



SoBigData RI

SoBigData RI: European Integrated Infrastructure for Social Mining and Big Data Analytics

Website

<http://www.sobigdata.eu/>

Headquarters

Consiglio Nazionale delle Ricerche - CNR -
Istituto di Scienza e Tecnologie dell'Informazione
(ISTI) "A. Faedo", an Institute of the National
Research Council of Italy
Area della Ricerca CNR di Pisa, via G. Moruzzi 1,
56124 Pisa, Italy

Legal Status

None

Type

distributed

Access

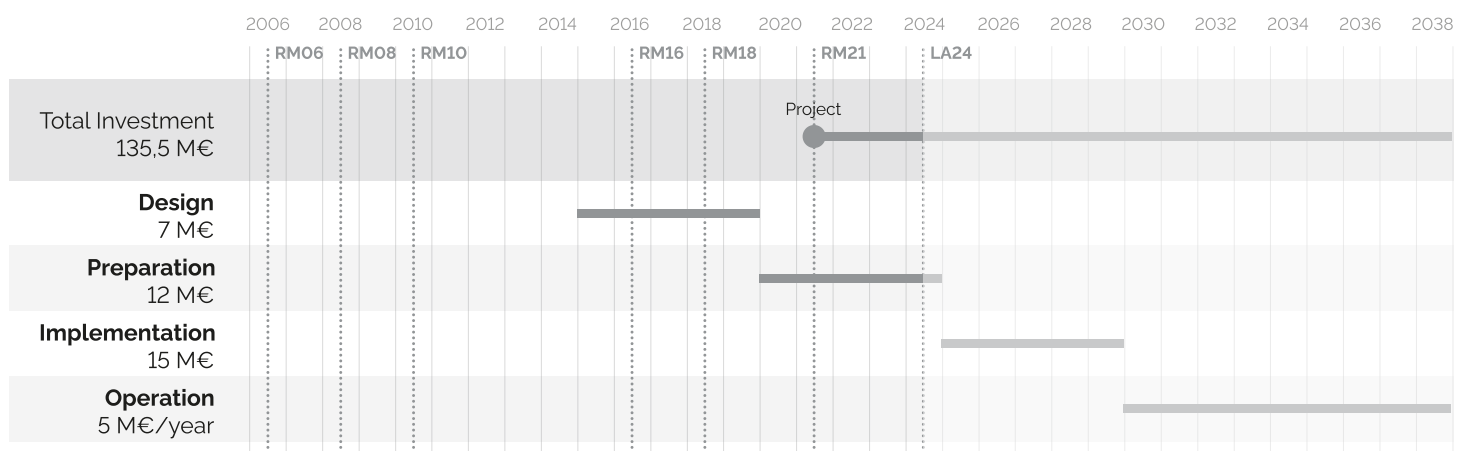
virtual

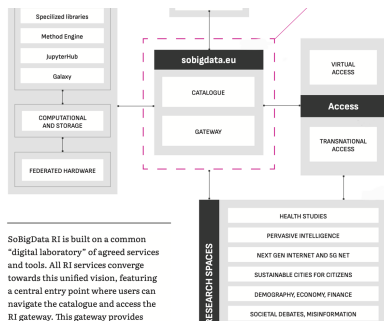
DESCRIPTION

SoBigData is a distributed, Pan-European, multi-disciplinary research infrastructure aimed at using social mining and big data to understand the complexity of our contemporary, globally interconnected society. The SoBigData RI leverages big data and artificial intelligence approaches to understand the complexity of society. The RI aims to create a multidisciplinary scientific community that adheres to the EU's ethical, legal, and open science vision by providing no-profit services to researchers, industry, public bodies, and citizens. Through its goals and collaborations, SoBigData strives to become the reference infrastructure in Europe for big data analysis and social mining, supporting data scientists, AI researchers, enterprises, and policymakers in their data-driven tasks. SoBigData RI is built on a common "digital laboratory" of agreed services and tools. All RI services converge towards this unified vision, featuring a central entry point where users can navigate the catalogue and access the RI gateway. This gateway provides researchers and practitioners with a collaborative environment, promoting responsible open science practices. Under these premises, SoBigData RI can be seen as a resource for sharing datasets, methods, research skills, experiments, and computational resources for supporting the comprehension of social phenomena through the lens of big data. The RI's uniqueness lies in its ability to make the effort of data science and artificial

