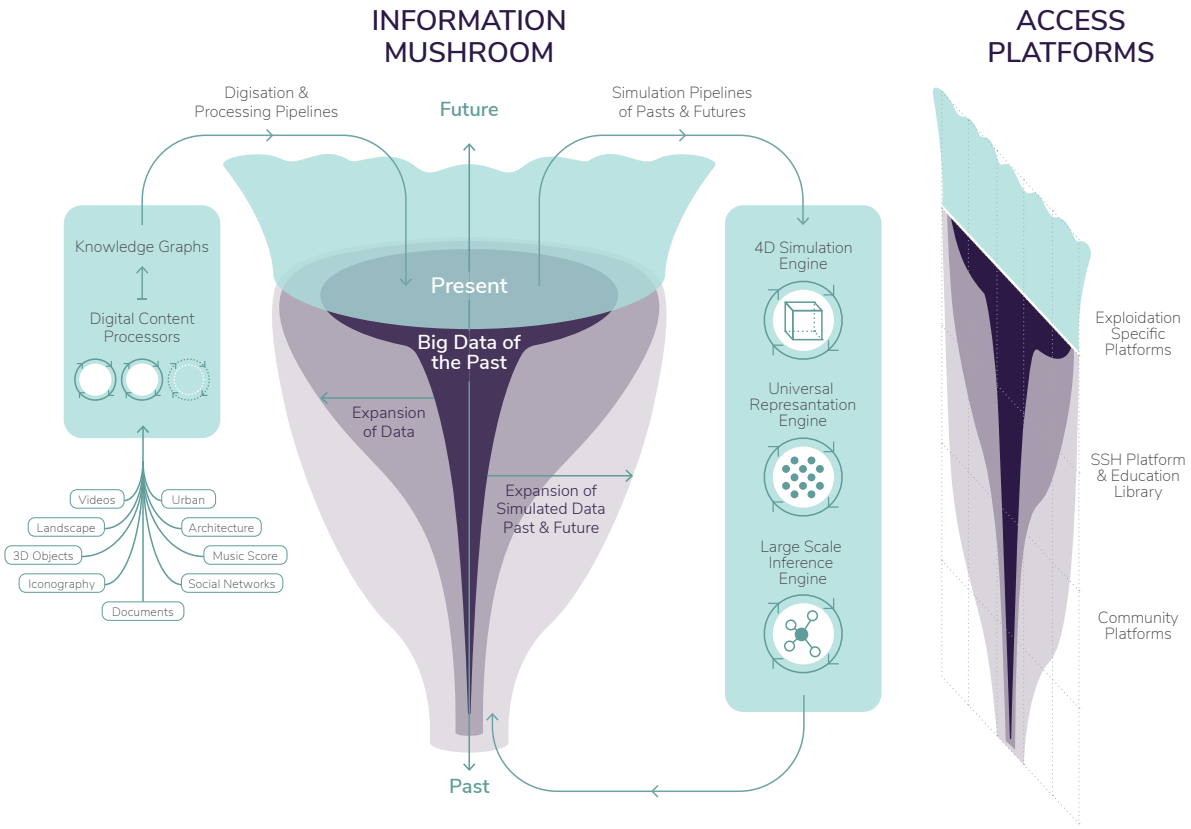

FIGURE 3: Schematic representation of the TM Digital Content Processor and the three simulation engines

- The 4D Simulator manages a continuous spatiotemporal simulation of all possible pasts and futures that are compatible with the data
- The Universal Representation Engine manages a multidimensional representation space resulting from the integration of the pattern of extremely diverse types of digital cultural artefacts (text, images, videos, 3D), and permitting new types of data generation based on transmodal pattern understanding
- The Large-Scale Inference Engine is capable of inferring the consequences of chaining any information in the database. This permits to induce new logical consequences of existing data. The Large-Scale Inference Engine is used to shape and to assess the coherence of the 4D simulations based on human-understandable concepts and constraints

All functions of the different components can be deployed through a fully distributed solution using a storage and computation architecture aimed at an integrated, long-term and sustainable storage of the processed content. This solution embodies our strategy for the long-term availability of processed content, even beyond the lifetime of the organisations hosting it, through predefined and legally binding agreements on licensing, redundant storage, automatic hand-over policies and long-term self-supporting investment initiatives to indefinitely extend the availability of the digitised content of TM.

Local Time Machines

The Time Machine Network is organised as an unlimited amount of **Local Time Machines (LTMs)**. Each LTM is anchored in the space of a city or a region, around which various partnerships can form, aiming to transform it into a zone with a higher density of "rebuilding the-past activities". The TMO provides help in their launch and growth. The Local Time Machines follow the rules of TMO. These are elaborated as a series of RFCs.

In the course of time, Local Time Machines pass through different maturity phases (indicatively: preparatory phase, submission phase, operation phase, with different levels of operational maturity). Each maturity phase permits to envision specific exploitation strategies. For instance, it will only be extremely dense Local Time Machine that can launch Mirror World interfaces.

Each LTM will implement **Projects with Time Machine Label (PWTML)** that aim to increase its density of rebuilding the past activities. Once funding is secured by the partners with, when needed, the help of the TMO, partners of a Local Time Machine can decide to gather around a common goal and create and finance a new PWTML.

As shown in Figure 4, various typologies of partners are central players for the development of both the Time Machine Organisation and the Local Time Machines. We can however distinguish two players' typologies, the actual partners within the TM network (meaning they did ratify and comply to TM's rules and recommendations and a specific "franchise" system), and the external communities that could be engaged with voluntary actions (in the context of a Local Time Machine or identified projects at the TMO level), or provide guidance and advices on specific topics (e.g. professionals communities or cultural-heritage networks).

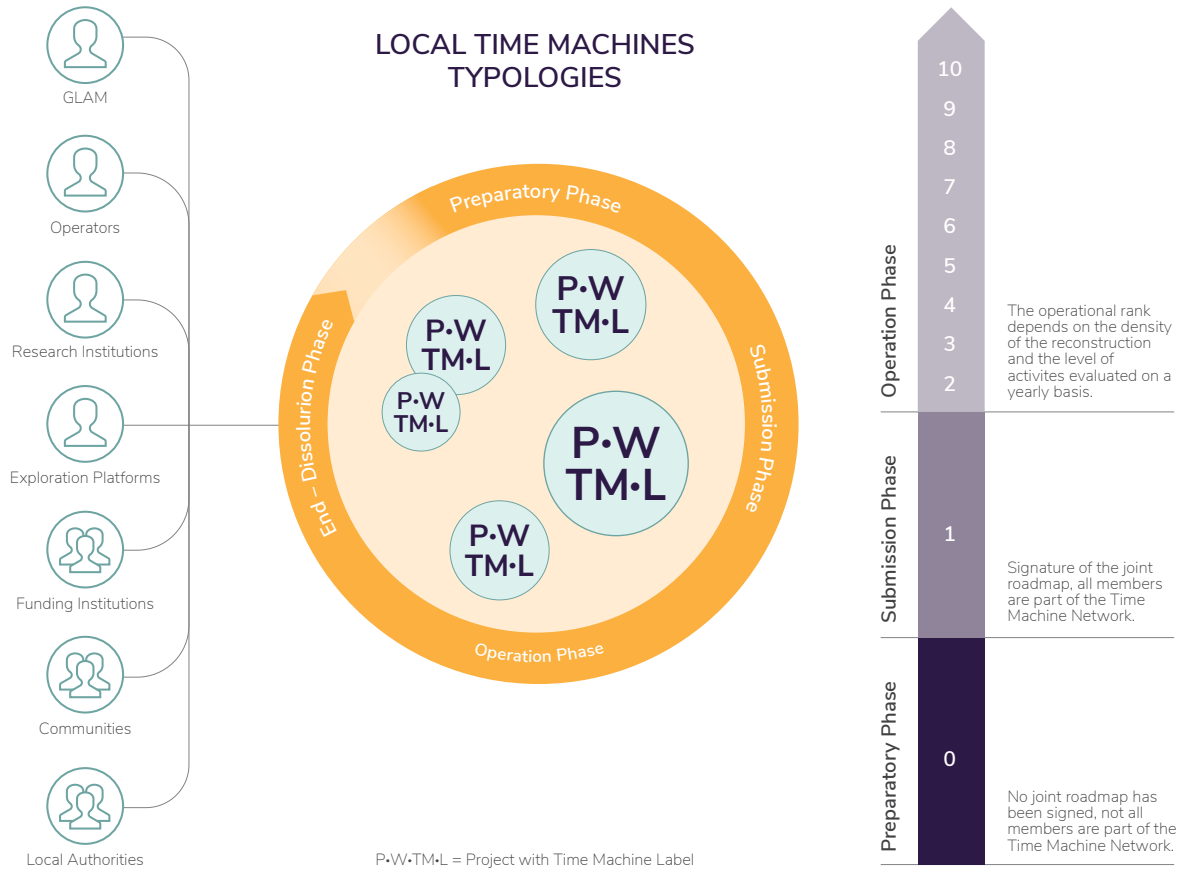

FIGURE 4: Local Time Machines

Following the example of the Venice Time Machine that was launched in 2013, several initiatives have been or are planned to be initiated. These projects are putting together the Big Data of the Past in the following cities (in parentheses are the historical periods covered):

- Venice TM (1000-2000)
- Amsterdam TM (1550-2000)
- Paris TM (1000-2000)
- Jerusalem TM (2000 BCE-2000)
- Budapest TM (1680-1990)
- Regensburg TM (1200-2000)
- Nuremberg TM (1000-2000)
- Dresden TM (1200-2000)
- Antwerp TM (1500-2000)
- Ghent-Bruges TM (800-2000)
- Naples / Campania Region TM (800-2000)
- Lower Austria TM (800-2000)
- Vienna's St Stephen's Cathedral TM (1100-1960)
- Utrecht TM (40-2000)
- Limburg TM (1775-2000)

The TM LSRI will enable these projects to move from their current more or less regional level of operation to a European partnership scale, benefiting from a common infrastructure and framework.

3 THE SCIENCE AND TECHNOLOGY PILLAR

3.1 Objective

The objective of the Science and Technology pillar is to develop cutting-edge computational methods, enhanced with Artificial Intelligence, to access, organise, and understand large-scale cultural heritage collections. This technology will enable virtual time traveling by extracting knowledge and establishing links over space and time. We aim to put together multidisciplinary work groups in Europe to radically transform large-scale humanities studies, (archival) data processing, user interfaces and the way we analyse the past to understand our future.

The logic behind the development of the roadmap involved first critically surveying the state of the art in related scientific domains, then identifying the domain-specific challenges which TM will address. In this way, targeted achievements are identified, along with the proposed methodologies, together with a series of realistic milestones and key performance indicators. In this process we, also, identified topics that need to be developed with strong interactions with the other TM pillars, listed in Box 2.

- **Technical Charter:** Which project-wide formats, protocols and infrastructure for data creation, storage, and exchange must be negotiated by the consortium members?
- **Digitisation Hubs:** Which technology must be developed to realise hotspots of local digitisation initiates, the results of which can be seamlessly aggregated into a pan-European CH data infrastructure?
- **Digital Content Processor:** Which standard processing techniques and generic routines will become available in the next 10 years for the automatic processing of digital CH content?
- **Time Machine Data Graph:** Which scientific and technological challenges must be addressed to realise a distributed super-computing and storage system, which can be efficiently indexed, searched, updated and exploited?

BOX 2: Topics requiring interaction across the TM pillars

3.2 Targeted Achievements

Data (Area 1)

A1.1. Digitisation Hubs

M1.1.1 – RFC for Digitisation Hubs: In order for the Digitisation Hubs to be implemented standards in terms of resolution, file formats, and metadata during acquisition need to be defined. These must be consensual and simple, in order to be easily implemented and fit into existing practices. The RFC also needs to evaluate relevant technologies and recommend affordable technology that does not damage the objects and provide the best possible results. We aim to distribute cheap technology at large scale using e.g. open design hardware. More costly and dedicated scan methods such as scan robots and tomographic methods should be available in dedicated specialised centres spread across the European Union such that their services are available to a maximum number of users. The objective of pillar 3 of achieving cheap and wide-spread digitisation should be a priority in this RFC.

M1.1.2 – Implementation of Digitisation Hubs: The Digitisation Hubs designed according to the results of M1.1.1 will start functioning. In the first stage, we aim predominantly at a wide-spread use of standardised and inexpensive technology. A review process and user consultation should take place 3 months after the beginning and then periodically. Aim of the review is to identify weaknesses in scanning recommendations, hardware, and software post-processing.

M1.1.3 – RFC on new scanning technologies: The cutting-edge technologies like automatic scanning machines with low human supervision, scanning robots and solutions for scanning films and books without the need to unroll/open them need to be considered and fostered by TM. A specific scheme to incentivise these technologies will be created. We aim at an appropriate mix of dedicated specialised scanning centres and development of mobile special use hardware, e.g. mobile CT scanners that are mounted on trucks.

A1.2. TM Box (Servers)

M1.2.1 – RFC for TM Box: The features of the distributed storage system where the Data Graph is to be hosted will be discussed by the community in this milestone. Important issues are the technical server infrastructure, the compliance with international standards, the creation of a system to prove trustworthiness via certification processes, de-duplication methods leveraging pattern-recognition across large datasets, and the implementation of digital observatory and digital archive layers. Also, connection to long-term storage, e.g. DNA storage and selection of the most important data to be stored in such archives is an important design question.

M1.2.2 – Implementation of TM Box: The data up to this point stored in different individual storage systems (LTM and partner institutions) are copied or linked to the TM Box and the correct access is assessed.

A1.3. TM Data Graph

M1.3.1 – RFC on priorities of objects to digitise: Due to the massive amount of European CH and the different states of conservation, availability, and proprietary status, it makes sense to establish criteria to determine priorities of objects to be digitised. The rationale behind this hierarchy can be varied. It could make sense, for example, to offer priority to endangered objects that could be lost in the near future, but maybe accessibility and low difficulty in the digitisation process would provide more material faster and be a more efficient strategy. In any case, the different possibilities need to be carefully evaluated and a plan outlined.

M1.3.2 – RFC on models and formats: Definition of guidelines and standards to follow regarding formats and protocols to store and query data. Proper data management and

curation enable interconnection and involvement of diverse research disciplines and therefore provide excellent environment for boosting innovation. Data will be made available with trustworthiness and FAIR (Findable – Accessible – Interoperable – Reusable) principles in mind. They will include content (primary data), metadata and derivatives (secondary data) as well as externally linked data. Primary Data should be preserved in a Digital Archive with persistent identifiers, usually called a Trusted Digital Repository. Secondary Data should be stored in the research infrastructure with data versioning and full provenance information. Linked Data should be available in Linked Open Data Cloud (LOD). All data should be stored in these FAIR repositories with possibility to get data in and out in the standardised way.

Computing (Area 2)

A2.1. Interface for Annotation

M2.1.1 – User studies of current annotation platforms: Good quality annotation is key to create a linked Data Graph for TM. In order to produce quality human annotations, a proper interface is required. As many such interfaces currently exists, an assessment of the landscape is a prerequisite for the creation of a new TM interface.

M2.1.2 – RFC on interface for annotation: The created interface must allow for easy but complex annotation, that comply with the standards set for data modelling (M.1.3.2). The principles of human-computer interaction (taxonomy task 2.4) and previous user studies (M2.1.1) will inform the development of the annotation tools.

A2.2. User Interface

M2.2.1– User studies of current platforms for historical data: Users of TM will be able to access the data and materials produced by TM through user interfaces. Some of them will be developed by the LTMs for their own purposes, but a central interface, as well as templates for the LTMs that so require, must be elaborated. The first step towards this goal is an assessment of the current interfaces being used in the LTMs and other projects on digitisation of CH.

M2.2.2 – RFC on user interface: The community will propose the features and requirements of the TM user interface. This milestone must be closely coordinated with the work on pillar 3 "Exploitation Avenues", as the interface will be one of the main methods users will interact with TM.

A2.3. Natural Language Processing Tools for Older Language Variants

M2.3.1 – RFC for classification and planning of languages to address: TM will handle documents in multiple European languages and dialects. Some of them might be more complicated to address than others due to pre-existing tools for modern variants or availability of materials. Considering the materials, the places of the LTM and Digitisation Hubs and the features of the languages a working plan of NLP tools development should be conceived.

