Time Machine: Big Data of the Past for the Future of Europe

Deliverable D8.2
TM Preparation Report 2

Abstract
The second Interim Progress Report describes how the CSA project advanced in the second 3-month period in relation to the quality criteria set out for the development of the Time Machine Pillar roadmaps and lists key activities planned for the next 3 months.

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

Project Identification

Project Full Title	Time Machine: Big Data of the Past for the Future of Europe
Project Acronym	TM
Grant Agreement	820323
Starting Date	1 March 2019
Duration	12 months

Document Identification

Deliverable Number	D8.2
Deliverable Title	TM Preparation Report 2
Work Package	WP8
Delivery Due Date	31 August 2019
Actual Submission Date	13 September 2019
Leading Partner	EPFL
Main Author	Kevin Baumer
Contributions	Andreas Maier. Mike Kestemont, Gustavo Riva, Thomas Aigner, Daniel Jeller, Harry Verwayen, Dorit Raines, Valerie Gouet, Julia Noordegraaf, Julia Fallon
Reviewer(s)	Frederic Kaplan

History of Changes

Date	Version	Author	Comments
10.09.2019	1.0	K.Baumer	Document ready for submission

Disclaimer

List of abbreviations

AI	Artificial Intelligence
CH	Cultural Heritage
CSA	Coordination and Support Action
ERIC	European Research Infrastructure Consortium (Legal entity for Research Infrastructures)
FAIR	Findable – Accessible – Interoperable – Reusable
GLAM	Galleries, Libraries, Archives and Museums
HE	Horizon Europe (The 9th Framework Programme for Research and Innovation of the EC)
ICT	Information and Communication Technologies
KPI	Key Performance Indicator
LSRI	Large Scale Research Initiative
RFC	Requests for Comments
SSH	Social Sciences and Humanities
TM	Time Machine
WG	Working Group
WP	Work Package

Table of contents

1 Introduction

Time Machine (TM) is a Large-Scale Research Initiative (LRSI) built around the vision of creating the big data of the past, a distributed digital information system mapping the European social, cultural and geographical evolution[1]. This huge 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.

The objective of the TM CSA project is to prepare a detailed proposal for the Time Machine LSRI. The design is organised around four pillars, namely science and technology (pillar 1), TM operation (pillar 2), exploitation avenues (pillar 3) and framework conditions (pillar 4). The methodology foresees the elaboration of draft roadmaps for each pillar by working groups composed of consortium experts, followed by a round of consultations with all targeted stakeholders. The draft roadmaps have been prepared in June 2019. The consultations, currently in progress till September 2019, 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 the Time Machine and is currently expanding at fast rate.

The LSRI proposal is being elaborated in WP8 of the CSA project. The work involves close coordination and follow-up of actions conducted in the different work streams of the CSA, in particular those of WPs 2 to 5. For this purpose, quality criteria have been defined (Deliverable D8.1) to ensure that the specifications for the final outcome are taken into consideration during the design process.

The present document is the formal deliverable D8.2 Interim Progress Report 2, prepared under Task 8.1 of WP8. It assesses progress in the different stages of the LSRI maturing process in the second quarter of the project against the quality criteria for the pillar roadmaps. The reference documents used for the assessment of work in Pillars 1 to 3 are the draft roadmaps (deliverables D2.1, D3.1 and D4.1). Pillar 4 started on Month 5 (July 2019). Its progress, will be discussed in the next interim report (D8.3).

Progress in pillars 1 to 3 is presented in section 2 and assessed against the quality criteria in section 3, while an overview of next steps is provided in section 4. The focus is on the pillar objective, the research and innovation plans, the funding sources and the stakeholders to be involved. The other two topics of the roadmaps, the framework conditions and the risks and barriers, are related to work in progress in pillar 4, so they will be included in the next interim report (D8.3).[2]

The deliverable is supported by Annex A that describes the structure of the TM LSRI together with the roadmap methodology and Annex B that lists the quality criteria for the final TM document.

The intended audience comprises all members of the CSA project team.

[1] The expression Cultural Heritage in the document will often refer to every trace of European social, cultural and geographical evolution, which is wider than its current scope. A unique characteristic of this project is to design solutions respecting the cultural wealth of Europe as these developed for cultural heritage.

[2] In each pillar, the working groups have identified a preliminary list of framework conditions and risks and barriers (D2.1, D3.1 and D4.1). These serve as input to the work in pillar 4 that will design the accompanying actions supporting the implementation and outreach of Time Machine.

2 Progress achieved during the reporting period (Months 4-6)

2.1 Pillar Objective

Time Machine proposes an integrated programme with concrete objectives to be reached in its different components. Each pillar of the LSRI focuses on a specific feature of the Time Machine vision, i.e. creating the big data of the past and turning it into a living social and economic resource for Europe:

- Pillar 1 addresses the scientific and technological challenges in AI, Robotics and ICT for social interaction, for developing the big data of the past, while boosting these key enabling technologies in Europe.
- Pillar 2 intends to develop the overall management framework, build the Time Machine infrastructure for digitisation, processing and simulation and create the basis for and engagement with the communities participating in the development and use of Time Machine across Europe and worldwide.
- Pillar 3 will create innovation platforms in promising application areas, by bringing together developers and users for the exploitation of scientific and technological achievements, and therefore leveraging the cultural, societal and economic impact of Time Machine.
- Pillar 4 will look at developing favourable framework conditions for the outreach to all critical target groups, and for guiding and facilitating the uptake of research results produced in the course of the LRSI.

Table 2-1 below provides the mission statements developed for the four pillars of the Time Machine LSRI.

Table 2-1: Mission statements for the Time Machine Pillars

Pillar	Mission statement
Pillar 1 – Science and Technology for the big data of the past	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. The aim is 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 focus is, therefore on three main areas: • Data • Computing and AI • Social Sciences and Humanities (SSH)
Pillar 2 – Time Machine Operation	To put in place the constituent parts of the Time Machine infrastructure and the management principles and processes for an ecosystem of Time Machine contributors and users extending across the EU. The sustainable management and operational model to be developed is based on: • A governance scheme around a Time Machine Organisation (TMO) that sets out the global rules for the organisation and operation of the Time Machine communities. • A Time Machine processing infrastructure, composed of a digital content processor and three simulation engines: a 4D simulator, a large-scale inference engine and a universal representation engine. • Local Time Machine (LTM) projects aiming to develop and exploit the big data of the past in local/regional geographic environments, as parts of a pan European network[3].
Pillar 3 – Exploitation Avenues	To demonstrate how the scientific & technological advances (Pillar 1) and operational models (Pillar 2) enable us to work towards the vision of developing the

[3] The LTM idea is further discussed below (subsection on Targeted Achievements, Box 2-3)

Pillar	Mission statement
	big data of the past, and in turn how that is foreseen to provide social and economic impact across a range of areas of potential exploitation avenues. The main areas explored cover: • Scholarship • Education • Specific exploitation areas and uses, including GLAM, Creative Industries, Smart Tourism, Smart cities & urban planning, and Land Use and Territorial policies.
Pillar 4 – Outreach and innovation	To anchor the Time Machine in an international scholarly and innovation community, to develop framework conditions to support the outreach of TM to all critical target groups and to identify and valorise TM's innovation output. Time Machine will have efficient mechanisms to transfer, disseminate and valorise IPR to all relevant stakeholder groups. It will unite and connect the stakeholder communities around cultural heritage, digital humanities ICT & AI and will successfully introduce these topics into all legal, standardisation and policy initiatives as well as in the funding landscape.