M2.3.2 – RFC for named entity recognition: Based on the plan outlined in M2.5.1, the community will develop tools for named entity recognition in older European languages and variants. The results of the tagging of entities will feed the **Dark Data Graph**[2] with new information.

M2.3.3 – RFC for orthographic normalisation: Based on the plan outlined in M2.5.1, the community will develop tools for orthographic normalisation of older European language variants. The results will improve the search functionality of the databases and be useful for M2.5.4.

M2.3.4 – RFC for machine translation: Existing algorithms for machine translation will be adapted to older language variants of European languages as outlined in M2.5.1 and taking advantage of the results of M2.5.2 and M2.5.3

2 Please refer to the definition of "Dark Graph" at the end of this document.

A2.4. Digital Content Processor

M2.4.1. – RFC for Digital Content Processor Level 1: Using Machine Learning from existing annotated data, the Digital Content Processor level 1 will be able to label mentions of entities. Results from M2.3.2 will be essential for this development.

M2.4.2. – RFC for Digital Content Processor Level 2: Level two will be able to create labels to establish relationships between entities to create linked data that improves the Data Graph.

M2.4.3. – RFC for Digital Content Processor Level 3: Level three is able to create re-useable models, that generalise from few observations and contribute to possible understanding of the patterns behind the available data.

A2.5. TM Engines

M2.5.1 – RFC for TM APIs: Algorithms and software integrated into Time Machine need to be able to communicate with each other. Thus, definition of joint APIs is required. It is likely, that TM Services are built on top of REST interfaces. In order to match TM's needs these will have to be adopted toward the need of large-scale machine learning. A likely addition is for example the option to provide gradient information of a specific module that is integrated using the API. This way also remote services can be integrated into large-scale training processes.

M2.5.1 – RFC for Large-Scale Inference Engine: The Large-Scale Inference Engine will support fact-based reasoning and logical deduction. It needs to provide a dedicated API to generate new insights from data. Therefore, it needs to enable addition of new evidence, hypothesis checking, and retrieval of related data nodes.

M2.5.2 – RFC for 4D Simulator: The 4D Simulator needs to interface with the TM Data Graph such that virtual worlds can be derived from known evidence. At the first stage, this will imply the loading of pre-configured world models and objects. At later stages, also the generation and adaptation of existing models according to new evidence is in the focus of the 4D Simulator. Also, generation of virtual agent inhabiting the generated world is a focus of the 4D Simulator. An appropriate API modular API has to be designed with respect to these requirements.

M2.5.3 – RFC for Universal Representation Engine: The Universal Representation Engine will make use of the earlier discussed Universal Representation Space. Aim of its use is to support creative and connotative research methodologies. Therefore, APIs have to be defined that enable conversation of e.g. images to text, description to 3D objects, or even maps to virtual cities and vice versa. Again, appropriate design choices have to be made to design APIs and enable intercommunication.

A2.6. Automatic Text Recognition

M2.6.1 – RFC on Text Recognition (1): A call to researchers and TM partners working in Text Recognition will be issued to use the digitised documents of TM to improve the existing models.

M2.6.2 – RFC on Text Recognition (2): Following on the results of M2.4.1 general models for text recognition should be created, i.e. models that work for the largest number of similar documents possible, so that no new models need to be trained to process texts in almost any European script.

A2.7. Automatic Graphic Document Recognition

M2.7.1. – RFC for map recognition (1-2): Using the digitised materials of TM the methods and results of automatic map recognition should be improved. Depending on the results of the first RFC on map recognition, probable another will follow.

M2.7.2. – RFC for music scores recognition (1-2): Using the digitised materials of TM the methods and results of automatic music scores recognition should be improved.

Theory – SSH (Area 3)

A3.1. Increased Acceptance of Quantitative Studies in Historical Research

M3.1.1. – Call for quantitative historical research with TM Data Graph (A1.3): TM should incentivise and create a framework for researchers in historical subjects (history, literature, art, musicology, etc.) to use the TM Data Graph to perform quantitative historical studies as well as facilitating a longue durée perspective. This will be achieved by organising dedicated conferences and open call for papers. There should be at least three organised phases in the strategy to enhance historical research with the Big Data of the Past. The first one will happen before the development of the TM tools for historical research (M3.1.1), in order to better assess the needs of the scholarly community. The second will happen right after the publication of those tools, so that researchers can test and use them. The third call will happen in the final phase of the project.

M3.1.2. – RFC TM tools for historical research: To engage researchers in social sciences and humanities to productively use the Big Data of the Past, TM can offer them a series of tools that facilitate the analysis. These tools will be enhanced by the Digital Content Processor and the Simulation Engines, which will enable scholar to work with historical data in an unprecedented way.

A3.2. Successful Historical Simulations Using the TM Data Graph

M3.2.1. – Call for agent-based simulation using linked data: Agent based simulation can be achieved, in a first stage, using the information from the Data Graph and models from outside TM. A call to teams working on simulation studies that want to test models with information from TM should be issued to assess the quality of the data and the existing models.

M3.2.2. – RFC for improved simulation using TM simulation engines: Researchers will be able to use the TM simulation engines (A2.5) to perform simulations studies, without having to rely on outside models and tools. The simulation engines have the capacity to improve the performance and reach of computational simulations for historical research.

3.3 Proposed Implementation Methodology

As indicated in section 2.4, the main instrument to achieve the proposed objectives is the Requests for Comments (RFC). The RFCs are accompanied by a set of fundamental research questions that need to be clarified by scientific project work (e.g. user studies required to prepare the development of a RFC). These projects need to be funded directly by the Time Machine LSRI. In this category, call for papers and conferences will also enhance the communication and help clarify the goals and assess the progress in different areas.

Once individual RFCs are developed, we propose to implement the roadmap via a modular design in the coming 10 years, in which various calls-for-proposal will attract bottom-up research proposals targeting specific milestones in a pre-specified time-frame. This research initiative can range from macro- to micro-level funding initiatives, from both national/regional/local and European sources.

The TMO will also create a network that enables collaboration of individuals and institutions working towards particular goals. Partner Institutions will develop technologies in their expertise, using TM as a hub fostering their collaboration. The local TMs will use their own institutional infrastructure and take advantage of the general TM architecture as agreed. The

periodical meetings (like the annual TM Conference) will be essential to share and evaluate progress and milestones. TM will receive input and suggestions from the local TM experiences, the users and the academic community, especially in humanities and social sciences in order to understand the necessities and demands of the public.

Key Performance Indicators

The Key Performance Indicators (KPI) are organised according to the taxonomy, although with different levels of granularity. In some areas, each topic is specific enough to have their own KPI, while other areas have only one set of KPIs. At the same time, some topics can be easily measured in many different quantified ways, while others can only be quantified as a vague indicator of the actual state of affairs.

Data Acquisition: Number, diversity, and types of objects digitised and quality of the digitisation.

Data Storage: Data-loss probability; overall system operating cost. Number of total / well-formed and valid file formats within an archive; ratio file formats in archive / available tools for analysis.

Data Modelling: Publication of the TM guidelines for data and metadata as part of the TM Official Components. Scope of the integration with other initiatives. Number of certified digital archives. Speed of development, speed of adoption, percentage of assets making use of models.

Text Recognition: Accuracy in terms of Word/Character Error Rate (CER/WER). Variety in languages and type of documents. Free available tools.

Graphic Document Processing: Accuracy in terms of false positive rate (FPR). Variety type of documents. Free available tools.

Indexing and Retrieval: Number of public and private institutions making their collections searchable. Number of searches carried out by final users on these collections. Number of validated interconnected documents via search engines. Classical performance indicators (precision, recall, mean average precision, etc.) on cross domain and multimodal collections. Performance indicators versus required memory and search time. User studies.

Understanding and Interpretation: Accuracy and AUC for classification. Recall@{1,5,10} for metric learning and localisation. Distance in meters for localisation. User studies.

Recognition and Detection: For classification, accuracy and AUC; for detection, average precision. Intersection over Union (IoU).

Person & Face Identification: Face detection performance in different content domains (as precision/recall, MAP) compared to human (in identification and verification tasks). Face recognition performance in different content domains, across persons' lifetimes (as precision/recall, MAP).

Audio Recognition and Transcription: WER for speech recognition, Number of institutions and media providers that make their archives searchable. Number of searches carried out by the final users of the archives. Number of enriched archives.

Machine Learning and AI: Speed and efficiency of technologies. Performance on large-scale benchmarks. CH bots accuracy in human understanding, language generation and human understanding. Avoidance of biases. User studies.

Computer Graphics: Faithful Renderings of historic artefacts in their original context, in real-time and thus applicable for VR and AR. Quality of visualisation, supported platforms. Tracking offset. Perceived lighting artefacts. User studies.

Natural Language Processing: Error rate of methods (accuracy, F1 score, BLUE scores, etc.). Language and variants where they are effective. User studies.

Human-Computer Interaction and Visualisation: User studies. Number of users of TM interfaces.

Humanities and Social Sciences: Engagement of academia and research (Bibliometrics, Alt Metrics) with TM through mentions in journals and books, initiatives and projects using the TM data or infrastructure.

Funding Sources

Many of the technologies presented in our state of the art analysis are already being developed using a variety of funding sources in the involved institutions from European, national, and industrial resources. Yet, none of the funding sources is able to support a large-scale project such as Time Machine. Furthermore, as project-driven research is typically limited in budget and time, there is no readily available mechanism for projects to implement TM APIs. A central, large-scale funding mechanism is clearly required to implement Time Machine as a whole.

In order to increase the impact of TM and to create synergies with the many European efforts already in place, we propose that TM provides "on top funding" to otherwise supported projects, enabling them to develop the correct implementation of APIs and methodologies proposed in the RFCs. Additional TM funding can be linked to preferred data and license models enforcing open access policies.

Using the proposed "on top funding" methodology, research labs, universities, and private companies are incentivised to integrate in the Time Machine project. Doing so, enables to combine developments already in progress by their existing funding scheme into the grand vision of Time Machine. Yet, to fully develop the required technologies for the Big Data of the Past, a series of specific funding for the development of RFCs, user studies, and light-house projects is needed. This can only be achieved using a large-scale research initiative.

4 THE TIME MACHINE OPERATION PILLAR

4.1 Objective

The TM Operation pillar has 3 main directions each dealing with the development of one of: the TM infrastructure, the TM community and the Local Time Machine network.

For infrastructure, the aim is to develop:

- A digitisation-hubs network that will facilitate and standardise digitisation processes at European scale
- A distributed super computing infrastructure for processing the Big Data of the Past
- A distributed storage system

Regarding the TM community, it should be stressed that the massive involvement of scholars, developers, cultural heritage professionals, service providers and citizens constitutes a key success criterion for the overall initiative. The main idea is that a system of platforms will connect Time Machine with such external communities that will benefit from and can provide input in various forms to Time Machine. The aim of the thematic areas is, therefore, to build a strategy for Community management and an associated Community Management System, responding to well identified requirements for staff and processes leading to mutually beneficial and sustainable interactions with the TM communities.

The LTMs are meant to be at the centre of the TM network, rallying different partners behind the TM objectives and helping the growth of both TM Data and Operation Graph. Shaped as independent networks of projects, they're still active within a common global TM network, and as such, benefit from the TM infrastructure and contribute to its development. Finding the proper balance between TM network governance and the LTM activities is an important factor in the TM development.

4.2 Targeted Achievements

Overarching Achievement: TM Request for Comments

TM Request for Comments: Publication and reviewing process for TM infrastructure and Local Time Machine is incrementally defined

As the development of the TM operations is an iterative process, the entire step-by-step negotiation will be done through the writing, editing and publication of TM Request for Comments, inspired by the process by which the Internet Society establish its international standards. Establishing this process is one of most urgent achievement as it organises the process for all the others infrastructure developments.

Infrastructure

Digitisation Hubs: Facilitate, standardise and optimise digitisation process at European Scale.

The TM digitisation infrastructure will be composed of a network of digitisation hubs and will be organised on a European scale. A peer-to-peer platform will be in charge of managing and optimising digitisation strategies at European level, and will also be tasked with the development of generic solutions for archiving, directly documenting the digitisation processes, and swiftly putting the digitised documents online.

Processing and Simulation: Development of distributed super computing infrastructure for processing Big Data of the Past.

TM Super Computing Architecture is composed of distributed computing resources for Digital Content Processor and at least three simulation engines, will allow users to generate possible pasts and futures from the TM Data Graph.

Virtual/Augmented Reality and Discovery: Development of innovative infrastructure grounding for new experiences based on the TM Data Graph.

The data generated by Time Machine, will allow creation of new experience for the users, both in term of content and discovery interfaces.

Distributed Storage: Development of the distributed storage infrastructure for both public and private data

TM Distributed storage infrastructure will be a fully decentralised highly redundant architecture based on the shared resources of **TM Infrastructure Alliance**[3] partners. Its main aim is to store the core public datasets of the project (referred to as **The Great Commons**) and possibly private datasets associated to TM.

Communities

Community Interface: Design TM infrastructure enabling its interfacing with existing communities

The different parts of the TM infrastructure, for instance the source code powering the various platforms, the design of software portals for crowdsourcing or the reference mechanism for scholarly articles will be designed in a way that enables the connection to existing community infrastructures. For example, source code will be published on GitHub to enable direct contact to open source developers and user accounts for crowdsourcing volunteers can be connected to other social media or community accounts. Data can be shared.

Community Inclusion: Tailor communication to various communities

TM includes external communities in communication efforts of various milestones and designs specific communication activities tailored to different communities. For instance, descriptions of new software releases, internal architecture or the selection of various technologies is communicated via articles, forum posts or similar means to developer communities. Hackathons with the TM APIs will be organised. Success or post-mortem reports will be produced on crowdsourcing activities. Guest articles by scholars involved in TM are invited and published in the appropriate TM communication channels.

Community Statistics: Enable tracking of community involvement

To enable tracking of and reporting on community efforts, mechanisms to collect statistical data on individual participation, impact on the TM data and similar metrics are included in appropriate parts of the TM infrastructure, mainly user accounts and contribution transparency.

Local Time Machines

TM's Rules and Recommendations: Ensure cohesion and network's operation by the implementation of general values and technical standards.

Gathering all data generated by a digitisation process or already existent ones, will not be possible without agreement on minimal technical specifications, bounded with technical needs of the TM infrastructure. Defining a common framework and a proper starter kit implies also to set clearly some values and perimeters criteria aiming to support common objectives and regulate data acquisition, data sharing and data publishing.

3 Please refer to the definition for "TM Infratrstucture Alliance" at the end of this document.

Framework: Set up a support structure aimed at launching an LTM and ensuring a regular and smooth development of the TM network.

Any partners wishing to integrate or launch an LTM should be redirected to a clear path and find requested guidance during the whole process. As active partners, they would integrate a network and should therefore respect the TM rules and TM recommendations.

Labelling System: Set up a value scale in order to precisely evaluate the progression of an LTM and support the TM network's dynamics.

How to encourage any LTM to enter in the operational phase and contribute to the growth of the TM Data and Operations Graphs, and how best guarantee member's commitments towards TM's objectives? The labelling system should contribute to answer such issues.

Legal Setting: Guarantee the respect of national and European policies and laws in order to build a coherent and standardised contractual and licensing system for TM's network operations aiming to regulate data acquisition, data sharing and data publishing.

Many trade agreements are expected within the TM network. Standardisation will prevent malicious use, while easing the processes, and guaranteeing conformity with national and EU requirements.

Financial System: Foster financial independence of all TM's instances.

Funding sustainability plays a central role in the building of any PWTML, LTM's partners would need to conduct efficient financial research, and the services offered by the TMO (infrastructure and coordination) will also require support. A franchise model will be shaped consequently.

Incorporation, Collaboration and Enhancement: Enabling existing initiatives to be aligned and consolidate the global network by encouraging connections amongst TM's partners. Contribute to the future development of the network by supporting the creation of new LTMs (top-down process). Ensure knowledge transfer amongst the partners.

Existent LTM and PWTML initiatives will be included in the TM's network. Global heritage networks are partners of the TMO and therefore should be invited to contribute in the network, and share their expertise and best-practices, enhancing a global European collaboration system. Europeana in particular, will serve as backbone to Time Machine, guarantying the use of pan-European standards and international solutions for cataloguing and identifying the digital resources. Top-down initiatives will also contribute to the system's growth, and as it does, knowledge and technical achievements will be shared amongst network's partners.