2.2 Research and Innovation plans

a) State of the art

In **Pillar 1**, a variety of domains in science and technology are involved, each having its own methodological traditions and discipline-specific challenges. A clear-cut taxonomy was developed, following the overall three-branch structure of the main thematic areas, i.e. Data, Computing and SSH. In this way, the Working Group was able to identify and analyse the areas in science and technology which can be expected to be most relevant for the science and technology advances in the Time Machine initiative (Box 2-1).

Box 2-1: Taxonomy of Relevant Areas in Science and Technology (Pillar 1)

1. DATA

1.1. Data Acquisition: 2D digitisation**;** 3D digitisation**;** Audio digitisation**;** Film and video digitisation**;** Scientific analysis

1.2. Data Modelling: Knowledge Modelling**;** Data formats**;** Metadata Formats and Mapping between Standards; Annotation

1.3. Long Term Preservation: Bitstream layer**;** Functional layer**;** Semantic layer**;** Trustworthy archives

2. COMPUTING AND ARTIFICIAL INTELLIGENCE

2.1. Computer Vision and Pattern Recognition: Text recognition**;** Graphic document processing**;** Image processing and analysis**;** Indexing and Retrieval**;** Understanding and Interpretation**;** Recognition and Detection; Person, Face Identification; Modelling, Registration, and Reconstruction; Audio recognition & transcription

2.2. Natural Language Processing: Methods for Resource Scarce Languages; Orthographic normalisation and variation handling; Machine reading / Document understanding / Question answering; (Structured) Metadata extraction, manipulation, and translation/mapping; Discourse analysis

2.3. Machine Learning and Artificial: General Artificial Intelligence**;** Supervised Learning**;** Unsupervised Learning**;** Weakly Supervised Learning**;** Transfer Learning**;** Deep Learning**;** Universal Representation Space**;** Explainability**;** Bias / Fairness / Ethics in AI

2.4. Human-Computer Interaction and Visualisation: User-centred Interfaces; Access to large-scale information retrieval and recommender systems; Virtual / Augmented / Mixed Reality; Accessibility and Learning, Adaptive, and Cognitive Interfaces; Motivational Design; Big data visualisation; User Experience; Virtual research environments

2.5. Computer Graphics: Rendering; Animation; Immersive, Virtual, and Augmented Reality; Interactive Computer Graphics and Computer Games; Procedural Content Generation

2.6. Super Computing: Scaling and distribution; Dynamic provision of computing platform; Cloud computing; Secure distributed computing

3. SOCIAL SCIENCES AND HUMANITIES

3.1. Theory: Qualitative vs. quantitative studies: resistance and acceptance**;** Increase research scope in SSH**;** Simulation studies**;** Digital methods

3.2. Disciplines: History; Language and literature; Archaeology; Art history & media studies; Geography and demography; Musicology; Digital humanities; Urban studies

In **Pillar 2**, a detailed analysis was carried out for the different aspects related to the Time Machine infrastructure, targeted communities of users and the LTMs.

For the infrastructure, the focus was on current practices, available technical solutions and recent developments. The areas covered concerned:

- Mapping of cultural heritage resources
- Digitisation infrastructure and technologies
- Storage infrastructure and technologies
- Linked open repository
- Generic document processing
- High performance computing
- Deep learning frameworks
- 4D technologies
- Inference engines

For the communities, the analysis looked into the characteristics and needs of different groups, comprising developers, educators, professionals in GLAM and similar institutions, scholars and volunteers.

For the LTMs, the aim was to examine the current state of play for the key aspects required for a sustainable and synergetic pattern of operation. These include the data standardisation and interoperability, the selection and connection of resources, the legal framework, the financial model and the labelling system for identifying the cultural heritage resources involved. Emphasis was also given to recording the approaches and achievements of current local TM initiatives.

The state-of-the-art analysis in Pillar 2 enabled the WG to clarify concepts, identify operational objectives, and define the main lines of intervention for the Time Machine initiative within the existing framework characterising activities related to cultural heritage.

In **Pillar 3**, the baseline review was organised around the three thematic areas: scholarship, education and specific exploitation avenues and uses.

For scholarship, the investigation addressed the current bottlenecks in SSH research and the limitations of existing platforms to address them. The analysis enabled the WG to develop an overview of existing SSH research infrastructures that have taken the form of ERIC and to stress the synergetic effects of the collaborative schemes that could be developed: ERIC structures will serve to provide sustainable access to the big data of the past and make it available to researchers and other users in the various associated domains/disciplines, while Time Machine will offer new methods and tools for digitisation and information extraction. This mutually beneficial cooperation is expected to open new research approaches and methodologies in SSH.

With regards to education, Time Machine is geared to all forms of learning, from pre-school to higher education and including lifelong learning, vocational training and all forms of informal learning. The main target group are the educators, while the primary objective is to enable the developers of educational materials to produce new tooling within education platforms to enhance learning. Therefore, the base line review examined: (a) the general “web of knowledge”, including search engines, wikis and other open sources of information; (b) platforms and tools designed to extend educational content, such as virtual learning environments and massive open online courses; (c) tools intended to managed education environments, including teaching and studying practices; (d) tools designed to extend education skills, including critical and analytical thinking, by way of, e.g., analytical tools.

The specific exploitation avenues and uses were chosen based on criteria related to the relevance for Europe, including the opportunity to develop European leadership, the potential of technology breakthroughs leading to disruptive effects, and the substantial societal and economic impacts that can be expected. The specific exploitation avenues that were chosen (Box 2-2) are not to be considered as mere silos; hence open innovation can rely on transversal results.

Box 2-2: Selected exploitation avenues

- **GLAM**: With GLAM institutions acting as one of the main contributors in storing, collecting, describing, curating, discussing, exhibiting, and sharing Europe’s Cultural Heritage, but also as processors and users of Time Machine data, the interplay between the Time Machine initiative and GLAM institutions is a complex and multifaceted relation.
- **Creative industries**: the capacity of creating and mobilising people imagination is key for our future to reach a vision aligned with our values.
- **Smart tourism**: a specific domain where technology and demand readiness are at good levels and Time Machine has a huge potential of return on investment.

- **Smart Cities and urban planning**: underlining the capacity (individually and as a society) to interact with the environment and design it. Cities concentrate most human activities and are a privileged exploitation avenue.
- **Land use and territorial policies**: shares the same stakes as smart cities, but territories in general are also studied to address sustainable development challenges in a context of climate change.

In each case, the baseline analysis aimed at determining the key aspects that could be relevant to innovative business models: stakeholders, technology and demand readiness, as well as needs in terms of new tools and processes to be developed from the TM processing and simulation infrastructure.

b) Targeted achievements

The design of Time Machine builds on the following key concepts (Box 2-3):

- The Time Machine Organisation for the overall governance of initiatives and actions related to the big data of the past.
- The Time Machine digitisation and processing infrastructure.
- The Local Time Machines.

Box 2-3: Key concepts for the overall design of the Time Machine

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 based on organisational requirements formulated in each pillar.

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 cultural heritage (Europeana, Archive Portal Europe, etc.). Time Machine will introduce new processing pipelines for transforming and integrating cultural heritage 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 composed of a digital content processor and three simulation engines: a 4D simulator, a large-scale inference engine and a universal representation engine:

- 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 the TM 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 Time Machine.

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 high *density of "rebuilding the-past activities*".

The Local Time Machines follow the rules developed by the TMO. 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). The TMO provides help in their launch and growth.

The corresponding management structures, methods and tools are planned to be implemented in pillar 2, following advances made in the other pillars, as discussed below.

Pillar 1

A number of specific breakthroughs and innovations are targeted to address the scientific and technological challenges related with the big data of the past. Each of these targeted achievements involves work in particular areas of expertise according to the taxonomy for Science and Technology developed during the state-of-the-art analysis (Box 2-1). The order of relevance of these areas of expertise for each specific goal has been specified in the pillar 1 roadmap (Deliverable D2.1).

Data

The aim is to enable persistent digital access to more than 3 millennia of linked historical data, which requires the development of:

- The Time Machine Data Graph: the formal representation of knowledge extracted by human or automatic processes, represented with semantic web technology that needs the elaboration of: (a) criteria on priorities of objects to digitise, taking into account the different states of conservation, availability, proprietary status and degree of emergency for endangered objects; and (b) guidelines and standards to follow regarding formats and protocols to store and query data, promoting trustworthiness and FAIR principles.
- Digitisation Hubs: hotspots of local digitisation, allowing digitisation outcomes to be seamlessly aggregated into a pan-European Cultural Heritage data infrastructure, with the appropriate standards in terms of resolution, file formats, and metadata during acquisition, using dedicated scan technologies, such as scan robots and tomographic methods.
- The Time Machine Box: the distributed storage system where the Time Machine Data Graph will be hosted, characterised by adapted technical server infrastructure, compliance with international standards, certification processes, de-duplication methods leveraging pattern-recognition across large datasets, together with a 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.

Computing and AI

Generic methods will be developed to explore, connect, and simulate historical information, including:

- Interface for Annotation: an interface to the Time Machine Data Graph, allowing for easy but complex annotation, complying with the standards set for data modelling.
- User Interface: a central interface and templates for specific applications, enabling users of the Time Machine to access the data and materials produced, with features and requirements defined by the user communities.
- Natural Language Processing Tools for Older Language Variants: Processing tools allowing for handling documents in multiple European languages and dialects, named entity recognition in older European languages and variants, orthographic normalisation of older European language variants, as well as machine translation adapted to older European language variants.
- Digital Content Processor: Processor with capabilities that will evolve from labelling mentions of entities (Level 1), to creating labels to establish relationships between entities, improving the Data Graph (level 2) and then to creating re-useable models that generalise from few observations and contribute to possible understanding of the patterns behind the available data (Level 3).
- Time Machine Engines: the design of the Time Machine digitisation infrastructure and its components, as described in Box 2-3 above.
- Automatic Text Recognition: general models for text recognition 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.

- Automatic Graphic Document Recognition: improved methods and results of automatic map recognition, including automatic music scores recognition, developed using material of the Time Machine Graph.

SSH

Explanatory models of historical evidence opening the way for new, plausible narratives, radically transforming the manner in which SSH engage with and interface with the past will be elaborated, leading to:

- A new framework for researchers in historical subjects (history, literature, art, musicology, etc.) using the Time Machine Data Graph to perform quantitative historical studies with a 'longue durée' perspective. Increased acceptance of quantitative studies in SSH research will be achieved by organising dedicated conferences and open call for papers. The implementation strategy for this framework rely on tools that facilitate and enhance scientific analysis, like the Digital Content Processor and the Simulation Engines.
- Enhanced research methods, such as agent-based simulation, using linked data from the Time Machine Data Graph. Researchers will be able to use the Time Machine engines to perform simulations studies, without having to rely on outside models and tools.

Pillar 2

Pillar 2 aims to put in place the constituent parts of the Time Machine infrastructure and the management principles and processes for an ecosystem of Time Machine contributors and users extending across Europe.

Infrastructure

Specifications will be developed for the Time Machine hardware and computing infrastructure that will define the research challenges to be addressed in pillar 1. The research results will then be used to design and develop:

- A network of digitisation hubs on a European scale, managed by a peer-to-peer platform in charge of managing optimising digitisation strategies at European level, also tasked with the development of generic solutions for archiving, directly documenting the digitisation processes, and swiftly putting the digitised documents online.
- A distributed storage infrastructure for both public and private data in the form of a fully decentralised highly redundant architecture based on the shared resources of a purpose-built network forming the Time Machine Infrastructure Alliance partners.
- The distributed super computing infrastructure for processing big data of the past as described in Box 2-3, equipped with specially designed content and discovery interfaces for accessing the Time Machine Data Graph for all intended uses and applications.

Community Management

Work deals with the organised interaction of TM with scholars, developers, cultural heritage professionals, service providers and citizens. A system of platforms will connect Time Machine with such external communities that will benefit from and can provide input in various forms to the TM, so the aim is to:

- Build a strategy and an associated Community Management System, responding to well identified requirements for staff and processes leading to mutually beneficial and sustainable interactions with the Time Machine communities.
- Develop interfaces facilitating connection of TM infrastructure with those of existing communities[4].
- Design and implement transparent mechanisms for tracking community involvement and reporting on community contributions, through metrics on individual participation and overall impact on the Time Machine Data Graph.

[4] 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.

Local Time Machines

The thematic area is related to the governance scheme of the Local Time Machine network and by extension to the overall governance scheme of the TMO. So, this thematic area constitutes the core of the Time Machine sustainability model and is strongly related to the growth of the entire initiative, as it will set out the principles and processes for a network of Local Time Machines, including:

- The LTM common framework ensuring cohesion in the network's operation, through the definition of general values, common objectives, as well as technical standards and guidelines regulating data acquisition, data sharing and data publishing.
- The support structure that will oversee the smooth development of the LTM network, so that institutions wishing to launch or integrate an LTM are given a clear path and guidance during the whole process. Particular attention will be paid to enabling existing initiatives to be aligned with the LTM framework, as well as encouraging TM partners to launch new initiatives.
- The labelling system or value scale to be used for evaluating the progression of an LTM. The labelling system should encourage the progression of the LTMs through the different grades, while providing a means to assess members' commitment towards the LTM objectives.
- The legal setting based on a coherent and standardised contractual and licensing system for all LTM network operations, guaranteeing conformity with national and European policies and laws.
- The financial system to foster financial independence and, therefore, longer term viability of the LTM initiatives, including shaping a LTM franchise model. In this respect, emphasis will be on utilising local assets to enhance / develop new exploitation avenues for the big data of the past (in cooperation with pillar 3).

Pillar 3

Pillar 3 is designed to leverage the societal and economic impact of Time Machine. Two thematic areas, scholarship and education, focus on the disruptive effects on scholarly methods and learning, respectively. The third thematic area groups sectors of activity that are important for the EU, for which the LSRI is expected to introduce new approaches and transformative business and cultural models.

Scholarship

Time Machine 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, with a *longue durée* perspective. The new and more critical methods of analysis will open the way to meaningful outlooks for the future.

Because of its integrated approach that combines digitisation, interpretation & data gathering and management, as well as methodological innovation, the Time Machine infrastructure will drastically speed up advances in the state of the art in most SSH fields (and probably also in ICT), by multiplying the pace and explanatory power of scholarly and scientific progress. 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 micro-level 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 the following ways:

- 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.
- The analyses based on big data of the past can be used to empirically test existing assumptions based on smaller, sample data.
- The quantitative methods (including simulation) will allow for the combination of different types of data and thus for more complex analyses.

Education

Time Machine for Education will offer unique enquiry and experience-based blended learning, citizen science infrastructure and approaches based on revolutionary digital technologies (VR, AR, AI). Pedagogical content will be developed for schools, universities, and lifelong learning in a mix of free, sponsored and paid services, largely based on the big data of the past and associated simulation technologies. The emphasis will also be on accelerating the learning of SSH, through swift availability of many facts on a single subject, as well as 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 will also benefit, as Time Machine will enrich teaching material associated to SSH, the sciences, health and practical technologies.