Smart Cluster: Ensure the identification and future creation of local competitive assets.

Once data gathered within the Big Data of the Past Graph, how to encourage future exploitation avenues to identify opportunities, build on their regional particularities and shape innovative cultural-heritage experience platforms and services?

4.3 Proposed Implementation Methodology

The pillar 2 implementation is based on milestones to be achieved as indicated in table 2 below.

Milestone	Due year	Means of verification
RFC Coordination	1	Each identified RFC have been assigned a specific coordinator by the RFC Committee.
Framework: Definition	1	Close of the TM RFC on LTM/Framework (e.g. typologies and routines, perimeter).
Labelling system: Definition	1	Close of the TM RFC on Value scale systems. V1 of the value scale systems for both LTM and PWTML.
TM's Rules and Recommendations: Time Machine Box, development	1	Close of the TM RFC on Time Machine Box. Prototype V1.
Framework: Documentation	1	Close of the documentation and trainings phase for the implementation of the framework, including welcome guide, starter kit, LTM and PWTML's forms.
TM's Rules and Recommendations and Legal settings: Models	1	Close of TM RFC on Vision Mission and Values Charter, TM RFC on Technical Charter, TM RFC on data lifecycle, TM RFC on intellectual property rights and licenses, and contractual documents phase. V1 of the respective documents.
TM's Rules and Recommendations: Time Machine Box, assessment	2	End of the evaluation of the technologies. When needed, new RFC.
TM's Rules and Recommendations and Legal settings: Roll-out	2	Close of the test phase for implementation with a representative percent of the TM partners aligning with its processes and prerequisites and making feedback to the development team.
Labelling system: Assessment	2	End of the evaluation of the system. When needed, new RFC.
Financial system: Definition	2	Close of the TM RFC on Franchise system, franchise V1. And close of the TM RFC on Training.
TM's Rules and Recommendations and legal settings: Adaptation	2	According to the result of the roll-out phase: Data selection Model, Data acquisition Model, Data Sharing Model, Vision Mission and Values Charter, Technical Charter, contractual documents, decision on opening a new RFC.
Financial system: Assessment	2	End of the evaluation of the system. When needed, new RFC.
Standardisation of the TM network	3	Launch of the first 10 LTM officially complying with the TM's Framework and TM's Rules and Recommendations and Legal settings, advertisement and tracking on the TM's Operation Graph and on the TM' networks landing pages (use of the proper labels).
Standardisation of the PWTMLs	3	Launch of the first 10 PWTMLs officially complying with the TM's Framework and TM's Rules and Recommendations and Legal settings, advertisement and tracking on the TM's Operation Graph and on the TM networks landing pages (use of the proper labels), payment of the proper franchise fee.
Alignment	4	All TM partners now comply with TM's Rules and Recommendations and Legal settings (comprises alignment of previous initiatives).

Milestone	Due year	Means of verification
TM Cooperation	4	Close of the TM RFC on Solidarity, TM RFC on Enhancing collaboration, TM RFC on Knowledge transfer, TM RFC on Top-Down initiatives, TM RFC on Smart Cluster. V1 of the respective models.
Community – Density growth	4	Assessment of the network's growth performances in terms of partners actively involved within the TM, number of LTMs and PWTMLs launched and density of "rebuilding the past activities" in a geographical area. Definition of supportive top-down initiatives or new RFCs.
Assessment	5	Evaluation of TM's Framework, Labelling system, Rules and Recommendation, Legal settings, Financial system accuracy and suitability with TM needs. Decision on updating initiatives or new RFCs.
Community – Density growth	5	Assessment of the network's growth performances in terms of partners actively involved within the TM, number of LTMs and PWTMLs launched and density of "rebuilding the past activities" in a geographical area. Definition of supportive top-down initiatives or new RFCs.
Assessment	6	Evaluation of TM's Framework, Labelling system, Rules and Recommendation, Legal settings, Financial system accuracy and suitability with TM needs. Decision on updating initiatives or new RFCs.
Community – Density growth	6	Assessment of the network's growth performances in terms of partners actively involved within TM, number of LTMs and PWTMLs launched and density of "rebuilding the past activities" in a geographical area. Definition of supportive top-down initiatives or new RFCs.
Assessment	7	Evaluation of TM's Framework, Labelling system, Rules and Recommendation, Legal settings, Financial system accuracy and suitability with TM needs. Decision on updating initiatives or new RFCs.
Community – Density growth	7	Assessment of the network's growth performances in terms of partners actively involved within TM, number of LTMs and PWTMLs launched and density of "rebuilding the past activities" in a geographical area. Definition of supportive top-down initiatives or new RFCs.
Assessment	8	Evaluation of TM's Framework, Labelling system, Rules and Recommendation, Legal settings, Financial system accuracy and suitability with TM needs. Decision on updating initiatives or new RFCs.
Community – Density growth	8	Assessment of the network's growth performances in terms of partners actively involved within TM, number of LTMs and PWTMLs launched and density of "rebuilding the past activities" in a geographical area. Definition of supportive top-down initiatives or new RFCs.
Assessment	9	Evaluation of TM's Framework, Labelling system, Rules and Recommendation, Legal settings, Financial system accuracy and suitability with TM needs. Decision on updating initiatives or new RFCs.
Community – Density growth	8	Assessment of the network's growth performances in terms of partners actively involved within TM, number of LTMs and PWTMLs launched and density of "rebuilding the past activities" in a geographical area. Definition of supportive top-down initiatives or new RFCs.

Milestone	Due year	Means of verification
Assessment	10	Evaluation of TM's Framework, Labelling system, Rules and Recommendation, Legal settings, Financial system accuracy and suitability with TM needs. Decision on updating initiatives or new RFCs.
Community – Density growth	10	Assessment of the network's growth performances in terms of partners actively involved within TM, number of LTMs and PWTMLs launched and density of "rebuilding the past activities" in a geographical area. Definition of supportive top-down initiatives or new RFCs.

TABLE 2: Implementation methodology for the TM operation

Key Performance Indicators

Specific Key Performance Indicators will be used for monitoring progress in each thematic area. As one of the overall objectives of Time Machine is to increase the amount of cultural-heritage digitised datasets available across EU, the TM KPI system will also use KPIs already put in place to monitor the progress on digitisation of cultural heritage across Europe (global study "Enumerate" by Europeana) such as:

- Presence of a digitisation strategy
- Presence of digital heritage collections
- Necessity to reproduce analogue heritage collection in digital format (per object type)
- Copyright conditions

Infrastructure

- Number of units and relation in the TM Data Graph (measured by typologies in Bright and Dark Graphs).
- Number of rules in the Large-Scale Inference engines (measured by typologies of rules)
- Number of digitised images
- Number of digitised artefacts
- Number of digitised sites
- Number of linked resources

Communities

General:

- Community outreach efforts (AMAs, Blogposts, Hackathons, Giveaways)

Developers:

- Opened Issues/Commits/Forks/Pull Requests of Open Source Developers to TM Repositories
- Stars given to TM Repositories
- Known reuse of TM Software and developed Standards

Educators:

- Lectures/Courses given with use of TM Data

Professionals:

- Numbers of professional organisations that join the Time Machine Organisation
- Number of agreed policy measures about exchange of cultural heritage
- *Scholars*
- Scientific articles using TM Data or about TM

Volunteers:

- Gamification of Crowdsourcing
- Members in the TM Portals

Local Time Machines

Collaboration:

- Average of partners involved per PWTML within an LTM
- Contentment of the various partners regarding collaboration
- Capacity of the collaboration in place to enhance the meet of project's goals
- Degree of partners' participation towards decision-making
- Perception of the different partners regarding collaboration
- Members participation's rate at project's stakeholders meeting
- Mix of line businesses involved in the project

Growth of the TM network:

- Number of new partners of the TMO (measured per membership's categories)
- Number of new LTM
- Number of new PWTMLs (in total)
- Activity of an LTM (number of PWTML launched over a period of time)
- Engagement rates: Active contributors in the TM Networks

Labelling system:

- Label adoption (number of labelled LTM and PWTML per label's categories)

Guidance and Training:

- How many new requests are coming in?
- Attendance: How many people are coming to special events-formations?
- Training completion, percentage rate

Funding Sources

For **Infrastructure**, there are two main fields of activity, one dealing with public data (referred to as Great Commons) and one dealing with private data (referred to as Private Time Machine)

Great Commons: the main funding source would be in-kind contributions for both storage and computing powers. Institutions, industries, wishing to give servers' space running with the chosen distributed storage protocol or computing power for TMO technologies will be welcomed as members of a Time Machine Infrastructure Alliance.

Private Time Machine: Industries, institutions that cannot expose all (some of them might still feed TM pipelines) their data under an open – creative commons licence yet, will benefit from TM infrastructures, on payment of the service's costs. As those data are probably more sensitive (e.g. private, financial, bank information), specific security means will need to be deployed (such as cryptography, authenticated access, erasing data right etc.). The TMO will encourage a dedicated and regional network of operators to provide required technical components and ensure security and quality criteria, and as the coordinator of this network, will perceived royalties.

For **Communities Management**, regional or EU funding can be used for the development of the RFCs and the launch of put projects.

For **Local Time Machines**, before becoming an official partner of an LTM, candidate institutions should become an official member of the TMO and therefore chose amongst memberships categories and their related fees (currently under development) for the different services offered by the TMO. This financial system to take place between the TMO and the LTM will be further developed within the TM RFC on Franchise aiming to clarify usage infrastructure and services costs.

LTMs and PWTML will operate essentially with public and/or private funding that relates to regional development. An example of public funding is through the current European Structural and Investment Fund (ESIF) instruments, for which LTMs and PWTLMs offer concrete projects ideas with very focused development goals to be supported. The PWTMLs are core to the operation phase of an LTM. All technical operations are meant to happened within a PWTML. When launching a PWTML all partners should ensure its financial sustainability and provide a financial plan, verified and validated by the TMO. They can be helped in the process by the TMO for identifying a funding strategy and in developing grant application.

The creation of smart clusters will enable the LTM to create job opportunities within both local industries and exploitation avenues and contribute to the overall sustainability of the TM network. An ambitious turning point in the business model of the LTMs will be when all actors of local regional development recognise the role of LTMs as powerful motors for growth, which will open the way for new models of sustainable operation, first for the nodes and then for the whole pan-European LTM Network. This aspect will be further elaborated in the next months, before the finalisation of pillars 2 and 3.

5 THE EXPLOITATION AVENUES PILLAR

5.1 Objective

The roadmap for pillar 3 seeks to demonstrate how the scientific & technological advances (pillar 1) and operational models (pillar 2) enable us to work towards the vision of developing the Big Data of the Past, and in turn how this approach can provide social and economic impact across a range of areas, comprising:

- Scholarship
- Education
- Specific exploitation areas and uses, including GLAM, Creative Industries, Smart Tourism, Smart Cities & Urban Planning, and Land Use and Territorial policies

Common denominators facilitating significant Impact to be made in these areas are the following enabling factors that will be created by the TM LRSI:

- **Cheap Digitisation:** Enabling the provision of cheap and cost-efficient solutions for the further digitisation of resources through standardised offers and services and easily replicable open hardware technologies.
- **Generic Automation:** Enabling the automation of the mark-up of these resources tagging concepts, named-entities, relations and rules.
- **Connection:** Facilitating the intelligent connection of existing fragmented data resources using, adopting and building on existing legal frameworks and developing standards for distributed storage solutions.
- **Simulation:** Transforming sparse data into continuous 4D representations capable of representing multi-worlds.
- **Experience:** Enabling new paradigms for the restitution of the data to the end-user including spatio-temporal search engines, geo-historical services and Mirror Worlds.

5.2 Scholarship

Time Machine will make available a macroscopic observatory for cultural history, where cultural production can be studied as a four-dimensional macro-object, as opposed to the microscopic scale which is nowadays common in SSH. TM thus has the potential to realise a radical expansion of the “zooming” capabilities of scientific research: from the microscopic level of historical anecdote to the macroscopic level of high-level cultural patterns and their interrelations with socio-economic trends.

TM for Scholarship platform will host projects that allow researchers to navigate history in radically new ways, ranging from:

- “Google Street View”-like augmented reality experiences
- Agent-based simulations of large-scale virtual re-enactments of major historical events
- Ontological aggregators which extract multiple layers of metadata and content to cross-reference multiple sources for trustworthy scholarship

These interfaces will allow accurate modelling procedures with a longue durée perspective that opens the way for new and more critical methods of analysis, with the aim of developing meaningful outlooks for the future.

Targeted Achievements

- First and foremost: the amount of connected, multimodal data and its accessibility in/through TM will in itself reduce the limitations in a wide array of disciplines that employ historical data (most SSH fields) related to the fragmentation of the data and tooling. As such, **TM will provide SSH scholars with a much more comprehensive knowledge base for the study of longitudinal trends across various sectors and locations.**
- Because of its integrated approach (i.e. digitisation, interpretation & data gathering and management, as well as methodological innovation), **the TM infrastructure will drastically speed up advances in the state of the art in most SSH fields (and probably also in ICT). The pace and explanatory power of scholarly & scientific progress will multiply.** This innovation will be supported by the conceptual and methodological framework developed in pillar 1, for SSH research that combines the strengths of the tradition of hermeneutic research (interpreting the complexity of human culture and society at the microlevel of individual sources, places, people or events) with the advantages of quantitative methods (seeing patterns in large datasets and analysing those with statistical methods). This 'scalable' approach to SSH research methods will innovate scholarship in three ways:
    1. The methods can be used heuristically, whereby the patterns observed lead to new hypotheses on the phenomenon under investigation, that then subsequently are analysed with traditional, interpretative methods
    2. The analyses based on Big Data of the Past can be used to empirically test existing assumptions based on smaller, sample data
    3. The quantitative methods (including simulation) allow for the combination of different types of data and thus for more complex analyses
- The location-based approach of TM allows scholars to query datasets of different origins and types, and relating historical data to present-day data on a particular location. **This facilitates research geared towards present-day societal challenges**, providing those challenges with a longitudinal perspective on their historical roots. **This will significantly boost the societal value of humanities research**
- **Conjectural prediction of the future becomes possible** because of huge training data for the prediction algorithms that TM will provide, and which can be tested through predictions based on past data (then we can check the results against the existing historical knowledge of the past). Simulation of alternatives past(s) becomes possible because of vast data and trained simulation engines
- Sustainable access to tools, that are also maintained – focus on specific software solutions that are also maintained by communities of users

Specific strengths of TM in boosting scholarship include:

- The integration of existing fragmented resources (examples include the architectural images and databases like SAHARA, ArtStor, geospatial databases)
- International reach (our project focused more clearly on connecting historical information in a supra-national way)
- The localisation of historical information (ability to use the sources in relation to the locations to which they relate, which now has to be manually added)
- Direct access to non-textual content
- Infrastructures for integrated research with multimodal sources (text, images, AV, 3/4D etc.)
- Implementation MethodologyApart from the obvious R&D challenges in creating a

user-friendly Big Data of the Past (pillar 1), for uptake and exploitation in Scholarship, **accessibility and awareness are probably the most important challenges,** because progress against the state of the art methods will strongly depend on the actual use of the new methods and tools by the largest number of researchers

In order to achieve the foreseen impact on SSH scholarship, we will adopt the following strategy:

1. Conduct a number of **predefined use cases**, that focus on the value of a longitudinal, comparative perspective on present-day societal challenges and serve to advance and validate the Time Machine platform for scholarship
2. On the basis of the use cases, we will generate a **set of best practices and training and dissemination material for SSH scholars,** that we will distribute via the relevant research infrastructures (DARIAH and CLARIN, as well as their national representations (e.g. CLARIAH-NL, CLARIAH-LUX, CLARIAH-VL, CLADA-BG); EHRI; E-RHIS) and via domain-specific professional organisations (including papers and workshops at their annual conferences, e.g. the European Urban History Association, Digital Humanities, etc.)
3. Subsequently, we will issue two rounds of **open calls for proposals to test and validate the platform** and its services and tooling, and to open up the project to new stakeholders

The use cases are large-scale, **long-term (*longue durée*) and comparative research projects, transcending existing subfields and focusing on the relations between economy, society and culture.** They innovate existing SSH scholarship by making it comparative in nature, extending its scope in four dimensions:

- Time (enlarging temporal scope)
- Space (enlarging geographical scope)
- Disciplines (combining methods and tools from SSH, Computer Science and other relevant disciplines)
- Sources (combining different data types)

This extended scope allows scholars to study present-day societal phenomena and challenges in their inherent complexity, considering their economic, social, legal, technical and cultural dimensions over time, across space. Possible topics include:

- The role of local cultural values for belonging and social cohesion
- Democracy and democratic values
- Welfare and wellbeing
- Financial markets and crises
- Security
- Populism
- Migration and social and cultural integration
- Climate change and environmental issues
- Adoption and impact of new technologies

The selection of use cases will be prepared by relevant expert groups and proposed to Time Machine Organisation for approval. The open calls can be on any topic but have to use one or more components of the Time Machine Toolbox, in order to ensure that these are tested and improved with the results of the projects. Each round of projects is followed by thorough evaluation and implementation of the results.

Milestones

Milestone	Year	Description
M1	2020	Use cases selected
M2	2020-2023	Use cases executed
M3	2023-2025	Development of best practices and training and dissemination material
M4	2024	First open call for proposals
M5	2027	Evaluation CfP projects round 1 & implementation results
M6	2027	Second open call for proposals
M7	2030	Evaluation CfP projects round 2 & implementation results

Key Performance Indicators

Social impact

This is the most relevant impact dimension for scholarship that could be assessed by looking at:

- Usage of the TM data and tooling (user statistics on TM infrastructure; papers; publications and other scholarly output)
- Integration of TM data and tools in higher education curricula in SSH field (e.g. via the CLARIN-DARIAH Digital Humanities Course Registry[4])
- Monitoring the impact of SSH publications on crucial topics (e.g. as evidenced by attention in the media)

Economic impact

Could be measured by looking at increase in number of startups that initiate in the SSH field (e.g. via Venture Labs as this one at UvA in Amsterdam[5]).

Funding Sources

The proposed match-making system to stimulate SSH scholars to include the TM data and infrastructure in their research projects and funding applications (with the purpose of obtaining requirements for further developing the infrastructure) can be realised by issuing calls from European/national funds, and/or by convincing funders to make use of TM data and/or tools a requirement for certain calls.

5.3 Education

Current challenges in the field of education comprise inclusiveness and life-long learning, designing customised educational material, information competency as well as an increasing complexity and change rate of knowledge intensive labour requirements. Facing these challenges, the TM platform for Education will offer unique enquiry and experience-based blended learning, citizen science infrastructure and approaches based on revolutionary digital technologies (VR, AR, AI).

TM for Education will focus on developing pedagogical content for schools, universities, and lifelong learning in a mix of free, sponsored and paid services. Content will be largely based on the Big Data of the Past and associated simulation technologies. Important objectives will

4 https://registries.clarin-dariah.eu/courses/

5 https://www.uva.nl/en/faculty/faculty-of-humanities/humanities-in-the-city/humanities-lab-avs/humanities-lab-avs.html

be to accelerate the learning of SSH, through swift availability of many facts on a single subject, as well as emphasis on epistemological and methodological issues and critical analysis. Students will be in position to study complex societal and urban challenges and thus to learn informed decision-making, considering and balancing relevant facts, interests, values, costs and benefits.

Teaching and research is also considered as a third exploitation domain (after private use and professional use), for which TM can enrich teaching material associated to SSH, the sciences, health and practical technologies.

Targeted Achievements and Implementation Methodology

The state of the art does not point so much to a gap in terms of availability of software in education, but more so in terms of the efficiency of such software.

On the one hand, popular reference websites are used abundantly, both in and outside of formal education environments. On the other, there are complex and exhaustive software packages that are designed for education environments with the intent to as if replicate or extend certain aspects of education processes and activities, including studying, communicating with students and staff, applying one's analytical and critical skills through computational techniques, or extending lab settings through virtual experimental environments, such as in VR or various applications of 3D models.

The exact procedure and method to innovate the state of the art is outlined below, followed by an indication of the utilities this will yield.

The objective of Time Machine in education is to **develop pedagogical content for different groups and institutions.** Such content could complement existing curricula with additional data for history and history-based courses; offer students analytical tools and big data of the past analysis training; and, by extension, offer students and users at large with seamless and integrated access to historical data.

Time Machine's personalised, localised access to the Big Data of the Past is ideally positioned for the current trend towards more self-directed learning, whereby the nature of the instructor shifts 'from transmitter of knowledge to facilitator and curator' (EDUCAUSE Horizon 2019, 19). From that perspective, it makes sense to design an infrastructure that provides direct access to the TM data in ways that match the infrastructures for education currently in use and the new opportunities for innovating access to cultural and historical information.

At the same time, **we need public and private partners to develop and maintain (technically and in providing service) the services that provide access to the TM data. The Time Machine also needs to invest in engagement of the educational communities, raising awareness of the potential of TM data** (by creating showcases and best practices) and facilitating interaction (iterative, user-driven, inclusive, value-sensitive co-design approach to researching and developing the infrastructure).

Given the main objectives of Time Machine's approach to education, there is a need of intense cooperation with education professionals and education certifying bodies in these fields. To this end, **Time Machine aims to apply pilot projects in a select number of institutions from a representative sample of the European education landscape, including primary and secondary schools, technical schools, universities and other. Via its Community Interfaces (pillar 2) Time Machine will allow us to also test the use of TM in informal learning**. For each of these four areas of learning, we will start pilot projects consisting in experimenting and monitoring the uses of Time Machine components, including historical data, training in analysing big data of the past, accompanying analytical software and training in using and developing Time Machine interfaces for Big Data of the Past.

As Time Machine is centred on the use and application of big data of the past, it aims to consolidate the above with **pilot projects primary schools, secondary schools, higher education and local volunteer-expert community groups within pilot project, focused on three core aspects of the educational potential of Time Machine:**

1. **Encyclopaedic use: granting students and educators at large with access to Big Data of the Past through web-based reference techniques,** such as a "History Look Up" function that can be activated to consult historical background information about various (or any) information students encounter. Here, the envisioned application is intended to be general and seamlessly integrated in existing and habitual reference consultation practices
2. **Engaging explorations of and experiences with the past: providing students and educators with specific applications and interfaces through which to make use and visualise big data of the past, including the simulation of those pasts using advanced visualisation techniques.** These applications can include maps with integrated 3D models, AR/VR applications, search engines and other information systems based upon big data of the past. While these applications are not exclusive to students, they constitute specific pedagogical approaches to introduce students into historical data analysis and application development
3. **Critical thinking and digital literacy: supporting these applications are code and big data analysis training, or "Time Machine analytics".** Such training is not intended to be specific or exclusive to Time Machine, but to all students and educators engaged in studying and teaching historical disciplines through data analysis. Time Machine can offer use cases and material for educators to teach students how to study history with respect to such data – implying, here, that they are also be offered material on how to teach historical data analysis. Critical thinking and digital literacy required for using such data can be developed in cooperation with pillar 1 – Theory

Our strategy is to develop four pilot projects around these three core aspects, with each pilot addressing a specific level of education. The pilots address either:

- One level of education with multiple aspects
- Multiple levels of educations for one of the aspects
- Acombination of the above

An example of how the pilot projects may be distributed is given in the table below.

Level TM aspect	Primary	Secondary	Higher	Informal
Encyclopaedic	Pilot 1	Pilot 3	Pilot 4	
Engagement				
Literacy	Pilot 2			

Organisationally, we will use the infrastructure around the existing Local Time Machines to develop the pilots. The local Time Machines can provide the data and the services for very targeted projects focused on local history (or on the local links to broader historical developments) that can be tested in practice with local partners with access to educational institutions.

The pilots will yield best practices and training and dissemination material which, with the help of the dissemination activities in pillar 4, will be promoted among the stakeholder groups in the educational field.

Since the pilots depend on the availability of sufficient Time Machine Data and Infrastructure, they will start somewhat later in the project. The pilots are preceded by a preparation phase

which focuses on stakeholder organisation and collection of requirements for the necessary infrastructure via interviews, expert meetings and focus groups.

The application of this pilot project can result in the following utilities:

1. Introducing students to Time Machine as a platform of Big Data of the Past and cultural information
2. Introducing students and educators to Time Machine as an open-source platform for Big Data of the Past and cultural information
3. Introducing students and educators to Time Machine as a third-party source of historical and cultural information for reference tools and features
4. Introducing students and educators to Time Machine as localised, personalised history (or, a bridge between the exotic and the local)
5. Using Time Machine as a source of new engagements with historical information

Milestones

Milestone	Year	Description
M1	2020-2022	Preparation phase: stakeholder organisation, requirement analysis
M2	2022	Selection of pilot projects round 1
M3	2023-2025	Pilot project 1 & 2 executed and monitored
M4	2025	First round pilot projects evaluated
M5	2025	Development of best practices and training and dissemination material
M6	2025	Selection of pilot projects round 2
M7	2026-2028	Pilot projects 3 & 4 executed and monitored
M8	2028	Second round pilot projects evaluated
M9	2028	Refined best practices and training and dissemination material
M10	2030	TM infrastructure for education in place

Key Performance Indicators

- Number of users (both students and general users) using Time Machine web-based tools for reference and developing applications on the basis of big data of the past
- Number of non-partnered high-schools relying on TM-supported applications and tools and teaching TM-provided Big Data of the Past analytics
- Number of national ministries of education to have committed to TM content
- Number of educational publishers to take up TM content

Funding Sources

For the pilot projects, dedicated TM funding from European and national funding schemes is required. Where possible, we can benefit from grants obtained by the Local Time Machines.

At the higher education level, we can link with existing European programmes for knowledge exchange:

- Erasmus program (BA/MA)
- Marie Curie Initial Training Networks (PhD)

There have been fresh proposals to tap onto the Erasmus programme by funding pan-European courses and diplomas. Time Machine would complement these initiatives well, as it is already equipped with educational facilities and staff and comprises leading institutes in areas such as cultural heritage and the digital humanities. It could standardise a set of BA or MA courses in, e.g., digital cultural heritage, AI for the digital humanities and other methods-based courses.

5.4 Other Exploitation Areas and Uses

Galleries, Libraries, Archives, Museums (GLAM)

GLAM institutions are central in the collection, description and making Europe's Cultural Heritage accessible. In the process of the digital transformation, GLAM institutions face multiple complex challenges and operate in completely different contexts concerning societal discourse, economics, and technology. Novel trends in fields like AI, AR, VR, machine learning, automatic or semi-automatic description of (digital) objects, and immersive experiences provide sheer endless possibilities, yet are only slowly being adopted throughout the GLAM sector. In addition, just a small proportion of GLAM collections have been digitised, being partly due to limited financial and technological resources. Also, institutions are confronted with significantly higher and more diverse expectations from the general public and constantly need to reposition themselves.

Improving the efficiency and processes for the handling, storage, description, exploitation, exhibition, discussion, and interconnection of digital objects will play an ever increasingly important role in the digital transformation of GLAM organisations. Time Machine – essentially the Big Data of the Past and all of its surrounding services – will address all kinds of issues GLAMs tackle today. This research and innovation plan will examine possible fields of application for Time Machine in the GLAM sector.

The overarching vision for this research and innovation plan is to further **strengthen the role of GLAM institutions as central Cultural Heritage providers;** as driving forces for cultural experiences; as platforms and rich sources for education, research, entertainment, creativity, and innovation for current and future generations by adopting Time Machine data, knowledge, processes, services, and tools. In more detail, we propose the following objectives:

- **Promote the adoption of processes, services, and platforms by the Time Machine initiative** as the standardised backbone in enabling GLAM institutions to open up, enrich, share, and exploit their (digital) collection(s)
- **Create synergies** between developments and strategies already underway in the GLAM domain and Time Machine initiative, e. g. connect GLAMs – and especially smaller institutions – with Local Time Machine initiatives
- **Develop the frameworks, pipelines, and business models** to enable GLAMs to actively contribute to, process and/or re-use the Big Data of the Past

Targeted Achievements

Time Machine will boost, aid and accelerate many developments that are already underway in GLAM and introduce completely new **transformative effects**. Since **collections constitute a key element of GLAM institutions,** we propose to categorise Time Machine's **transformative effects in four areas dealing with (digital) collection(s).** This concept should help to cluster Time Machines various developments and exploitation possibilities in GLAMs.

Collection Custodianship & Enrichment

Nowadays, **digitisation** of analogue objects is a labour- and cost-intensive process. Large quantities of analogue objects still reside within GLAM institutions that have not been digitised yet. With Time Machine, versatile and affordable digitisation hardware and techniques will be introduced. This includes novel 2D and 3D digitisation techniques as well as innovative approaches both for large- and small-scale digitisation initiatives. Contrary to state-of-the-art digitisation techniques, Time Machine provides a more effective way of scanning the context of analogue artefacts in order to be used for the Big Data of the Past. Also, born-digital material can be stored and linked as well.

In an ideal scenario, Time Machine's novel digitisation techniques will **lead to a larger body of digitised material** that can be made accessible to the general public according to FAIR data standards. Also, since funding for GLAM institutions (and here especially smaller, local organisations) is limited, Time Machine will lead to more affordable and flexible digitisation services.

The main beneficiaries of these developments are mostly GLAM institutions themselves, mostly through knowledge transfer and hardware innovations.

In order for collections to be queryable, they have to be semantically enriched. However, the description of collections is a tedious and cost-intensive task that often is undertaken as an isolated initiative – and: by humans with special expertise. With Time Machine, various new methods of algorithmic enrichment will be introduced to annotate and describe collections. This enables easier findability of data, e. g. single objects and collections.

With a vast amount of newly acquired metadata, the demand for curation will increase. Time Machine provides intelligent tools for helping GLAM professionals to **select and further refine metadata.** Again, this is expected to demand new skills and create **new job profiles** in the GLAM sector, e. g. a "data curator".

Also, Time Machine fosters automated **information extraction, machine learning, and AI** as the main drivers of innovation in GLAM institutions. Document understanding and **automated translations** (including translations from ancient languages to modern languages) will vastly increase the accessibility for all kinds of audiences and is expected to have a transformative effect not only on education and scholarship but also tourism.

Lastly, GLAM institutions will hugely benefit from Time Machine's vast source of sound and **robust training data.**

Collection Access

Today, **access to (digital) collections** is limited due to legal, financial, technological, or strategic reasons. Time Machine provides frameworks for dealing with the above aspects and vastly increases the visibility and accessibility of collections by helping institutions to streamline the process of opening (digital) collections. In an ideal world scenario, collections can be accessed with the least number of barriers possible.

Also, information will be easier to find and retrieve by novel **query mechanisms**. Time Machine will completely rethink and innovate current methods to query both digitised and born-digital content. This includes data types and objects that cannot be even digitised at the current moment. The main beneficiaries – from a stakeholder perspective – will be the general public – daily users of GLAM institutions – and, further, especially researchers.

Collection Curation, Engagement & Experience

Today, immersive experiences in GLAM institutions are separate initiatives. Their realisation is cost- and labour-intensive and requires an interdisciplinary team of curators, branding experts, storytellers, digital strategists, programmers, technicians, among others.

Time Machine will provide frameworks for enabling institutions to provide **richer and more diverse experiences** for their users, both in a physical, augmented, and virtual setting. This is not limited to collections, it can also tackle talks, performances, or other events. In the physical realm, this is also not limited to GLAM institutions, it can even extend to urban spaces.

Time Machine will introduce **ground-breaking multisensory experiences,** that are elegant, authentic, nuanced, unobtrusive, and customisable according to the user's needs – a truly positive experience by explicitly adopting multimodal interfaces and feedback mechanisms.

Time Machine services let users experience collections by providing the ability to dive deeper, augment or generalise when needed, the ability to set in a context and provide room for imagination.

Further refinement of these proposed avenues for exploitation will be fostered in expert interviews with stakeholders. Examples for sparking discussions and inspiration in expert interviews could be of the following:

- **Mixing of physical and virtual spaces, GLAMs as smart spaces[6]:** Imagine to let users experience different versions of the same exhibition. This could be achieved by letting users choose context: their time resources ("I have limited time"), knowledge about a topic ("I am familiar with the basics"), level of detail ("I just need an overview"), their mood, emotional state, or even different versions programmed by curators. These user preferences could completely change the virtual and augmented realm of a given space. By exploiting the Big Data of the Past, Time Machine could foster unique experiences
- **Feedback of user experiences into Time Machine:** to share knowledge among institutions by linking institutions via standardised Time Machine services and tools
- **Shift between various modalities of objects:** Since Time Machine provides a huge amount of metadata, users could choose their preferred way of perceiving an object. This could lead to a vast reduction of barriers in GLAMs
- **Digital twins**[7] at GLAMs, noting that digital twins can be regarded as a smaller version of the "mirror-world", referred to in section 2.1.