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'[5]. From that perspective, it makes sense to design an infrastructure that provides direct access to the Time Machine data in ways that match the infrastructures for education currently in use and the new opportunities for innovating access to cultural and historical information.

As the Time Machine is centred on the use and application of big data of the past, the key advances are expected in the following areas:

- 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.
- 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, including maps with integrated 3D models, AR/VR applications, search engines and other systems based upon big data of the past.
- Critical thinking and digital literacy: supporting these applications are code and big data analysis training, or "Time Machine analytics" for all students and educators engaged in studying and teaching historical disciplines through data analysis. Critical thinking and digital literacy required for using such data will be developed in cooperation with pillar 1 - Theory.

<u>Other exploitation avenues and uses</u>

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 boost, aid and accelerate many developments that are already underway in GLAM and introduce completely new transformative effects in four areas dealing with (digital) collection(s):

- Collection Custodianship & Enrichment; (a) a larger body of digitised material will be made accessible to the general public according to FAIR data standards, with more affordable and flexible digitisation services; (b) the vast amount of newly acquired metadata will increase the demand for curation. Time Machine tools will help GLAM professionals to select and further refine metadata; (c) automated information extraction, machine learning, and AI will increase document understanding and automated translations (including translations from ancient languages to modern languages) and so the accessibility for all kinds of audiences.
- Collection Access: Novel query mechanisms will innovate current methods to query both digitised and born-digital content for the general public and researchers.
- Collection Curation, Engagement & Experience: Time Machine will enable institutions to provide richer and more diverse experiences for their users, both in a physical, augmented,

[5] EDUCAUSE Horizon Report, 2019 Higher Education Edition, page 19

and virtual setting: multimodal interfaces and feedback mechanisms will give ground-breaking multisensory experiences, that are elegant, authentic, nuanced, unobtrusive, and customisable according to the user's needs.

- Collection Linking, Reuse & Remix: (a) 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; (b) customised frameworks to reuse and remix data in intuitive ways will foster exploration by humans on crowdsourcing platforms, GLAM labs, and raw data APIs, or semi- and fully automated methods through the use of machine learning. These initiatives will further feed data and new knowledge back into Time Machine's databases; (c) monetisation and distribution of single objects and entire collections, including a discussion of "levels of openness" will introduce new business models for GLAMs.

Creative Industries

The European creative industries contribute 6.8% of GDP and 6.5% of employment in the EU[6], 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 will introduce scientific and technological breakthroughs that will significantly impact the production cycle of the creative, media and entertainment industries, through interventions to the creative value chain as the ones below:

- Creation - elaboration of ideas, contents and products: (a) 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; (b) the open and interoperable infrastructures for data exploration will enable creative freedom and diversity; (c) artificial intelligence will also support new forms of creativity, including computational creativity.
- Production/Publishing - the making of original, non-reproducible or reproducible work: (a) production processes will be supported by easily finable, high-quality resources, while smart metadata models will support the ability to combine and seamlessly integrate digital objects in different variations to tell different stories; (b) storytelling will be enhanced using ground-breaking simulations and visualisations and possibilities to query granular properties of digital objects will support the emergence of new kinds of storytelling techniques that appeal to different senses; (c) reuse of data will be supported by clear copyright acquisition and licensing mechanisms, and newly developed business models will ensure that both data providers and creators can benefit from these transactions.
- 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
- 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.

[6] http://www.teraconsultants.fr/en/issues/The-Economic-Contribution-of-the-Creative-Industries-to-EU-in-GDP-and-Employment

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.[7] To face strong competition from other world regions, Europe largely invests in smart tourism, which refers to smart, innovative and inclusive approaches to touristic development, paying particular attention to CH and creativity.

Taking into consideration the Smart Specialisation Strategy (3S) framework for regional development in the EU, Time Machine will work with territorial clusters to develop specific technological innovations and tools for local cultural-heritage experience platforms that raise smart tourism to a key local/regional priority. This approach is fully compatible with the development of LTMs (pillar 2) that could serve as backbones to these local "smart clusters" and is expected to lead into the following achievements:

- Synergy models for core re-users, enablers and infomediaries to offer products and services that are enhanced by Time Machine technologies that boost touristic demand based on cultural heritage.
- Innovative clusters working with LTMs to create a sustainable ecosystem of smart tourism.
- Increased awareness and respect toward cultural heritage destinations through TM narratives.
- Economic sustainability of destinations, locations and institutions (e.g. GLAM) through the Time Machine smart tourism model.

Smart cities, urban planning, land use & territorial policies

The aim is to use Time Machine technologies to achieve more inclusive societies as well as sustainable development in our cities and territories and to support the elaboration of common visions and projects for our cities, territories and Europe based on common values. The innovative solutions to be developed will support people in understanding their environment dynamics, identifying what are the choices they have to make when they design their environment (incl. European, regional and local regulations), as well as enable them to assess options, connect to other inspiring experiences, and learn to use data and state of the art knowledge. The main targets are described 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 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. "Culture ready" information systems that integrate cultural specificities of different information sources and contexts of use.
- "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.).
- 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-

[7] UNWTO (2018). European Union Tourism Trends: https://www.e-unwto.org/doi/book/10.18111/9789284419470

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.
- 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.

c) Proposed methodologies

Reaching consensus on the technology options to follow in a programme as large as Time Machine is a complex issue. To ensure an open development and evaluation of work, a process inspired by the Request for Comments (RFC) used for the development of the Internet protocol[8] will be adapted to the needs of Time Machine.

Time Machine Requests for Comments will be used as the main methodological process for establishing rules, recommendations and core architectural choices for the Time Machine components. The TM RFCs will be freely accessible publications, identified with a unique ID. Their basic principles are presented in Box 2-4.

Box 2-4: Basic features of the TM Requests for Comments

- **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

The TM RFCs will be accompanied by a set of fundamental research questions that need to be clarified by scientific project work, for example user studies required to prepare the development of a RFC. 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.

[8] https://en.wikipedia.org/wiki/Request_for_Comments

Once individual RFCs are developed, the roadmap will be implemented via a modular design covering a 10-year period, in which various **calls-for-proposal** will attract bottom-up research proposals targeting specific milestones in a pre-specified time-frame. The draft roadmaps for pillars 1-3 have already developed lists of RFCs and research topics leading to the targeted achievements discussed above.

The exploitation avenues will be based on a number of new capabilities or “impact facilitators” that pillars 1 and 2 will produce:

- Cheap and cost-efficient solutions for the further digitisation of resources through standardised offers and services and easily replicable open hardware technologies.
- Generic Automation for the mark-up of these resources tagging concepts, named-entities, relations and rules.
- Intelligent connection of existing fragmented data resources using, adopting and building on existing legal frameworks and developing standards for distributed storage solutions.
- New simulation capabilities, by transforming sparse data into continuous 4D data sets capable of representing multi-worlds.
- Innovative forms of experience, by enabling new paradigms for the restitution of the data to the end-user including spatio-temporal search engines, geo-historical services and Mirror Worlds.

The availability of these resources will create opportunities for social and economic impact for the potential users, which offers strong incentives for their participation in the design of solutions and their commitment in applying them. As examples, one can mention the cost saving potential for economic sectors like GLAM and the creative industries and the improved tools and access to high linked data featuring critical reference points that open new horizons for social science and humanities researchers.

d) Time plan - Milestones

The development of Time Machine is foreseen in four phases: bootstrapping, scaling, sustaining and globalising.

- Bootstrapping – Y1 to Y3 (2021-2023): During this first phase, the key technological components will be designed through RFCs (Pillar 1). This process will interact with the design of the TM infrastructure and community platforms; a number of new LTMs will be launched (Pillar 2). The Scholarship and Education thematic areas will design pilot projects for the new concepts introduced by Time Machine in the respective fields, while the other exploitation thematic areas will develop strategic plans based on detailed user needs analyses (Pillar 3).
- Scaling – Y3 to Y5 (2023-2025): Based on research progress, the industrialisation of next generation scanning and storage techniques will start (Pillars 1 and 2), and a first version of the three simulation engines will be launched (Pillar 2). The SSH and Education platforms in Pillar 3 will give first results, in terms of research and education methodologies, as well as new studies and curricula.
- Sustaining – Y5 to Y7 (2025-2027): The third phase of the project will be dedicated to addressing the challenges linked with open-ended sustainability of Time Machine (Pillar 2) and to new developments and demonstrations of the platforms for Scholarship, Education and Specific Exploitation Areas and Uses (Pillar 3). In the meantime, research in Pillar 1 will focus on the development of new integrated AI, combining progress made in the previous phases.
- Globalising – Y7 to Y10 (2027-2030): The last phase of the project will address the challenges linked with the extension to non-European archives and patrimony, while sustaining the European densification (Pillars 1,2,3).

The working groups have started producing detailed time-plans and milestones along these broad guidelines.

e) Key performance indicators

A set of Key Performance Indicators (KPIs) is being developed to assess performance. The current status is discussed below for each pillar.

Pillar 1

The KPIs are organised according to the taxonomy, although with different levels of granularity, as shown in Table 2-2. 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.

Table 2-2: KPI set to measure performance in pillar 1

Taxonomy area	Performance metrics
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 & 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	Results of user studies. Number of users of the TM interfaces.
Social Sciences and Humanities	Engagement of academia and research (Bibliometrics, Alt Metrics) with the TM through mentions in journals and books, initiatives and projects using the TM data or infrastructure.

Pillar 2

One key objective of the Time Machine is to increase the amount of cultural-heritage digitised datasets available across EU, therefore previous monitoring studies conducted within the cultural-heritage sector, for example the one by Europeana, will be used and supported.

As the targeted achievements of each thematic area of pillar 2, although complementary, have their own targets, specific Key Performance Indicators will contribute to monitor each of them, as indicated in Table 2-3.

Table 2-3: KPI set to measure performance in pillar 2

Thematic area	Performance metrics
Infrastructure	• Number of units and relation in the TM Data Graph • 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	The different communities will have different impact on the TM and therefore will have different indicators for efficiency and impact. Examples include for developers: known reuse of TM-Software and developed standards, for scholars: scientific articles using Time Machine data, for volunteers: degree of participation in the Time Machine Platforms. In addition to these user specific KPIs, a general one, referring to the actions aiming to increase the community outreach
LTMs	• Collaboration: ○ Average of partners involved per PWTML within an LTM ○ Degree of satisfaction of the different partners regarding collaboration ○ Capacity of the collaboration in place to enhance the goals of LTM projects ○ Degree of partners' participation in decision-making ○ Members participation's rate at LTM stakeholder meetings ○ Involvement of business sector in the projects • Growth of the TM network: ○ Number of new partners ○ Number of new LTMs ○ Number of new projects with LTM label ○ Activity of an LTM (number of projects launched over a period of time) ○ Engagement rates: active contributors in the TM networks • Labelling system: ○ Label adoption (number of labelled LTM and projects per label's categories) • Guidance and Training: ○ Number of new requests ○ Attendance to special events-trainings ○ Outcomes of training sessions

Pillar 3

The Key Performance Indicators designed for the exploitation avenues of pillar 3 are presented in Table 2-4.

Table 2-4: KPI set to measure performance in pillar 3

Exploitation area	Performance metrics
Scholarship	• Usage of the Time Machine data and tooling through indicators measuring: user statistics on Time Machine infrastructure; papers; publications and other scholarly output • Integration of the Time Machine data and tools in higher education curricula in the SSH and other fields[9] • Monitoring the impact of SSH publications on crucial topics (e.g. as evidenced by citation indices and attention in the media) • Economic impact can be assessed by looking at increase in number of start-ups that initiate in the SSH field[10]
Education	• 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 Time Machine-supported applications and tools and teaching Time Machine-provided big data of the past analytics • Number of national ministries of education to have committed to Time Machine content • Number of educational publishers to take up Time Machine content.

[9] For example, via the CLARIN-DARIAH Digital Humanities Course Registry, https://registries.clarin-dariah.eu/courses/

[10] For example, via Venture Labs as this one at UvA in Amsterdam: https://www.uva.nl/en/faculty/faculty-of-humanities/humanities-in-the-city/humanities-lab-avs/humanities-lab-avs.html

Exploitation area	Performance metrics
GLAM	• 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 in which GLAM institutions being a leading or contributing factor
Creative industries	• Collaboration and outreach o number of cross-sectoral collaborations o number of stakeholders from the creative industries joining the Time Machine organisation o growth in cultural participation and growth in social inclusions (e.g. number of products, experiences and services tailored for the disabled) • Proof of concepts o number of scenarios tested o number of stakeholders involved in Proof of Concepts • Incubation o representation of all creative industries' domains in the Local Time Machines o number of stakeholders connected to the Local Time Machines o number of entrepreneurial start-up and scale-up-stage businesses initiated as a result of the incubation efforts o number of strategic partnerships with relevant ancillary networks. • Support infrastructures o creative sector contribution to GDP o number of items available for reuse (high quality, using correct rights labels) o number of self-employed individuals and SMEs involved o employment in the creative industries • Sustainability o number of products and services developed for other sectors o income from licensing and use of TMO data and services o number of Time Machines with long-term financial stability
Smart tourism	The KPIs will be designed for any destination wishing to measure the sustainability of the smart tourism clusters along the following dimensions: • Raising awareness – emphasising the importance of obtaining relevant local political support for implementation • Creating a destination profile • Form a Stakeholder Working Group best suiting the destination and the groups involved • Roles and responsibilities in setting targets, action planning, and implementation • Data collection enabling the initiative to develop the destination's tourism industry • Analysing results and taking action on the basis of priorities • Enabling 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
Smart cities, urban planning, land use & territorial policies	• Number of European organisations related to urban planning and land use engaged • Number of national government bodies related to urban planning and land use engaged • Number of local government bodies engaged • KPI concerning linking and harmonisation of land use data • KPI concerning linking and harmonisation of urban planning data • Number of best practices regarding Big Data of the past for land use • Number of best practices regarding Big Data of the past for urban planning

2.3 Funding sources

The required extensive, long-term and sustained effort to reach the Time Machine ambitious objectives exceeds by far what can be achieved in typical national or European research and innovation projects. Moreover, as digital preservation of cultural heritage is a priority in almost every Member State, it is crucial to align national research agendas and manage multiple projects and networks in different European countries and cities.

Time Machine was initially conceived as a FET Flagship, which should receive half of its funding from Horizon Europe and the other half from other sources. The TM partners have, therefore,

examined the different forms of funding from the very start. Table 2-5 below indicates the different sources of funding, taking into account the current state of design of the various financial instruments that will be available in the upcoming programming period (2021-2027).

Most activities of TM pillar 1 and more research oriented contributions needed for TM pillar 3 relate to the Horizon Europe (HE) clusters Digital, Industry and Space and Culture, Creativity and Inclusive Society (HE pillar II). For the former cluster, Time Machine will develop multimodal historical and geographic datasets of an unprecedented semantic complexity that will give a new impetus for big data research, methods and application fields. Due to its transversal nature as a backbone for other critical technologies, Time Machine is expected to contribute to many other areas of HE, including AI, big data and machine learning in Pillars I and II of HE, and give a strong boost to SSH topics across pillar III of HE.