Collection Linking, Reuse & Remix

Time Machine will provide services and tools to make objects and collections travel beyond GLAM institutions. Through the adoption of **automated data linkage** based on customisable parameters, disparate data storages will be able to "communicate" and create **new bodies of knowledge.** These bodies of knowledge will be queryable by the general public and institutions alike.

Time Machine provides state-of-the-art and ready to be customised frameworks to **reuse and remix data** in intuitive ways that foster exploration, e. g. by humans on crowdsourcing

6 https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/

7 https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/

platforms, GLAM labs, and raw data APIs; or semi- and fully automated through the use of machine learning. These initiatives will further feed data and new knowledge back into Time Machine's databases.

Monetisation and distribution of single objects and entire collections are further main focuses that will be addressed by Time Machine, including a discussion of "levels of openness". With innovative business models, GLAM institutions will be provided new sources of income. However, when discussing business models, the Open Data Movement and FAIR[8] data principles have to be taken into account.

Beneficiaries of these innovative remix methodologies can be found in education, research, and creative industries.

8 https://www.force11.org/group/fairgroup/fairprinciples

Implementation Methodology

Despite sharing common characteristics, we propose to not treat galleries, libraries, archives, and museums as a single, unified entity when discussing potential exploitation. With GLAM institutions acting as one of the main contributors, but also as processors and users of Time Machine data, the interplay between the Time Machine initiative and GLAM institutions (as well as neighboring exploitation areas like creative industries and smart tourism) is a complex and multifaceted relation. In general, GLAM institutions interact with Time Machine in three ways, or "roles": data contribution, data processing, and data use.

We recommend pooling the following tentative roadmap in thematic clusters containing various activities for the coming years. Stakeholder consultations are expected to shape and refine these tentative roadmap topics. Some aspects of this tentative roadmap overlap and require coordination. Proposed roadmap activities also run in parallel and are not necessarily dependent on each other.

Cluster 1: Vision and Strategy

The goal is to place Time Machine as an essential initiative in the digital transformation of GLAMs. This first cluster not only deals with disseminating Time Machine concepts and developments to relevant stakeholders in GLAM institutions but also governs collaboration between TMO and participating GLAMs. Ideally, this leads to further exploitation opportunities and an overarching, refined understanding of Time Machine's vision, mission, and values.

Tentative activities include:

- **Strategic task force:** Establishing a dedicated think-tank on a strategic level is expected to lead to a jointly shaped vision and strategy for exploiting the Big Data of the Past. We propose that this think tank consists of GLAM institution's key stakeholders and representatives of the TMO and is linked to Local Time Machines where applicable.
- **User stories and use cases:** Formulating generic user stories and use cases within the GLAM domain facilitates the identification and clarification of system requirements on an operational level. At this point in the CSA-phase, we suggest to link user stories and use cases to the proposed GLAM roles (data contribution, data processing, data use) or to the expected areas of impact for GLAM (digitisation, indexing, connection, experience). We plan to establish the definition of user stories and use cases, as formal outcomes of ideation by the strategic task force, as a recurring, iterative process.
- **Dissemination for GLAMs:** Emphasising and tailoring Time Machine's key benefits and current developments for GLAMs especially is crucial when trying to foster the uptake of the initiative. We propose to base dissemination activities around the main areas of expected impact for GLAM: digitisation, indexing, connection, and experience.

Cluster 2: Experimentation and Exploration

A prototypical framework will allow GLAM institutions to test technological advancements of pillar 1 and operational breakthroughs of pillar 2. Activities grouped in this cluster are intended to foster agile experimentation. To minimise the organisational overhead required to enable this cluster, we propose to focus on a few select GLAM institutions.

Tentative activities include:

- **GLAM selection:** Defining a set of criteria to select GLAM institutions for pilot actions paves the way for fruitful results of exploitation. We propose a few preliminary topics influencing selection criteria to be further developed with stakeholders during the CSA-phase and later by the strategic task force in cluster 1. Areas could include the institu-

tion's profile (established vs. new player), framework conditions (varying depending on location), and alignment with Time Machine's vision, mission, and values (it could prove beneficial to reach out to GLAMs that are not familiar with Time Machine, e. g. to identify and address pain points that hinder adoption of Time Machine services).

- **Pilot action building-blocks:** Defining distinct scales for pilot actions within the TMO is crucial in serving a range of potential partners in the GLAM sector. For pilot actions, it is essential to reduce entry barriers to allow both established institutions and smaller organisations to participate. For instance, to strengthen GLAMs' role of data contributors to Time Machine, pilot actions could include the application of both top-down and bottom-up digitisation pipelines (as proposed in WP3).
- **Execution of pilot action / Test scenarios:** This activity deals with executing and handling the pilot action that consists of concrete, tailor-made test scenarios (e. g.: for novel business models and Time Machine services), as well as its documentation for other clusters. These test scenarios are based both on formulated top-level user stories and on the specific parameters of selected GLAM institutions for pilot actions.
- **Local Time Machines and co-creation platforms:** We propose to design and setup co-creation platforms and interlock them with Local Time Machines.

Cluster 3: Generalisation and Sustainability

The purpose of this cluster is to increase the likelihood of adoption of the Time Machine initiative in the medium- and long-term by creating generalised models deduced from learnings collected (in other clusters).

Tentative activities include:

- **Impact assessment:** We propose to establish assessment measures of pilot actions and tests based on the concept of impact, focusing on digitisation, indexing, connection, and experience. This impact assessment will mainly focus on activities and end users in the GLAM sector. However, also the Time Machine initiative as a whole is expected to be impacted by exploitation avenues.
- **Reporting and documentation:** Insights of pilot actions and smaller experiments should be continuously documented in order to be reusable.
- **Model and generalisation:** Creating generalised models based on outcomes of impact assessments and pilot actions will enable a large-scale roll-out of the Big Data of the Past.
- **Large-scale roll out:** Based on models and generalised principles, eventually, this activity will allow for large-scale EU-wide exploitation of Time Machine in a GLAM domain and will provide data, services, and tools for the greater good of European society. In conclusion, with previous findings and learnings, this activity's objective is to ensure Time Machine's long-term sustainability well beyond 2030.

Cluster 4: Collaboration and Outreach

In collaboration with neighboring exploitation avenues, such as scholarship, education, creative industries, and smart tourism, lies tremendous potential both for GLAMs as well as Time Machine – both in developing business models, as well as sharing resources and synergies. Beyond connecting to other domains, this cluster deals with the engagement of the general public.

Tentative activities include:

- **Creation of a smart cluster:** Based on the S3-framework[9] (Smart Specialisation Strategy), smart clusters will be formed to foster exploitation beyond GLAMs (see "smart tourism" for more details). This activity is intended to involve regional political stakeholders and is planned to be executed together with neighboring exploitation avenues creative industries and smart tourism.
- **Connection to Local Time Machines:** To reinforce Time Machine's sustainability across regions, GLAMs will be profoundly involved in Local Time Machines. We propose to align these collaborations with novel Time Machine business models.
- **Workflow for idea exchange:** Designing and implementing an agile workflow on an operational level within the smart cluster allows sharing ideas, developments, and corrective measures when needed.
- **Engaging the general public:** GLAMs could increase their potential in engaging the general public in a meaningful, truthful, and gratifying way (as already partially achieved with e. g. crowdsourcing platforms or labs). Also, this activity could include the transfer of Time Machine skills as proposed by creative industries.
- **Collaboration with scholarship and education:** The exploitation of Time Machine can also be fostered by linking scholarship, education, and GLAMs, e. g., by sharing intellectual property that can only be opened up through Time Machine.

Milestones

Clusters	Milestone	Year
1. Vision and Strategy	Strategic task force established	2020
	User stories and use cases formulated	2021
2. Experimentation and Exploration	Selection criteria defined	2021
	Selection process finished	2022
	Concrete test scenarios refined	2022
	First pilot action(s) launched	2023
	Impact analysis of pilot action concluded	2025
3. Generalisation and Sustainability	First model derived from experimentation and pilot actions	2025
	Large-scale roll-out started	2025
	Sustainability plan	2028
4. Collaboration and Outreach	First smart cluster created	2021

9 http://s3platform.jrc.ec.europa.eu

Key Performance Indicators

The Big Data of the Past will strengthen (and reposition) GLAMs as innovative players that create a huge impact on society and economics in Europe and beyond. We generally propose a few KPIs to measure this impact:

- **Heritage Accessibility:** Ratio of accessible material vs. inaccessible material (according to FAIR data principles)
- **Cultural and Economic impact:** Number of jobs in the sector created that are attributable to the Time Machine Initiative
- **Heritage Digitisation:** Ratio of digitised vs. non-digitised material; Speed of digitisation
- Levels of adoption of **Linked Open Data**
- **Impact on Creative Industries:** Number of collaborations with (local) creative companies contributing to providing state-of-the-art experiences in GLAM institutions
- Number of adopted **Time Machine services** and tools in GLAM institutions
- Number of **realised Local Time Machines** with GLAM institutions being a leading or contributing factor

Funding Sources

National Sources

Most GLAMS in Europe are funded through national (or regional) budgets. Most of the national budgets are earmarked within the organisation (e.g. staff, collection development, digitisation, marketing, etc.). Most organisations rely on projects funding for their innovation needs. These innovation needs can be funded through national funds that are often developed in project calls, which goes for digitisation as well. The TM platform could work as a catalyst for these types of funding, linking national funding to TM funding.

Other GLAMs function with governmental support as independent organisations, e.g. from endowments or entry fees. These cover mainly the basic needs of the organisation.

European Union Funding

The EU funds many research and innovation actions. This is often done in collaboration between institutions in different member states, linking research institutes, universities with the organisations in the field. There is a huge competition in getting this funding, but it's an important source of income to the GLAM-sector (especially research libraries, documentation centres, and large museums).

Funding by Non-Governmental Institutions/Individuals

Several non-governmental institutions function that support GLAMs. Often, they are of a philanthropic nature or public supported such as (national) lotteries or similar organisation. They tend to fund GLAMs for a longer period of time and provide basic income for the organisations. Or they fund specific types of work in the GLAMs (e.g. building extension, website development, marketing, outreach).

Creative Industries

Creative, media and entertainment industries are an integral part of a strong European economy and an engaged society. TM data and services will introduce transformative effects which will offer completely new avenues and innovation prospects for these industries. The

Big Data of the Past has the potential to become a rich resource for inspiration and creativity and will be exploited to create new works, experiences and products. LTMs will open opportunities to experiment with this novel data and technologies, create opportunities for cross-sectoral collaborations and foster frameworks that support the remuneration of creative outputs. This will have a transformative impact on the creative value chain across the creative industries and beyond, enabling organisations and individuals to take part in a competitive market and deliver high-quality creative products for commercial exploitation.

The roadmap for exploitation creates pathways for stakeholders in the creative, media and entertainment industries to successfully take up these and experiment with these innovations and materialise the envisioned social and economic benefits of the Time Machine project. The objectives of this research and innovation plan are to:

- Create pathways for creative industries to exploit Time Machine data and services via Local Time Machines
- Identify and address framework conditions that will enable and accelerate experimentation and exploitation
- Establish connections with stakeholders in other sectors for cross-sectoral collaboration and scaling

Targeted Achievements

1. Creation: Elaboration of ideas, contents and products

Large quantities of multimodal data made available through the use of advanced computing technologies and data visualisation techniques will support the exploration and retrieval of yet undiscovered patterns, connections and observations which will serve as an inspiration for the development of new creative ideas. The open and interoperable infrastructures for data exploration which will enable creative freedom and diversity. Artificial intelligence will also support new forms of creativity, including computational creativity.

2. Production/Publishing: The making of original, non-reproducible or reproducible work

Production processes will be supported by easily finable, high-quality resources. Rich cultural data will be available as assets for creative reuse according to the FAIR data principles (findable, accessible, interoperable and reusable), providing sufficient context and level of granularity. Smart metadata models will support the ability to combine and seamlessly integrate digital objects in different variations to tell different stories[10].

Storytelling will be enhanced using ground-breaking simulations and visualisations. Possibilities to query granular properties of digital objects (including spatial, temporal, tactile, visual and aural qualities) will support the emergence of new kinds of storytelling techniques that appeal to different senses. The increased computational processing capabilities for big data processing will also reduce the complexity of production processes (e.g. use 3D visualisations and modelling in design).

In addition, Time Machine infrastructures will offer an alternative to the current gatekeepers and intermediaries in the market who set high barriers around access to high-quality content; this will significantly improve opportunities for SMEs and individuals in the creative industries. Reuse of data will be supported by clear copyright acquisition and licensing mechanisms. Newly developed business models will ensure that both data providers and creators can benefit from these transactions.

10 For example, BBC developed toolkit for production teams to create personalised object-based experiences and narratives. See https://www.bbc.co.uk/rd/projects/object-based-media-toolkit

3. Dissemination/Trade: Dissemination of cultural products to make them available to consumers and distributors

Supported by new business models, licensing frameworks and high-quality resources, creative industry players will have more bargaining power to enter the market and promote and disseminate their creative outputs. Smaller and much more diverse players are likely to emerge, further fostering creative circulation in the digital single market.

Machine learning and natural language processing technologies will support the delivery of high-resolution experiences at a massive scale for broad audiences and over various platforms. Other sectors, including the tourism industry, GLAMs and education, will benefit from novel services and experiences designed for their end-users. With more creative products to offer, the role of European online platforms in the digital market will gain a prominent role and attract much more traffic and investment. Sectors

4. Transmission/exhibition/reception: Provisioning access to creative products for consumption

Metadata about the Intellectual Property of new works will be managed in a machine-readable way to track copyrighted content on a granular level (tracking of individual elements or excerpts) and support remuneration, rescue and reuse. Collective licensing frameworks and other security mechanisms (e.g. smart contracts) will support smaller actors in the sector and provide sustainable revenue streams.

Time Machine will also develop models that will help to incorporate user-generated content, in this way increasing cultural participation and raising awareness about the potential of cultural heritage.

Implementation Methodology

Given the broad scope of the creative, media and entertainment industries, each domain might be dealing with different framework conditions and existing infrastructures that would influence their capacity for exploiting the Time Machine data and services. Their level of readiness to enter into the Time Machine ecosystem might be very different and cannot be generalised. They need a gateway which would initiate and accelerate exploitation.

The Local Time Machines would act as this gateway, a launchpad for bringing stakeholders from the creative industries to the Time Machine ecosystem, enabling them to exploit the Big Data of the Past and benefit from the tools and infrastructures that it provides. Each Local Time Machine would act a smart (thematic) cluster that would invite stakeholders from the creative industries to develop products and services around it. Based on their individual strengths, technological developments and regional support, Local Time Machines would build and grow communities of stakeholders from across the creative industries, as well as stakeholders from other exploitation areas, and create the conditions for them to exploit the Time Machine resources. Starting on a local/regional level with local stakeholders, the network of Local Time Machines would expand and support cross-sectorial and cross-regional collaborations and stakeholders to benefit from the pan-European Time Machine infrastructure and resources.

To bridge the gap between the state of the art and the desired targeted achievements in the creative value chain, the role of the Local Time Machines is to act as incubators or living labs for the creative industries. This will be achieved in two phases:

1. First, the proposed exploitation scenarios and frameworks will be **validated through Proof of Concepts and collaboration and outreach activities** will be conducted to engage relevant stakeholders
2. Building on the outcomes of this, Local Time Machine will **establish incubation hubs that enable creative industries** to exploit Time Machine data and services

Initiation Phase

Cluster 1: Collaboration and Outreach

Local Time Machines will position the creative industries as an integral part of the Time Machine ecosystem that can provide new products and services for GLAMs, tourism industry and education as well as other sectors, in this way increasing and opening exploitation opportunities and supporting the Digital Single Market. Local Time Machine will play a key role in facilitating these cross-sectoral connections to embed creative products in other industries as well as reaching new players in the creative industries who could benefit from Time Machine. The following actions for facilitating collaboration and outreach are proposed:

- **Time Machine Ambassadors.** Identify representatives from the different domains in the creative industries who could offer their expertise to the Time Machine consortium, act as mediators between Time Machine and the industry, help to mobilise new stakeholders and participate in the ongoing ideation for the roadmap
- **Creative residencies** to connect individual creatives and SMEs to other sectors - in particular, GLAMs, tourism industry and education
- Training/mentorship/peer-learning programmes that **target creative individuals and SMEs to support collaborative creation and capacity building.** Stimulate cooperation models that help smaller actors join and compete with larger players

Cluster 2: Proof of Concept

As a starting point, it is essential to validate the proposed roadmap actions that support the creative value chain as well as strategically position and demonstrate the exploitation impact in the creative industries and define priorities that need support from the decision-making bodies. To galvanise this process, Time Machine proposes to run **Proof of Concepts with one domain in the creative industries** and through this process, refine the roadmap and the proposed exploitation strategies that will pave the way for exploitation in the creative industries at large.

In order to achieve the greatest impact in a short period of time, we propose to invite stakeholders from the game industry to take part in the Proof of Concept stage. As an industry that holds a prominent position in the European market and already has connections to the cultural heritage sector, it is strategically positioned to efficiently embed the innovations introduced by TM into its exploitation mechanisms and business models. The proven success of video games that reuse cultural heritage resources developed by companies such as Ubisoft, Semantika and DROPSTUFF.nl, point to exploitation potential that other domains in the creative industries could tap into. Given the large number of independent and amateur game developers, it would also serve as a testing ground to see how could smaller and individual players find their place in the market with the help of TM. In addition, using cutting-edge technologies, game developers can take advantage of the multimodal cultural heritage resources to create immersive experiences and build rich historical narratives with a great level of detail.

A number of well-established as well as smaller players in the game industry would be invited to experiment with a number of scenarios that correspond to activities in the creative value chain[11]. Based on the realisation of exploitation scenarios in the game industry, business case studies and user stories will be developed to serve as a source for inspiration for other stakeholders in the sector, demonstrate the gains and impact as well as challenges related to exploitation that need to be addressed.

The results of these initial experiments **will inform the activities** in (1) "pillar 1: Science and technology for the Big Data of the Past and (2) pillar 2: Time Machine Operation". Furthermore, the Proof of Concept phase will inform the realisation and strategy for setting up Local Time Machines as innovation hubs for the creative industries to support stakeholder engagement and exploitation models.

Execution Phase

During the execution phase, Local Time Machine will (1) **facilitate incubation** of new ideas and tools for the creative industries, (2) establish **infrastructures that support reuse and exploitation** of the Time Machine data and services and (3) **ensure sustainability** of these activities on a pan-European level. Collaboration and outreach activities (see above) within the creative sector as well as with other industries in the European market will continue.

Cluster 3: Incubation

The central role of the Local Time Machine in this roadmap is to act as incubation hubs for the creative industries and provide knowledge, resources and networks of partners and audiences needed to test innovative ideas and exploit the Time Machine data and services. Their role is to:

- Support creative entrepreneurship
- Foster experimentation with new data, technologies and business models
- Provide an environment for testing new ideas (ideation bootcamps) and scaling up
- Facilitate the exchange between the industry, researchers, decision-making bodies and other related sectors
- Introduce the potential of the creative industries to other sectors and potential investors, and help to identify new opportunities

In setting up the incubation hubs, TM will build upon insights from tested collaboration models and methodologies. For instance, the Sandbox hub initiated by public broadcasters across the EU has developed a model to validate new technologies[12]. Also, collaborating with EBN, the network of over 140 business and innovation centres, and ImpactHub with more than 16,000 members will help to maximise the impact of Time Machine[13]. With respect to design methodology used, insights from frameworks such as 'Design Sprint' by GV and 'Future Visioning' developed by Business Models Inc. will be used[14].

Cluster 4: Support Infrastructure

To support the incubation activities, Local Time Machines will negotiate and establish and infrastructure that foster innovation and experimentation, and support the remuneration of creative outputs. These support mechanisms will also ensure that the exploitation activities carried about by the Time Machine consortium and their stakeholders are reaching their objectives. While Local Time Machine will provide targeted support for their stakeholder groups, it is essential that these support infrastructures are coordinated on a pan-European level.

11 These will be defined in coordination with other pillars and depend on the resources and infrastructure available.

12 https://www.mediaroad.eu/about-sandbox-hub

13 http://www.ebn.eu, https://impacthub.net

14 https://www.gv.com/sprint, https://www.businessmodelsinc.com/strategy-design/future-visioning

The following infrastructures are proposed:

- Clinics that help stakeholders in the creative industries to develop **"Time Machine skills"** (e.g. digital skills necessary to work with the Time Machine data and services). Local Time Machines could offer a certification programme to encourage the development of these skills
- An observatory that **monitors trends and measures the impact of exploitation.** To ensure that Local Time Machines provide the necessary support for the creative industries, it is essential to continuously monitor technological innovation and trends in the sector. This observatory would provide recommendations that would enable decision-makers and the Time Machine consortium members to respond in time to the changing conditions in the market. The monitoring should be done on a domain-level to ensure that the framework conditions of each domain in the creative industries are addressed
- Pan-European licensing hubs[15] **that oversee fair licensing regulations,** ensure remuneration for creative products and provide support for individuals and organisations in the industry

Cluster 5. Sustainability

This activity identifies ways of enabling take-up of project results in order to achieve the expected outcomes in a sustainable way and at scale. Models for running incubation hub will be developed so that exploitation activities could be supported at a large scale across Europe and efficiently adopted by in new Local Time Machines.

15 For example, licensing hubs have been successfully established in the music industry: https://www.bmat.com and https://www.armoniaonline.com

Milestones

Milestones are provided in the table below.

Clusters	Milestone	Means of verification	Year
1. Collaboration and outreach	Hub infrastructure established	Partners in the Game Industries approached and strategic collaborations with relevant ancillary networks established.	1
	Outreach Strategy in place	The Outreach Calendar published and the marketing strategy is in place. The marketing strategy includes: (1) appointing Time Machine Ambassadors, (2) hosting of creative residencies, (3) training, mentorship and peer-learning programmes	1
2. Proof of Concepts	Selection criteria User stories refined	Successful execution of the Proof of Concept in the Gaming Industry.	2
	Concertation efforts across the TMO pillars	Outcomes from the Proof of Concept are discussed with pillar 1: Science and technology for the Big Data of the Past and pillar 2: Time Machine Operation. This results in updated activity plans across the pillars	2
3. Incubation	Launch of the incubation activities	Methodology established and First incubation activities launched.	2
4. Support infrastructure	Launching and operating the monitoring observatory	The observatory monitors trends and measures the impact of exploitation.	3
	Licensing hubs launched	The hubs oversee licensing regulations, remuneration and provide support to its users.	5
	First clinics launched	Content of the clinics co-designed with end-users.	2
5. Sustainability	Large-scale roll-out of incubation activities	Start-ups and scale ups identified, support scheme in place.	6

Key Performance Indicators

The exploitation of the Time Machine data and services will have a significant impact on the European economy and society. Time Machine will give a strong boost to the creative, media and entertainment sector itself but more importantly, it will produce much broader spill over effect. Notably, it will generate new ideas, knowledge and products that will benefit other industries, the public sector and European society at large. The KPIs indicate the impact within the sector as well as demonstrate these much wider effects of exploitation to ensure continuous engagement from the industry and support from the decision-making bodies. The list below is a non-exhaustive list of the most KPI's.

Collaboration and Outreach

- Number of cross-sectoral collaborations
- Number of stakeholders from the creative industries joining the Time Machine Organisation
- Growth in cultural participation and growth in social inclusions (e.g. number of products, experiences and services tailored for the disabled)

Proof of Concepts

- Number of scenarios tested
- Number of stakeholders involved in Proof of Concepts

Incubation

- Representation of all creative industries' domains in the Local Time Machines
- Number of stakeholders connected to the Local Time Machines
- Number of entrepreneurial start-up and scale-up-stage businesses initiated as a result of the incubation efforts
- Number of strategic partnerships with relevant ancillary networks

Support Infrastructures

- Creative sector contribution to GDP
- Number of items available for reuse (high quality, using correct rights labels)
- Number of self-employed individuals and SMEs involved
- Employment in the creative industries

Sustainability

- Number of products and services developed for other sectors
- Income from licensing and use of TMO data and services
- Number of Time Machines with long-term financial stability

Funding Sources

The creative, media and entertainment industries can benefit greatly from the already existing European and national funding programmes. However, these programmes often concentrate on creative content creation but do not support other activities in the creative value chain, namely distribution/trade and exhibition/reception/transmission, that are key to the exploitation envisioned by Time Machine. What is more, this funding is often inaccessible to smaller players and self-employed individuals.

The Time Machine consortium should advocate for funding that provides support for the following:

- SMEs and individual players in the sector for whom barriers to apply for European funding schemes are often too high
- Synergies between actors in the creative sector and cross-sectoral collaboration to stimulate capacity building
- Activities throughout the creative value chain, in particular, distribution/trade and exhibition/reception/transmission

We foresee four main sources of funding that could support these research and innovation needs in the creative, media and entertainment industries:

- National Funding Sources: Funding sources vary per EU country. For example, in the Netherlands, the following funding streams are relevant: Creative Industries Fund NL, NWO, Mondriaan Fund, Fonds21
- European Funding Programmes: Creative Europe, Horizon Europe, Digital Europe, ERASMUS+, COSME, SME instrument, Structural Funds
- Private sector investments: Public private partnerships, equity investment, artist in residence
- Crowdfunding: Reward-based crowdfunding, tax shelters, match-funding schemes

Smart Tourism

The smart tourism exploitation platform has as objective to:

- Reach out to creative industries (core re-users) defining specific needs for the tourism industry to create technology-driven CH tools and services
- Use innovative 3S clusters to develop TM tools and services for smart tourism according to local priorities
- Identify a synergy model for core re-users, enablers and infomediaries to propose TM technology-driven CH products and services to end-users interested in touristic destinations, thus re-shaping their approach to sustainable/responsible touris

Targeted Achievements

- Raising awareness and respect toward CH destinations through TM narratives
- Innovative clusters working with local TMs to create a permanent ecosystem of smart tourism
- Economic sustainability of CH destinations, locations and institutions (GLAM) through TM smart tourism model
- Smart tourism through the TM products and services contributes to smart cities
- Enhance life-long learning programs through the TM smart tourism model

Implementation Methodology

Today, awareness to the overwhelming growth in tourism, its economic potential in the context of globalisation (the fourth industrial revolution) and its impact on territorial, urban and social transformations, coupled with the conviction that cultural tourism is tightly linked to education for diversity, to intensification of the European identity and to respect of CH artefacts and sites seems a fertile ground to revolutionize the whole sector by creating through the TM exploitation model a smart CH ecosystem which takes into consideration the whole pipeline: the decision makers creating needed legal framework and defining priorities, the creative sector (core reusers) with its technology-driven products, the TM platforms enabling core re users and end-users to enjoy the Big data of the past, the tourist industry's stakeholders who define their type of business model or cultural open data and the Web's infomediaries who reach out to end-users.

TM is based on its local Time Machines which create through technological innovations in AI and machine learning the Big data of the past specifically set to tell the history of specific area (city, site, province, region). Following the 3S (Smart Specialisation Strategies), TM proposes to identify the components of territorial clusters which can be interested in developing specific technological innovations and tools for local TM cultural-heritage experience platforms (SMEs, universities, start-ups, regional administrators) and create the conditions for smart tourism to be considered a local/regional priority. The creation of a "smart cluster" following the 3S framework envisages the participation of regional/municipal political stakeholders which set up the priorities and create the optimal conditions for the formation of the cluster (which is considered a network of start-ups, SMEs, research institutions, cultural institutions, regional/municipal administrators that share common goals and standards and create, on the basis of agreed priorities, tools and services). This "smart cluster" should by no means be limited to cultural smart tourism, as it includes cultural institutions and GLAM and above all, local creative industries that help shape together with GLAM and the local cultural smart tourism policy the output of local TMs. It is highly important to stress that local TMs are the backbone of this local "smart cluster" and that their relationship with TMO is twofold: they are given a TM franchise from the TMO, but also share through it standards, tools and services common to other TMs. The local "smart cluster" is also an enabler as it is responsible for the creation of a local "smart tourism" platform that unites all tools, services and products to be proposed to core-users, end-users and infomediaries.

The example of the Welcome City Lab, a French platform which aggregates various startups whose products are designed for smart tourism[16], is a case in point. Created in 2013, its founding members are the City of Paris, the BpiFrance, the Tourist Office, the DGE (Direction Générale des Entreprises), the Paris Airport, Air France, a Caisse des Dépots, Galeries Lafayettes, Compagnie des Alpes, Paris Inn Group, RATP, Skyboard, Sodexo and Viparis. It is an incubator offering an innovating platform to experiment together ideas and technology regarding smart tourism. Today it has 100 start-up companies, 600 jobs created, and 140.000.000 euro raised. The model has been copied in other cities: Deauville, Angers, Nimes, Aix-Marseille and Troyes and is expanding.

Milestones

After 1 year: Local TM with local 3S cluster define priorities regarding targeted tourist profiles, CH prioritized narratives, CH local destinations

After 3 years: First set of technology-driven CH narratives to be tested on targeted tourist profiles

After 4 years: Infomediaries to be approached with customer satisfaction test results

After 5 years: Measuring social, cultural and economic impact according to ETIS

Key Performance Indicators

Europe has launched in March 2016 ETIS: The European Tourism Indicator System ETIS toolkit for sustainable destination management[17]. The ETIS is a management, information and monitoring tool specifically intended for tourism destinations. It is designed as a locally owned and led process for collecting and analysing data with the overall objective to assess the impact of tourism on a destination.

The ETIS is based on 27 core indicators and 40 optional indicators, subdivided into four categories:

1. Destination management
2. Social and cultural impact
3. Economic value
4. Environmental impact

The KPIs are designed for any destination wishing to measure the sustainability of the tourist industry:

1. Raise awareness: Emphasising the importance of obtaining relevant local political support for implementation
2. Create a destination profile
3. Form a Stakeholder Working Group: There is no one set formula that works for every destination. It is important to be flexible and take an approach that best suits the destination and the group of people involved

16 https://welcomecitylab.parisandco.paris

17 https://ec.europa.eu/growth/sectors/tourism/offer/sustainable/indicators_en

4. Establish roles and responsibilities: It is the role of the local destination coordinator to steer stakeholders towards an agreement on setting targets, taking action and planning how to achieve these aims
5. Collect and record data: Data collection should simply be a process of bringing the various data sources together in one place to build a detailed picture of the destination's tourism industry
6. Analyse results and take action on the basis of priorities
7. Enable ongoing development and continuous improvement: The data collected should help tell a story about the destination that can be integrated into marketing and communication plans, as well as informing long-term strategy and policy

Funding Sources

A reasonable balance between public money, especially during the launching phase and further private investment based on the creation of 3S clusters built around local TMs which generate OD is the key to success.

The chain value to be adopted in the case of open data reuse lies in a resource (one or many datasets), released according to OD paradigm (without technical, legal and price barriers) which, if elaborated, becomes the enterprise-specific asset and, integrated into the enterprise's value proposition to the market, is "packaged" and embedded in a bundle of products and services.

The potential business models of each of the actors on the value chain are[18]:

- Core re-users (those facing directly the consumer):
    1. Premium product/service: Offering the end-user (high-end market) a product or a service characterised by high intrinsic value in two modes: (a) à la carte – pay-per-use; (b) recurring fee – all inclusive
    2. Freemium product/service: One of the offerings is free-of-charge and entails only classic features, while customers (low-end market) willing to take advantage of refined features or add-ons are charged
    3. Open source like: Costs incurred for free offering of unpackaged open-format data are covered by revenues stemming from supplementary business lines open-data-based
- Enablers (those operating behind the front lines):
    1. Infrastructural razors & blades: The value proposition hinges on an attractive, inexpensive or free initial offer that encourages continuing future purchases of consumable follow-up items or services, characterised by inelastic demand curve and high margins: datasets stored in cloud accessible via APIs and re-users charged only for computing power they employ on-demand
    2. Demand-oriented platform: Platforms capable to convert datasets in data streams by using metadata, harmonised formats exposed through standardised APIs. The earned revenue is in exchange for advanced services and refined datasets or data flows
    3. Supply-oriented platform: Open data holders are charged in lieu of developers. Pen data holders become platform owners making advantage of handy features like cloud-storage, rapid upload of brand new datasets, format standardisation, tagging with metadata and automated exposure of data via APIs and GUI (graphical interface)

18 Yannis Charalabidis et al., The World of Open Data. Concepts, Methods, Tools and Experiences, Cham (CH), Springer, 2018, pp. 115-156.