Table 2-5: Possible funding sources for the Time Machine pillars

Time Machine Pillar	**Funding sources**							
	Horizon Europe			Erasmus	Digital Europe	Creative Europe (II)	ESIF	Private funds
	Pillar I	Pilar II	Pillar III					
Pillar 1								
Data	x	x						
Computing	x	x						
SSH	x	x	x					
Pillar 2								
Infrastructure					x			x
Community Mgt					x	x		
LTM						x	x	x
Pillar 3								
Scholarship		x					x	
Education		x		x			x	x
Exploitation areas&uses								
GLAM		x	x		x	x	x	x
Creative industries		x	x		x	x	x	x
Smart tourism		x	x		x	x	x	x
Smart cities-related areas		x	x		x	x	x	x

Time Machine is particularly relevant to the Digital Europe programme broad area dealing with ensuring the wide use of digital technologies across economy and society that has a pillar specifically dedicated to education and culture. A number of activities in TM pillars 2 and 3 are fully aligned with these objectives. The Time Machine agenda will also have synergies with other broad areas, including supercomputing, AI, and advanced digital skills.

The Time Machine approach offers concrete methodologies and tools in line with the objectives of the Creative Europe successor programme. In this respect, Time Machine will also be crucial in giving a more active role to civil society and in developing a structured dialogue among communities of users, based on online collaboration opportunities.

Time Machine will also establish synergies with the Regional Development and Cohesion Funds that are well adapted to support the development of local Time Machines, as well as actions related to:

- Developing the TM infrastructure,
- Applications related to LTMs, as well as the exploitation avenues and uses considered in pillar 3.

Moreover, there are opportunities for raising private funding, especially in pillar 2 and pillar 3 actions. Such private investments are expected to be leveraged by specific instruments of the next programming period, including the COSME successor and the future InvestEU Fund, particularly in digital infrastructure, digital transformation of small businesses, research on digital technologies and

helping the social economy to benefit from digital transformation. Related funds will be relevant for developing the Time Machine infrastructure, initiatives related to new education programmes, as well as the specific application areas for GLAM, creative industries, smart tourism and smart cities.

2.4 Stakeholders to be involved

In addition to the **33 partners** of the CSA project, the Time Machine partnership network comprises **449** organisations from **34** countries (as of August 2019), including **associate members** that have stated their strong commitment to participate in the LSRI:

- Europe's top-level **academic and research expertise** – 184 academic and research institutions – for all key science and technology challenges in the project.
- A huge representation – 102 organisations – from **Galleries, Libraries Archives and Museums (GLAM)** providing cultural, historical and geographic material and expertise to the TM.
- **Private sector partners** – 84 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...)
- **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.

3 Assessment – next steps

3.1 Pillar Objective

Quality Criterion: A clear mission statement is developed for each pillar describing how the pillar contributes to the Time Machine vision.

The “mission statements” in Table 2-1 give a clear overview of the work to be conducted in each pillar and the way each pillar links to the broader objective of the LSRI, so no further work is needed in this topic.

3.2 Research and Innovation plans

Quality Criterion: A master plan is developed describing the baseline, the expected ambitious outcomes and the methods to achieve them.

More specifically:

a) A thorough state-of-the-art review of relevant scientific, technological and business fields constitutes the pillar’s baseline.
b) Concrete targets are defined for addressing the relevant scientific, technological and innovation challenges.
c) The path to the targeted achievements is elaborated to a sufficient level of detail - interdependencies with other pillars are identified.
d) Critical points to evaluate progress in the proposed developments are determined.
e) A set of metrics is developed to assess progress made in the different pillars / thematic areas.

a) State of the art

As discussed in section 2.2a, the state-of-the-art analysis was adapted to the requirements of each pillar:

- Pillar 1: the relevant domains in science and technology have been identified. The taxonomy that was developed enabled the WG to identify the scientific and technological challenges to be addressed to reach the Time Machine objectives.
- Pillar 2: The different aspects related to the Time Machine operation have been assessed, comprising the current practices, available technical solutions and recent developments in infrastructure, the characteristics and needs of the different target groups of Time Machine, and the key aspects needed for the viable operation of LTMs, based on their state of development and lessons learned.
- Pillar 3: The analysis was directed to identifying current limitations, opportunities, stakeholders, technology and demand readiness, as well as needs in terms of new tools and processes to be developed to facilitate new business models.

The WGs have been formed with experts covering the required knowledge-base in terms of science and technology fields and application areas. Unless otherwise concluded from the findings of the web consultation that is currently open to existing and potentially new members of the Time Machine network, no further work will be planned in this topic.

b) Targeted achievements

Clear targets to be achieved have been defined for all thematic areas covered by the three pillars. In most cases, these serve as intermediate steps contributing to the achievement of the objectives of each pillar, pointing to a well-defined implementation plan that can be outlined as follows:

- Time Machine is based on 3 basic concepts: the TMO, the TM digitisation and processing infrastructure and the LTMs.
- Initially, pillar 2 sets out the specifications on how to develop the basic concepts, in a way that respects the main idea of the Time Machine: creating a distributed digital information

system mapping the European social, cultural and geographical evolution that can be used by a number of communities to create momentous social and economic impact.
- These specifications are used in pillar 1 to define scientific and technical objectives that address the underlying challenges.
- Pillar 2 designs the physical and management infrastructure and undertakes the Time Machine operation according to the results of the designing process. Operation also includes overseeing the development of and providing support to the LTMs.
- Pillar 3 brings together communities that work on exploitation avenues in three thematic areas: two of them refer to scholarship and education and are targeting to demonstrate the benefits of new approaches in scientific methods and in learning; the third thematic area comprises important sectors of activity for the EU, for which the Time Machine is expected to introduce new approaches and transformative business and cultural models: GLAM, creative industries, smart tourism, as well as smart cities, urban planning, land use & territorial policies.

The overall aim of the target-setting process was to have a balanced approach when dealing with requirements for research and innovation actions on the one hand and the operational aspects relating to the implementation of new structures, processes, tools and modes of cooperation on the other. Feedback from the stakeholders in the currently on-going consultation process will enable to confirm that such a balance has been achieved.

The proposed LSRI covers the entire range from basic research to marketable innovations. Emphasis was given to the internal coherence of the programme, by setting out targets for the three pillars that are logically integrated and consistent with each other. In addition to the interrelations pointed out above, the current design identifies other areas where synergies can be further worked out:

- The thematic area SSH in pillar 1 is directly connected to the scholarship exploitation avenue in pillar 3 and to some extend also to the education exploitation avenue. The framework and tools to be developed in pillar 1 will be implemented/validated in pillar 3, by being applied to concrete test cases related to current research topics in SSH.
- The LTM profile can be significantly raised from synergies with the 3 topics of exploitation avenues referring to smart tourism, smart cities, and land use, based on the premise that: (a) LTMs will enable smart cities to add a ‘time dimension’ to their ICT-based management and development plans; (b) cities promoting smart tourism can more easily exploit the potential offered by LTM, while cities that invest in LTMs will be in position to develop smart tourism strategies.
- The corresponding TM platforms in these application areas will extend the socio-economic impact of LTMs, strengthening links between academics, tourism professionals, city authorities, policy makers and planers.

c) Proposed methodologies

The Time Machine Requests for Comments will be the main methodological process for setting out specifications for the frameworks, rules, core architectural choices and corresponding research actions to meet the Time Machine objectives. These will be related to:

- Scientific, technological and business studies required for their preparation.
- Research and innovation actions through open calls of proposals to develop solutions meeting the specifications.

At this stage of the LSRI design, particular attention should be paid to interdependencies across pillars and thematic areas, by clearly specifying the links between expected advances in a pillar/thematic area and actions in other pillars/thematic areas that will use these advances.

d) Time plan - Milestones

The Time Machine initiative is planned to be implemented along the four phases that have been broadly defined in section 2.2d.

Initial time-plans for the different milestones have been defined for each pillar (deliverables D2.1, D3.1 and D4.1). These have to be further fine-tuned taking into account the phased structure of the LSRI, as well as the interdependencies across pillars and thematic areas discussed above.

These time-plans have been developed based on intermediate results that serve as milestones, assessing the progress in achieving the overall objectives of each pillar. Attention should also be paid to developments outside the Time Machine, calling for continuous updates of the state-of-the-art during the implementation of the initiative.

e) Key performance indicators

The WGs have developed a set of metrics corresponding to the different characteristics of the pillars / thematic areas. For the final roadmaps a common system will be developed, adapted to the different levels of effects that are intended to be produced, distinguishing:

- Scientific and technological outputs and outcomes
- Innovation outputs and outcomes
- Level of outreach (dissemination and use by the targeted groups)
- Socioeconomic benefits and impact.

The programme has a modular design covering a 10-year period, so it is crucial to ensure coordination of work in the different areas, as well as close monitoring of opportunities to develop synergetic affects. The final set of metrics will, therefore, have to include KPIs for the efficiency and effectiveness of the overall management of the initiative.

3.3 Funding sources

Quality Criterion: Realistic options, including alternatives, are identified for the funding resources of the proposed R&I plans.

Possible sources of funding have been identified from the examination of the financial instruments of the next programming period (2021-27). The analysis will be followed by a more detailed study of the opportunities offered in each case. In parallel, having defined the technical objectives and broad methodological steps, the working groups will look into the required financial envelopes for each pillar. The results of this work will enable the CSA project to enter into more concrete discussions with potential funders from the public and private sector, with regards to the amount and form of support to be committed.

3.4 Stakeholders to be involved

Quality Criterion: The key actors that will validate, contribute and/or support the proposed actions have been identified and are informed/involved in the elaboration of the Research & Innovation plans.

Time Machine is designed to put in place the required large-scale research coordination in all relevant fields, creating the conditions for simultaneous progress to be made in the different fronts, by bringing together key stakeholders from academia, archives, museums, and businesses, and creating partnerships within and between Member States and European bodies involved with cultural heritage.

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).

In the first six months, the CSA managed to more than double the number of its supporting institutions, reaching 449 from the 213 that were present when the CSA proposal was submitted. A strong representation of these institutions has been achieved in the pillar WGs, whereas all of them

have been invited to participate in the web-based consultation that is currently active for receiving comments / input on the draft pillar roadmaps. At the same time, positive contacts have been established with national bodies in charge of financing research and innovation actions.

In the second half of the project, communication actions will intensify to reach the targets set in terms of institutions and funding organisations.

4 Forward planning

The time plan for the TM CSA is shown in Figure 4-1.

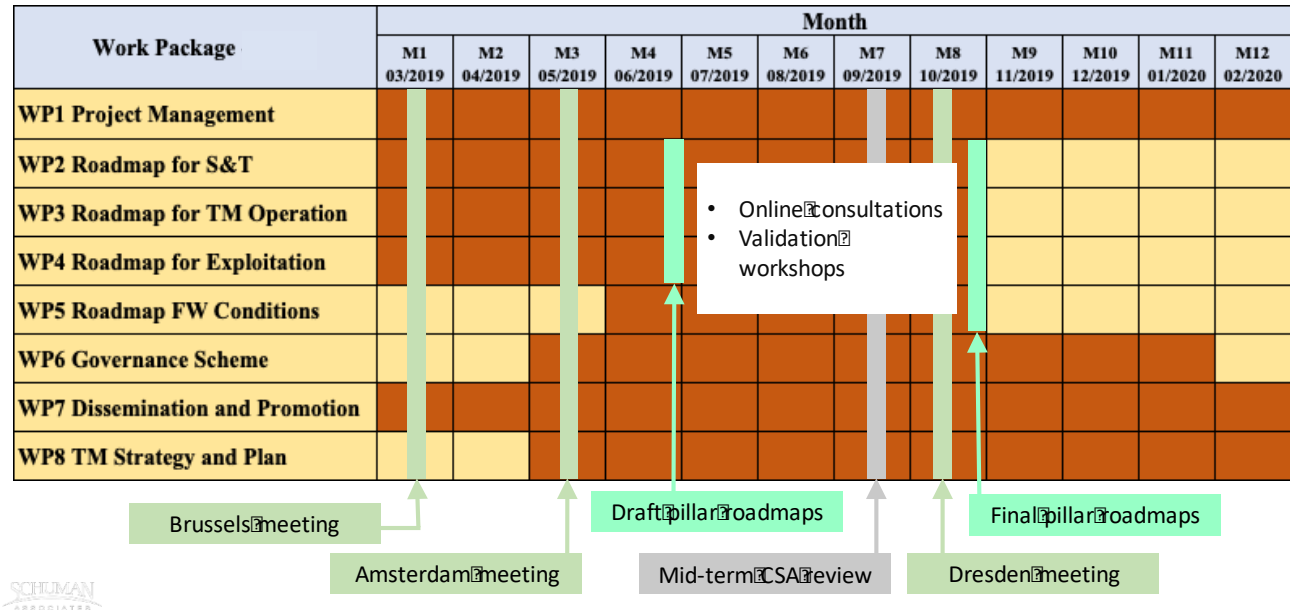

Figure 4-1: TM CSA Time plan – reporting period: Months 4-6.

The reporting period was characterised by the following events:

- The draft roadmaps for pillars 1-3 have been submitted at the end of Month 4 (June 2019).
- The work in Pillar 4 has started, by defining a detailed scope of work, forming the corresponding working group, assigning responsibilities and setting out deadlines for the different contributions.
- An online consultation targeting a wide group of identified existing and potential stakeholders was launched and will remain open in the period July – September 2019. The objective is to collect feedback – assessment and comments – on the proposed roadmaps in pillars 1 to 3.

The wok in the next three months will be organised around the following targets:

- Feedback will be received by the European Commission in September 2019 during the CSA mid-term assessment.
- The second annual Time Machine Conference is planned for 9 to 11 October 2019 in Dresden with an expected attendance of about 200 participants.
- Validation workshops will be organised in this conference, with the participation of external stakeholders expected to be major actors for the implementation of each TM pillar.
- The working groups will use conclusions of the on-line consultation, validation workshops and mid-term assessment to prepare the final pillar roadmaps.
- The final pillar roadmaps will be submitted by the end of Month 8 (October 2019).

Annex A: The Time Machine roadmap methodology

Overview of Time Machine

Time Machine is designed as a LSRI, understood to be a structured cooperation of an already broad network of stakeholders involved with developing/drawing value from the big data of the past, and, therefore, aiming to create substantial socio-economic impact, by implementing a strategic research and innovation agenda. The initiative is articulated around four pillars shown in Figure A-1, together with their corresponding thematic areas.