- Infomediaries (organisations positioning themselves between open data producers and users):
    1. Single-purpose apps
    2. Interactive apps
    3. Information aggregators
    4. Comparison models
    5. Open data repositories
    6. Service platforms

Smart Cities, Urban Planning, Land Use & Territorial Policies

The main objectives are listed below.

Integrated, inter-connected information systems for cities and lands, across time, space and scales, across administrations, across authorities and citizens, that supports not only browsing (in a way Wikipedia does) but also queries. We target intensified and more relevant (smart) information exchange in smart cities with new data sources, including exchange with other cities with comparable infrastructures, and with more focus on historical depth (longitudinal perspective provided by the Big Data of the Past).

Multi-scale and culture friendly city and land information systems. Information system should be "culture ready" in a sense that they can integrate cultural specificities of different information sources as well as of different contexts of use. They can adapt their interface and functionalities (iterative, co-design approach to development of responsive interfaces). They should support zooming in and out between the perspective of Europe and more local perspectives as well as to embrace focus (fine level of details) and context (lower level of details but wider coverage) in analyses.

"Affordable and sustainable" solutions to build specific cities or lands information systems (Time Machine projects) that integrate into a wider framework, whatever a city or rural territory or country resources (in terms of funds but also expertise and communities), including in emerging countries, and available also for transversal themes (e.g. Glaciers Time Machine, Wetlands Time Machine etc.). It is important that the studied cost comprises ecological footprint as well as how much of private information we are ready to share. In particular, promoting a shared knowledge graph and delivering new archiving principles and strategy.

User-centred retrieval of facts and data in Europe history (other cities, other territories) to favour exchange and mutualisation as a bottom up process to find solutions to sustainable development challenges, that may complete existing a top-down process using the state or using the European Commission. Users also need meaningful documentation of uncertainties and hypotheses.

Recommendations for decision makers to support their planning and design solutions: suggesting connections, presenting situations from the past that are related to the present-day experience of specific localities and phenomena can support and inspire decision makers, citizens, scientists to invent new solutions and approaches, e.g., regarding choices in urban development or land use. TM can also support cities in finding out which other cities are facing similar challenges, e.g., managing tourism, water management, social cohesion, and share data and solutions.

Enhanced scienceS-policy interface as well as scienceS-stakeholders interface either in cities or in land management in general: to connect stakeholders who seek a longitudinal perspective on a present-day problem with the relevant scientific communities to sample history and space and design training data set with regards to a given issue, apply machine learning method, trained on these samples from the past, and using Time Machine Knowledge graph to make recommendations on his specific problem. It is important that stake holders can use

an appropriate language to express their questions and visions, soft concepts and not quantitative thresholds. Different scientific communities need to be involved; from digital humanities (e.g., urban historians, information specialists, archaeologists), social sciences (e.g., urban planners, geographers, statisticians), and artificial intelligence.

Debating platforms related to cities and territories design present historical information and heritage in the contexts that are relevant to the experiences of the different audiences (bringing history and heritage to the people, rather than the other way around). As such, these platforms can be leveraged to connect present-day experiences and problems to different past events that make sense to different citizen groups. These platforms should be 'polyvocal', allow for multiple perspectives on the past, creating room for the often-unrecorded stories of minority groups, including newly arrived citizens who may not share the dominant culture. These platforms will also benefit from the capacity to share and compare hypotheses, thanks to story-telling functionalities.

Inclusive and transparent platforms to write and revise policies related to territories: Supporting interactions for stakeholders with different background and perspectives, considering data available to associate trustable dashboards to the policies.

Targeted Achievements

On **scientific bottlenecks** like "heterogeneities management", "unifying uncertainty framework design", "knowledge graph design", "recommendation models", "including soft values in regulation", we need to engage scientists to work on this. This can be done with task 4.1 but also task 4.2 if we think of setting up European master program to train students capable of undertaking such PhD. Pluri-disciplinary benchmarks to study models supporting the identification of similar cities' states -across space and time-, models to compare strategic measures and their impacts -between different cities and possibly different states.

We want to **set up specific Tim[19] Machine calls** (see milestones year 1 and year 4) to fund the digitalisation and indexation of existing archives based on proposals made by communities who describe what they will do with the data, the same way European Space Agency is doing before launching satellites .

The core infrastructure components needed must be confronted to the existing information systems in place. Very soon **core TM metadata for datasets** must be identified, an important element will be the documentation of provenance information and quality information from TM digital assets. These metadata will support the search for data sets in Diamond.

We will rely on existing networks among data providers to design (identify) a **production process of historical land cover products** out of archives and associated tutorials, at different representative scales, in Europe but also in Africa.

Very soon in the project, we will draft a **collaboration platform** based on words and more natural language that supports dialogues and debates using local concepts in connection to European concepts (words, mail, basecamp, picking one pivot language etc. aren't enough), either on an asynchronous way or associated with translators. Wikipedia could be a good candidate to start from. There exist several thematic wikis that could be interconnected. A key item is to have unambiguous URIs to align objects as well as a model to store links between comparable places.

At the level of Europe, we target at a **new implementation of European culture friendly spatial data infrastructure** articulated with existing data and metadata (INSPIRE, EEA, Europeana, etc.) new sources (remote sensing, collaborative content) as well as with needs from scientists (AI, Humanities) and with EAGLE, including a broker component to be able to

19 http://sci.esa.int/cosmic-vision/60498-call-for-a-fast-f-mission-opportunity-in-esa-s-science-programme

cope with member states heterogeneities and a model to document in a meaningful way the uncertainty of patchwork European data products.

Methodologies to foster innovations are needed and we will rely on our participants experience in the organisation of challenges and hackatons around data:

- A framework to design application oriented scientific challenges: relating to actual, present-day challenges (so that representative users can assess the value of a demonstrator even if it is not exactly their specific problems) and that can be used by scientists in their work (e.g. scientists working on data alignment, scientists working on simulation etc. may not have engineering expertise to prepare the datasets and the infrastructure) and incentive for them to do so. e.g.: In domains such as recommendations, or integration
- A framework to design application-oriented development challenges: relating to a range of actual, "real-world" problems (so that representative users can assess the value of a demonstrator even if it is not exactly their specific problem) and that can be used by developers of mature technologies. e.g.: Collaborative design of historical land cover products out of old maps, integrated information system

And prior to this we target in 2020: 3 days Lab, using teams or a platform supporting collaborative solution design between different stakeholders that will aim at:

- Selecting and prioritising use cases in terms of increasing complexity, and in terms of stakeholders' expectations (stakeholders incl. citizens): This will be done based on highly generic scenarios sent one month before
- Listing relevant technologies for territory description and the associated uncertainty: Paper maps, statistics, classifications, remote sensing, digital vector databases, gazetteers, etc.

We will need communication media to evangelise the SDG (sustainable development goals) because they really concern everyone (eradicate poverty on planet earth). We also need media to explain why they are monitored through land cover and land use data and criteria. An idea could be to organise escape games, YouTube video on these.

Milestones

After 1 year: "TM Land Use Digitalisation Missions" call for proposal template, which will invite proposals to get digitalisation, interconnection of given archives for Land Use and Territorial Policies (selection criteria will include the benefit compared to the cost)

After 3 years: Mock-ups of TM Land use integrated and multi-scales information systems adapted to selected thematic areas and to local stakeholders in similar places.

After 4 years: "TM Land Use Learning From the Past Missions" call for proposal template, which will invite proposals to set up machine learning experience (incl. Identify with correct scientific communities what are the relevant places and data to learn from, curating the data, running deep learning algorithms)

After 5 years: Mock-up of TM Land Use debating platform, presented to EU politicians and voting platforms

Key Performance Indicators

- Need to be identified with stakeholders. (see UN Habitat, DG Grow)
- Need to select among the indicators associated to the UN Agenda for 2030[20]: in goal 2 (zero hunger), goal 6 (clean water), goal 11 (sustainable cities); goal 13 (climate action), goal 15 (life on land)

20 https://www.un.org/sustainabledevelopment/sustainable-development-goals

- X European organisations related to urban planning and land use engaged
- X National government bodies related to urban planning and land use engaged
- X Local government bodies engaged
- KPI concerning linking and harmonisation of land use data
- KPI concerning linking and harmonisation of urban planning data
- X Best practices regarding Big Data of the past for land use
- X Best practices regarding Big Data of the past for urban planning

Funding Sources

- National funding schemes that focus on societal challenges (e.g., the Netherlands National Science Agenda)
- European funding schemes that focus on societal challenges (e.g., relevant calls in the new HE Programme)
- Bank and fund management dedicated to rural development
- Several bodies fund initiatives in the domain above (energy efficiency etc.).
- Crowd funding could also work on this type of solution
- Important programming and funding organisations: EEA
- Insurance companies

6 OUTCOMES AND IMPACT

A Thorough Dialogue With Our Past

The digitisation of European cultural heritage represents a formidable challenge, both in terms of cost and in its implementation. Moreover, only a small number of artefacts in collecting institutions is actually made publicly available, while billions of digital records still need to be indexed, so that they can be searched and analysed.

Time Machine will offer new technologies, methods and protocols to address these challenges through mass digitisation campaigns across the whole of Europe, covering also landscapes, cityscapes and architecture, enabling the EU Member States to reach their ambitious objectives for Cultural Heritage preservation and access.

Due to barriers such as being able to read old writings, to understand other languages, or to visit the relevant archives or libraries in person, citizens are, for the most part, excluded from accessing, navigating, and benefitting from this Big Data of the Past. This data is every citizen's shared heritage; it should not be 'exclusive'.

Time Machine now changes the game radically: for the first time in history, every citizen will be able to access historical data and actively engage with it by contributing their own stories and interpretations. This is also important because everything in our lives today builds upon a narrative that Time Machine will allow to be documented.

A Transformational Impact on SSH

SSH accounts for well above 40% of students in European Higher Education. SSH is also the largest European research community, with more than 30% of EU researchers in Higher Education, corresponding to about 500.000 Full Time Equivalent (FTE) positions. The SSH research spending, however, is substantially lower than 30% of the overall research spending, and is often lower than 20% in many countries.

The main reason seems to be that research projects in SSH are traditionally more limited in scale and scope compared to the exact sciences. This limitation stems primarily from the lack of easily accessible digital datasets that cover multiple modalities (text, sound, image) indexed using uniform metadata systems. Efficiently exploiting the available datasets as linked open data in SSH still requires a considerable degree of expert domain knowledge, for example in ancient languages, which prevents scholars from answering large-scale research questions.

A crucial leap forward is TM's promise of delivering unified access to Europe's past as linked open data. This will revolutionise the individual researcher's search capabilities, ranging from yielding immensely rich results for simple term queries, to the option of exhaustively tracking individual cultural artefacts through time and space, including texts, paintings, ideas or places.

Moreover, Big Data of the Past will enhance our ability to deal with historical information, and in this way, it will drive SSH towards larger problems of world events and institutional development over longer periods of time. The abundance of new data about the past and the development of a new generation of AI will allow new interpretative models to be built on a superior scale.

Identifying larger patterns, correlations and connections will contribute to important advances in scientific approaches and methodologies that open new frontiers in our capacities for in-depth analysis and informed decision making.

Time Machine will, therefore, drastically raise the scale and scope of SSH research, ena-

bling it to effectively contribute to developing strategic answers to major challenges, such as sustainable growth, social welfare, migration and integration of migrants, and safeguarding European democracy.

Making Education More Accessible, Interactive and Diversified

Time Machine will develop new ways of delivering education, moving away from the traditional approaches of classroom lectures, textbooks and printed materials, and making use of localised and customised sample data at different – and extremely enhanced – levels of detail, enabling learning to be accompanied with new levels of analysis. The interactive environments that will be developed will not only offer unprecedented access to the records of our shared past, it will also promote active engagement with that heritage, which will make learning an on-going and inclusive process that will bind generations and cultures.

In this way, European history will become much more accessible to citizens of all ages and backgrounds, raising awareness of European culture, and consolidating our shared European identities. At the same time, there will be a stronger focus on exploratory learning, encouraging reflection on long trends that have shaped our present.

New interdisciplinary methods will also be developed across the traditional scientific domains, making use of the advances in AI, the new modelling capabilities in SSH, and new simulation capabilities in most of the scientific and educational domains, and offering more depth to educational curricula, sharpening the critical thinking of learners, and contributing to informed decision-making at all levels.

A Strong Boost in European Competitiveness in Big Data, AI and Other ICT Areas

Time Machine will develop new smart algorithms that can meaningfully extract information and create knowledge from noisy, heterogeneous and complex data at a massive scale, from medieval manuscripts, collections of objects and cultural artefacts, to the recent smartphone and satellite images, and the multimodal content from websites and platforms. Due to this data's inherent complexity, an AI trained on Big Data of the Past will offer a strong competitive advantage for Europeans in the global AI race. The new processing and simulation technologies of TM, combined with its global curation and exploitation platforms, will enable Europe to increase its share of gains from the announced AI revolution. Time Machine will also introduce disruptive technologies in deep reading, linguistic and knowledge systems, multimodal (4D) simulation, HPC and long-term data storage.

Many of the advances developed by Time Machine would require advanced software, an area in which many European companies have world leading roles. Time Machine will further enhance their leadership, through scientific and technological achievements in rapidly growing segments of the global ICT industry. All the core components of the Time Machine infrastructure will be developed as open source software, ensuring that the LSRI creates a legacy of code that can be corrected, reused and adapted for continuous development and deployment already during and after the end of the proposed initiative.

The simulation capabilities substantially enhanced by Time Machine will have a major impact on all areas of research in science and engineering that use HPC to process large volumes of data, including life and environmental sciences, as well as SSH. These technological breakthroughs will also have a strong impact on the European software and software-based services industry, one of the top drivers of Europe's industrial performance, and a key contributor to EU growth.

Time Machine will develop novel scanning technologies to digitise massive amounts of fragile documents and artefacts, through new types of sensors, robots and automated processes.

These will provide rapid scanning solutions in science, industrial archives, public administration, and potentially in services for consumers. Massive increases in demand for scanning services would lead to economies of scale and falling costs, and falling costs would further boost demand, offering a clear opportunity for European start-ups to compete in the growing Document Process Outsourcing (DPO) market, currently dominated by American and Asian players.

Enhancing Key Sectors of the European Economy

Creative Industries

The European creative industries contribute 6.8% of GDP and 6.5% of employment in the EU, at the same time offering a strong potential for stimulating innovation in other sectors with a competitive edge, such as tourism, education and advertising. Time Machine introduces disruptive technological solutions which will transform cultural heritage into rich creative assets. Specifically, the following creative industries sectors will benefit from Time Machine:

Gaming and film industries. Due to the inaccessibility and costs of filming in historic locations, the use of detailed 3D reconstruction will become increasingly relevant for film and game makers. Mass-scale digitisation will remove spatial and temporal barriers to cultural heritage: using complex 3D models constructed from rich heterogeneous sources, creators will be able to produce immersive, multi-sensorial experiences of historic sites from different periods, and customise them according to their needs. Four-dimensional cinema and virtual reality experiences on large (amusement park) and small (AR/VR headsets) scales will represent a new market for photorealistic 3D models and digitised audio-visual assets.

Design. Time Machine will harness the potential of the European heritage as a resource for creative reuse. Tools developed to analyse complex multimodal digital objects will enable designers to extract individual aesthetic features and concepts and re-appropriate them for new creations, from fashion to architecture. Advanced 3D modelling technologies combined with decentralised storage solutions will present new opportunities for real-time visualisations of objects and enable transnational, remote collaborative design processes. These enhanced engagement possibilities will translate into new avenues for creators to exploit cultural heritage sources as digital capital.

Media. With online video content expected to constitute 82% of all internet traffic in 2019, media professionals will continue to develop competitive business models and services that focus on providing seamless user experiences across devices and media types, and increasing the reach of their content. Time Machine will introduce disruptive high-performance computing and storage technologies for the curation, enrichment and distribution of audio-visual content. Building on these technological innovations and the availability of much richer datasets, media industries will develop smart services of a new level of quality that will close the gap between media content and audiences by offering personalised experiences.

Journalism. The lowered barriers for content creation and distribution via digital platforms has given rise to an unprecedented abundance of data that currently lacks effective verification and curation methods. Moving forward, the industry will rely even more on citizen journalism as well as computational journalism, both of which require knowledge-validation mechanisms to maintain the trust of their audiences. Time Machine will fulfil this need via deep reading tools that can effectively analyse and assess the quality of digital objects. To remain competitive in the changing landscape, media agencies will utilise the distributed processing and simulation infrastructure introduced by Time Machine to synthesise disparate datasets into facts and narratives that offer new, more transparent and engaging perspectives.

GLAM

The largest part of European cultural heritage can be found in the many galleries, libraries, archives and museums spread across the EU. Digitisation and open access dramatically change the way these institutions operate, putting in question current business models and funding mechanisms. Time Machine will contribute to the emerging need to strengthen GLAM by offering these institutions and their collections new ways of exploitation, based on:

- Promotion of innovative services through research, educational and creative activities
- New opportunities in fundraising and brand licensing combined with cost savings associated with rights and reproduction management overhead
- Employment of staff with more mission-critical activities, resulting in more efficient and less costly digitisation functions

Smart Tourism

Europe is the most visited tourism region in the world, and in the EU, tourism contributes 10% to EU GDP and creates jobs for 26 million people, through its direct, indirect and induced effects in the economy. To compete with strong competition from other world regions, Europe intensively invests in smart tourism, i.e., in smart, innovative and inclusive approaches to touristic development, paying particular attention to cultural heritage and creativity. As a resource, Time Machine is uniquely placed to revolutionise smart tourism, through the rich, multi-faceted and comprehensive cultural and historical data and innovative tools that will allow for the creation of endless avenues of experience and exploration, in situ or remotely.

Time Machine will also help to diversify cultural tourism by creating tailor-made data platforms for each type of tourism: cultural, slow, rural, cruise-based, ethical, sport event-driven, wellness, medical, business, or adventure. By offering a full package (online preparation at home, experience-based visit, post-visit experience with materials gathered on site), the visit will become both educational and experience-driven, enhancing tourist retention of local history and identity, and thus strengthening the sense and understanding of European belonging.

Smart Cities

A Smart City is a place where traditional networks and services are made more efficient with the use of digital and telecommunication technologies, for the benefit of its inhabitants and businesses. Smart Cities are at present mainly based on the networking of current data from interconnected sensory devices. The exchange of information will be fundamentally intensified by Time Machine by, on the one hand, enriching it with so far unconsidered data, and, on the other hand, through data collections with historical depth.

The integration of urban information will further enhance risk mitigation in smart cities. For example, the large-scale provision of architectural layouts and infrastructural aspects facilitates the mapping of infrastructural needs. The inclusion of historical data makes it possible to track urban, land and infrastructure developments such as electricity networks, gas conduits, and drainage systems. The analysis of architectural, urban planning, land use, economic and social data collected over long periods can make it possible to predict the future development of specific areas, and thereby allow for more sustainable and efficient decision-making.

Land Use and Territorial Policies

Time Machine offers a long-term perspective for a transition to sustainable land management, based on our common history and with the perspective of the next generation. By increasing the possibilities of in-depth analysis of highly heterogeneous data (aerial and spatial images, land

database, soil quality database), Time Machine will develop methods and indicators to objectively monitor the consumption of natural, agricultural or forest areas at all territorial scales. The increased potential of Time Machine in digitising and indexing old data (silver prints, maps, tables, legal texts), will also make it possible to retroactively calculate these indicators, which offers a unique opportunity to analyse the impact of policies on land use over time.

The Time Machine platform for land use will enable a shared framework to compare territorial configurations across space and time, promoting the exchange of knowledge and experience between similar territories. It will also strengthen the science-policy interface, enabling land-use stakeholders in particular to simulate the effects of transposing European legislation into national and local levels.

The Particular Focus on Nurturing Talent and New Jobs

Time Machine will create a sharp increase in the demand for digital and traditional humanists and social scientists at a time where these disciplines and corresponding university degrees do not guarantee jobs in these fields. Time Machine will also promote and create jobs for the new profession of "Digital Humanities expert".

A Powerful Ally to Counteract Gender Bias

The enlargement of available historical data via the systematic digitisation campaigns under the Time Machine initiative will result in much more knowledge on women in societies of the past. New knowledge will also be available about many repressed groups and minorities in the pre-modern past.

TM-driven technological improvements in areas such as stylometry, or intellectual authorship recognition, will provide us with a much more accurate picture of the female contribution to human civilisation, which, because of TM's scale, will fundamentally alter historical gender studies.

The Strong International Perspective for European Excellence

Europe has a leading role in the digitisation of culture. Time Machine will strengthen this role at a time where this field gains momentum in Asia and the USA.

Time Machine will increase the scientific reputation of EU institutions in Digital Humanities and SSH, opening the way to launch or reinforce international collaboration.

Time Machine comes at a time where culture occupies a central role in the UN 2030 Agenda for Sustainable Development. For the first time, the international development agenda refers to culture within the framework of Sustainable Development Goals related to education, sustainable cities and peaceful and inclusive societies.

A strong opportunity exists to develop synergies with UNESCO's Culture Conventions on the safeguarding and promotion of cultural and natural heritage, and the cultural and creative industries, as well as joint programmes with other UN Agencies for the implementation of the 2030 Agenda for Sustainable Development.

7 THE EVOLUTION OF THE TIME MACHINE COMMUNITY

7.1 The Time Machine Consortium

The Time Machine concept has brought together a very broad Time Machine partnership of leading European academic and research organisations, cultural heritage institutions and private enterprises. The founding members that are also part of the CSA Consortium comprise the following 33 organisations:

- ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE
- TECHNISCHE UNIVERSITAET WIEN
- INTERNATIONAL CENTRE FOR ARCHIVAL RESEARCH
- KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN
- NAVER FRANCE
- UNIVERSITEIT UTRECHT
- FRIEDRICH-ALEXANDER-UNIVERSITAET ERLANGEN NUERNBERG
- ECOLE NATIONALE DES CHARTES
- ALMA MATER STUDIORUM - UNIVERSITA DI BOLOGNA
- INSTITUT NATIONAL DE L'INFORMATION GEOGRAPHIQUE ET FORESTIERE
- UNIVERSITEIT VAN AMSTERDAM
- UNIWERSYTET WARSZAWSKI
- UNIVERSITE DU LUXEMBOURG
- BAR-ILAN UNIVERSITY
- UNIVERSITA CA' FOSCARI VENEZIA
- UNIVERSITEIT ANTWERPEN
- QIDENUS GROUP GmbH TECHNISCHE UNIVERSITEIT DELFT
- CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
- STICHTING NEDERLANDS INSTITUUT VOOR BEELD EN GELUID
- FIZ KARLSRUHE – LEIBNIZ-INSTITUT FUR INFORMATIONS INFRASTRUKTUR GMBH
- FRAUNHOFER GESELLSCHAFT ZUR FOERDERUNG DER ANGEWANDTEN FORSCHUNG E.V.
- UNIVERSITEIT GENT
- TECHNISCHE UNIVERSITAT DRESDEN
- TECHNISCHE UNIVERSITAT DORTMUND
- OSTERREICHISCHE NATIONALBIBLIOTHEK
- ICONEM
- INSTYTUT CHEMII BIOORGANICZNEJ POLSKIEJ AKADEMII NAUK
- PICTURAE BV
- COMPUTER VISION CENTER
- EUROPEANA FOUNDATION
- INDRA
- UBISOFT ENTERTAINMENT SA

7.2 The Time Machine Ecosystem

In addition to the **33 partners** of the CSA project, the Time Machine partnership network comprises over **400 organisations from 34 countries** (as of July 2019), including associate members that have stated their strong commitment to participate in the LSRI:

- Europe's top-level **academic and research expertise** – 170 academic and research institutions – for all key science and technology challenges in the project
- A huge representation – 93 organisations – from **Galleries, Libraries Archives and Museums (GLAM)** providing cultural, historical and geographic material and expertise to TM
- **Private sector partners** – 74 enterprises – that will contribute to the actual implementation of the TM infrastructure and/or the development of services around TM
- **Institutional bodies**, including:
  - The Italian Ministry of Culture and the French Ministry of ecological transition and solidarity
  - The Regional Office for Science and Culture in Europe of UNESCO
  - Several national cultural heritage agencies (Belgium, Netherlands etc.)
- **Strong civil society and industry associations**, including:
  - Europa Nostra, an NGO with a long contribution to the development of heritage-related policies at EU level in cooperation with the EU institutions and the Council of Europe
  - The Big Data Value Association (BDVA) that represents 190 EU Data Users, Data Providers, Data Technology Providers and Researchers.

The members of this unique alliance are fully aware of the huge potential of digitisation and the very promising new paths for science, technology and innovation that can be opened through the information system that we propose to develop, based on the big data of the past.

This community has already established strong links through extensive joint research in EU and national projects. The CSA will enable the Consortium to create a dense TM ecosystem of leading scientists and innovators for the big data of the past. The impact of the CSA will be measured by a substantial increase in the current number of TM supporters, the target being to reach the figure 2000 TM supporting organisations by the end of the CSA (1st quarter 2020).

7.3 The Time Machine Organisation

The Time Machine Organisation (TMO), officially launched in June 2019, is the leading organisation for cooperation in technology, science and cultural heritage and is constituted by the TM ecosystem. It is an internationally oriented association under Austrian law and head-quartered in Vienna. As such, the association is open to any type of legal entity that deals with science, technology and cultural heritage.

TMO aims to provide the institutional framework that ensures the sustainability and economic independence of the Time Machine initiative, by creating a strong alliance between cultural heritage institutions, science, industry and civil society. In this capacity, TMO will foster the development of technologies for capturing, processing and using historical data and will create bridges between disciplines, while supporting open access to and economic reuse of cultural heritage.

More information on the Organisation`s constitution and access to membership can be found in: https://www.timemachine.eu/time-machine-organisation.

Definitions

Definition	Explanation
4D Simulator	2020-One of 3 TM Simulation Engines. The 4D Simulator manages a continuous spatiotemporal simulation of all possible pasts and futures that are compatible with the data. The 4D Simulator includes a multiscale hierarchical architecture for dividing space and time into discrete volumes with a unique identifier: a simulation engine for producing new datasets based on the information stored. Each possible spatiotemporal multiscale simulation corresponds to a **multidimensional representation** in the 4D computing infrastructure. When a sufficient spatiotemporal density of data is reached, it can produce a 3D representation of the place at a chosen moment in European history. In navigating the representation space, one can also navigate in alternative past and future **simulations**. **Uncertainty and incoherence are managed at each stage of the process and directly associated with the corresponding reconstructions of the past and the future.**
Big Data of the Past	A huge distributed digital information system mapping the social, cultural and geographical evolution. A key objective of Time Machine is that such system brings together dense, interoperable, standardised (linked data, preferably open) and localised (marked up with spatial-temporal information) social, cultural and geographical heritage resources.
Communities	2023-2025 Group of users, self-organised by territorial or transversal interests, offering various voluntary works and favours (annotation, digitisation, bibliographic research, development), according to the standards in place, to the partners. Those communities can elect a representative.
Digital Content Processor	Automatic process extracting information from documents (images, video, sound, etc.). Level 1 Digital Content Processor just label mentions of entities. Level 2 Digital Content Processor label relations between entities. Digital Content Processor of Level 3 label Rules. Each processing is fully traceable and reversible. The results of the processing constitute the core dataset of the Big Data of the Past and are integrated in the TM Data Graph.
Large-Scale Inference Engine	One of 3 TM Simulation Engines. The Large-Scale Inference Engine is capable of inferring the consequences of chaining any information in the database. This permits to induce new logical consequences of existing data. The Engine is used to shape and to assess the coherence of the 4D simulations based on human-understandable concepts and constraints. Its origin comes from more traditional logic-based AI technology, slightly overlooked since the recent success of the deep learning architecture, that can, nevertheless, play a key role in an initiative like TM.
Local Time Machine	Zone of higher **"rebuilding the past activities"** density. Constituted of a group of local partners and communities bound by a common territorial focus and a declaration of intent, which respect both graphical and values charters. Any institution who meets eligible criteria can integrate a Local Time Machine. The declaration of intent is reviewed on an annually basis (time for new partners to integrate TM)
Project with Time Machine label (PWTML)	Project respecting the technical charter, whose tasks are documented – modelled within the Time Machine graph. All the partners of a PWTML must have signed the declaration of intent of the related Local Time Machine.
Technical Charter	Should contain information about infrastructure standards required within any project with Time Machine label. The Technical Charter defines the Time Machines Rules, Recommendations, Metrics and Official software. The document is revised periodically.
Time Machine Box	Servers that allow partners to store their documents and metadata and integrate easily the Time Machine Network and be appropriately documented in the Time Machine Graph. The Time Machine Box is part of the Time Machine Official Components.
Time Machine Data Graph	Formal representation of knowledge extracted by human or automatic process, represented with semantic web technology

Definition	Explanation
Time Machine Index	The TM index is a global system indexing different type of objects: e.g. documents; iconography; 3D geometries. It gathers all information regarding documents and their contents. Could be used as a basis for other search engine infrastructures (allows backups).
Time Machine Infrastructure Alliance	Coalition of TM's partners regrouping in-kind donators for infrastructure components (server's space and computing power).
Time Machine Mirror World	One of the API of Time Machine using the processing of the 3 TM Simulation Engines to produce a continuous representation model that can be accessed as information stratum overlaying the real world.
Time Machine Network	Set of all the partners actually interacting in Time Machine. Each member of the Time Machine Network must have signed the Value and Technical Charter
Time Machine Official Components	Pieces of software (e.g. Time Machine Box) that help partners conforming to the Time Machine rules as they are directly embedded in the software.
Time Machine Operation Graph	Formal representation of the past, on-going and future operations of the partners in the Time Machine Network and the data pipelines.
Time Machine Organisation	Association regrouping the Time Machine Partners. Some maybe active and other not. Not all may have signed the Values and Technical Charters.
Time Machine Recommendations	Recommendation on technology which are not obligatory at this stage for the development of Time Machine (e.g. choice of a particular IIIF image server).
Time Machine Request for Comments	Main document for the progressive design of the Time Machine infrastructures, standards, recommendations and rules, inspired by the process used for 50 years for the development of Internet Technology, today administrated by the Internet Engineering Task Force (IETF) as part of Internet Society (ISOC).
Time Machine Rules	Standard and rules that need to be followed to be acceptable in the Time Machine Network and become a Time Machine operators. Any entity not following these rules are out.
Time Machine Standard Contracts	Set of standard contracts to facilitate the interaction between Time Machine partners.
Time Machine Standard Metrics	Measures helping partners of the Time Machine Network coordinate with one another to compare performance (for quotes of services, but not only, there are also use for research performances, etc.).
Time Machine Super Computing Architecture and Simulation Engines	TM Super Computing Architecture composed of distributed computing resources from the TM Network provided by the TM Infrastructure Alliance. On this distributed architecture, different typologies of computing process can run. For instance, Digital Content Processors are intrinsically easier to run in parallel, whereas Simulation engines, which allow users to generate possible pasts and futures from the TM Data Graph need for more specific computing architecture.
Universal Representation Engine	One of 3 TM Simulation Engines. The Universal Representation Engine manages a multidimensional representation space resulting from the integration of the pattern of extremely diverse types of digital cultural artefacts (text, images, videos, 3D), and permitting new types of data generation based on transmodal pattern understanding. In such a space, the surface structure of any complex cultural artefact, landscape or situation is seen as a point in a multidimensional vector space. On this basis, it could generate a statue or a building, produce a piece of music or a painting, based only on its description, geographical origins and age.
Values Charter	Conform to the principle of openness in EU law

List of Abbreviations

Abbreviation	Explanation
AI	Artificial Intelligence
CH	Cultural Heritage
GLAM	Galleries, Libraries, Archives, Museums
LTM	Local Time Machine
PWTML	Project with Time Machine Label
RFC	Request for Comments
SSH	Social Sciences and Humanities
TM	Time Machine
TMO	Time Machine Organisation

Time Machine

timemachine.eu/trailer

@TimeMachineEU

@TimeMachineEU

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 820323.

timemachine.eu