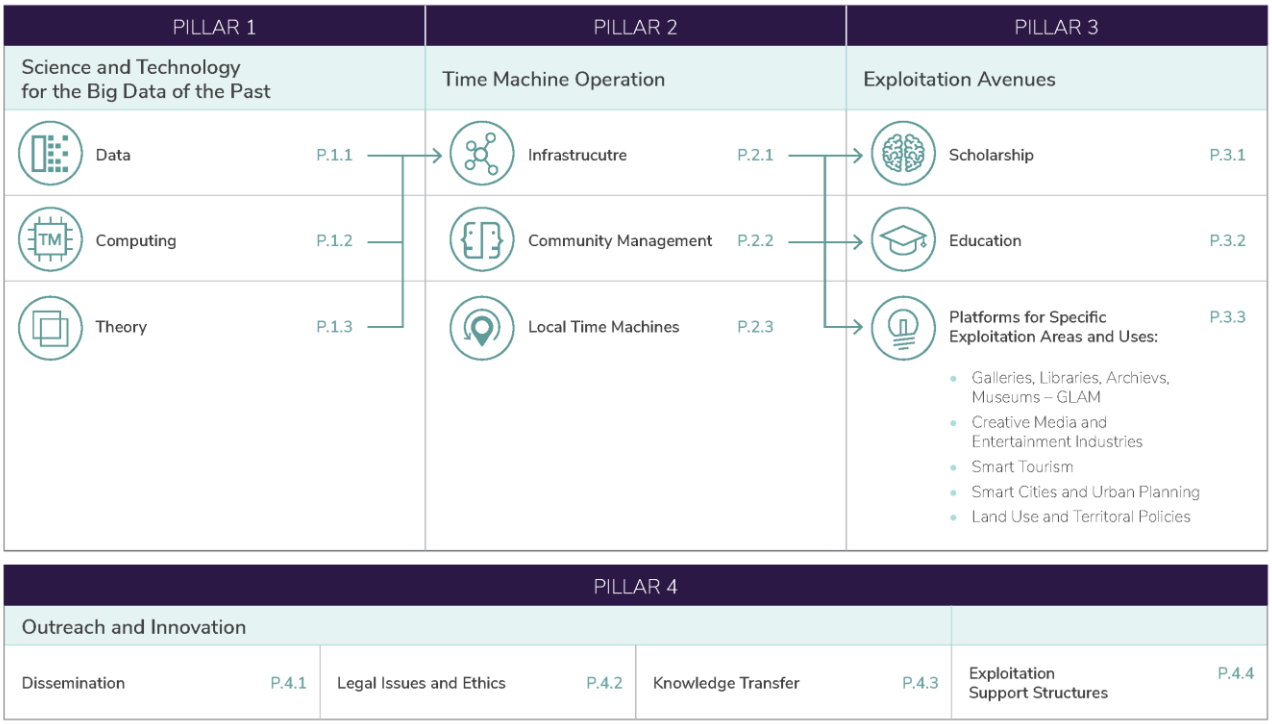

Figure A-1: The TM Pillars and Thematic Areas

Developing the roadmaps

Time Machine is based on a common understanding of the:

- Expected achievements in terms of science and technology that are ambitious, realistic and measurable
- Specific timeframes, where concrete intermediate results will be produced
- Substantial qualitative and quantitative effects for key existing and developing economic sectors
- Requirement to continuously interact with policy/decision makers in order to optimise the implementation process
- Necessity to combine different sources of funding, creating synergies across and promoting efficient utilisation of different funding instruments
- Need for an integrated and holistic approach to maximise impact.

The Time Machine roadmap is the reference document that consolidates and describes this understanding and, therefore, providing the guidelines for all actors involved in the implementation of the TM LSRI. It should, therefore, outline the strategy, the organisation of work to achieve stated goals and outcomes over specified time frames, necessary resources and criteria of success. This reference document will be produced following the TM structure in the pillars and their corresponding thematic areas shown in Figure A-1

In the CSA project, the development of roadmaps is organised in work packages (WPs). The pillar roadmaps are elaborated from detailed plans at the level of each thematic area, undertaken by

working groups (WGs) composed of subject matter experts from the TM consortium. Work in pillars 1-3 (WPs 2-4) has started in month 1 and will be concluded by Month 8 (October 2019), after the integration of feedback received from consultations with external stakeholders. Work for Pillar 4 (WP5) has started in Month 5 (July 2019), building on specific needs identified in pillars 1-3 and finish by Month 8 as well.

The starting point in each WP is a background document for the pillar describing the objectives, the current situation, and the priorities to be addressed. During roadmap development, the WGs describe and assess the state of the art, define priorities and examine alternative development routes for each thematic area, through internal workshops, own expert judgements, consultations with external experts as required, and document / data analysis. Their conclusions, as initial (Month 4) and, then, final (Month 8) drafts, are checked for coherence and compliance with the TM objectives, and are integrated into the pillar roadmap.

During this stage, the roadmaps for the pillars 1-3 also identify and describe:

- The funding sources, as well as corresponding mechanisms and processes to be followed in the contractual relations with the different funders;
- The large variety of stakeholders to be involved, their corresponding roles and, therefore, the needs in terms of management and coordination for the programme;
- The framework conditions relating to policy, legal aspects and ethics that have to be taken into account;
- The approaches and measures that address any barriers to market entry and/or facilitate the commercial exploitation of research results.

These aspects form the basis for the second stage of the design: the first two are used to shape a robust governance scheme (WP 6), while the last two will enable us to design the accompanying actions (WP5 - Pillar 4) supporting the dissemination and outreach of TM in the EU and internationally, as well as creating the enabling environment to maximise the societal and economic outcomes of Pillar 3.

Coordination also involves using the main conclusions from the pillar roadmap drafts to organise:

- Stakeholder workshops with selected representatives of academia, business and policy making.
- Stakeholder online consultations.

These actions aim to receive comments, integrate views, build consensus and ensure commitment to TM objectives and endorsement from a wide range of TM stakeholders. This is an important interaction of the WPs that develop the Pillar roadmaps and the horizontal WP7 that deals with dissemination and promotion:

- The Pillar roadmaps define the stakeholders to be involved (actors and funders).
- WP7 defines the overall dissemination strategy, part of which is to approach the stakeholders identified in WPs 2-5 and undertakes the promotion actions to raise interest and incite participation in TM.
- So, among other things, WP7 creates the favourable conditions for all stakeholders to participate in the road-mapping events organised in WPs 2-5.

In the final stage of the CSA, the findings of the previous stages are put together in the TM reference document presenting the TM strategic objectives, along with detailed methodologies and required resources, as well as the management and operational scheme that will drive the programme through its 10-year span.

Annex B: The quality criteria for the Time Machine roadmap

Taking into account the nature of the LSRI and design process (Annex A), the development of the pillar roadmaps is based on the quality criteria presented below.

Roadmap sections	Quality criteria
1. Pillar objective	A clear mission statement is developed for the pillar, describing how the pillar contributes to the Time Machine vision
2. Research and Innovation plans	A master plan is developed describing the baseline, the expected ambitious outcomes and the methods to achieve them
a. State of the art	A thorough review of relevant scientific, technological and business fields constitutes the pillar's baseline
b. Targeted achievements	Concrete targets are defined for addressing the relevant scientific, technological and innovation challenges
c. Proposed methodologies	The path to the targeted achievements is elaborated to a sufficient level of detail - interdependencies with other pillars are identified
d. Milestones	Critical points to evaluate progress in the proposed developments are determined
e. Key performance indicators	A set of metrics is developed to assess progress made in the different pillars / thematic areas
3. Funding sources	Realistic options, including alternatives, are identified for the funding resources of the proposed R&I plans
4. Stakeholders to be involved	The key actors that will validate, contribute and/or support the proposed actions have been identified and are informed/involved in the elaboration of the R&I plans
5. Framework conditions	Policy measures, legal issues, broader socio-economic developments that may have a positive or negative effect on the proposed R&I plans are identified and discussed
6. Risks & barriers and ways to address them	Risk factors and barriers for obtaining the targeted achievements are identified and mitigation strategies are formulated