intelligence scientific communities on a unique synergic work. The RI's services empower researchers to design and execute large-scale social mining experiments. SoBigData renders social mining experiments more efficiently designed and repeatable by leveraging concrete tools that operationalize ethics, incorporating values and norms for privacy, fairness, transparency, and pluralism also by pushing the FAIR (findable, accessible, Interoperable, responsible) and FACT (Fair, Accountable, Confidential and Transparent) principles. Moreover, SoBigData facilitates the synergy between research and industry by bringing together researchers from various disciplines and industry experts; new ideas can be generated, leading to innovative solutions and products. This collaboration between research and industry drives economic growth by creating new opportunities and markets. Under this view, SoBigData offers various services to support businesses in harnessing the power of AI, big data, and cloud computing. Our consulting services provide expertise in AI and big data analysis methods, while our storage and computational services ensure secure and efficient data management. We also offer evaluation services to address ethical and legal risks, as well as partnership opportunities for project development. Additionally, our staff training and education programs and digital integration services help companies stay at the forefront of technology.

TIMELINE & ESTIMATED COSTS





POLITICAL SUPPORT

Lead
IT
Prospective member
BG, CH, EE



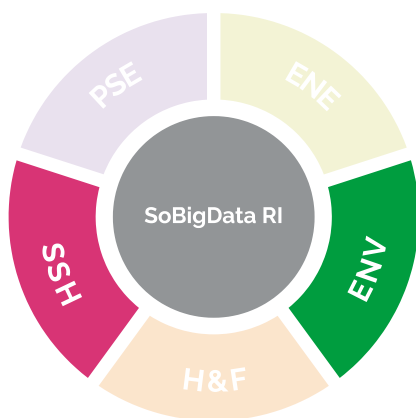
IMPACTS

Design: the first five years (2015-20) of the H2020 SoBigData aimed at consolidating the design phase and defining the functionalities of the e-infrastructure underlying the RI's services. Preparation: This phase ends in October 2025. The main objectives are to define operative strategies for: i) Modeling and definition of the ERIC legal entity, then acquiring legal status. The definition and design of all the aspects related to the definition of a European Research Infrastructure Consortium (ERIC). ii) Preparing the financial, legal aspects (for both central hub and national nodes). The RI plans to develop strategies for establishing partner agreements, to develop an effective and durable governance structure for both central and national hubs, and to define the involvement of the Member States and Associated Countries in the management structure (including the Observer states). iii) Producing and reviewing a Business Plan for long-term sustainability. Designing and engineering a formal business plan that describes the nature of our core business (related to RI services), background information on the organization, the RI financial projections, and all the strategies we intend to implement to achieve the stated targets. iv) Engineering, planning, and optimization of technical infrastructure. The organizational/technical challenge is delivering state-of-the-art dynamic digital assets to remote sites without expensive on-site expertise. v) Defining strategies for service design, community involvement, and partnerships with third parties. In this context, the objective is to develop a sustainable data and method integration strategy that enables the discovery and use of heterogeneous services. For this reason, we will identify analysis and plans to identify and involve all stakeholders in the technical integration work of the RI. We will develop specific communication strategies to involve new stakeholders and disseminate and advise our service beyond our reference community and stakeholders. Implementation: The main aim of this phase (2025-30) is to implement all management and legal structures expected, implement the cost book, and define and model the accesses and services related to the RI.

SERVICES

The SoBigData services are tailored to create interconnections between users and its distributed network of researchers, offering online and offline tools and opportunities. Visit our main page for more details of SoBigData activities and events: www.sobigdata.eu. Social Mining and AI Resources: Access a set of resources, including datasets, AI algorithms, and technologies for social mining. Users can access resources from the RI catalogue or request tailored solutions for specific problems. Visit: <https://sobigdata.d4science.org/>. Infrastructure as a Service: Collaborate in controlled virtual research environments (VREs) with tools like Dedicated Catalogues, Cloud Computing Platforms, and Interactive programming interfaces. VREs can be private or public based on user needs. Requests: info@sobigdata.eu. SoBigData Academy: Enhance skills in social big data and AI through self-learning courses linked to user accounts. Explore: SoBigData Academy. Requests: info@sobigdata.eu. In case of request for tailor-made training courses or masters for companies, policymakers or research institutions, request them at info@sobigdata.eu. Community Building and Networking: Access a network of institutions, public bodies, and industries for collaboration opportunities. SoBigData's mobility program supports users financially when visiting nodes across Europe. Explore: <http://sobigdata.eu/transnational-access>. Ethical and Legal Services: Evaluate ethical and legal risks for experiments and projects based on European principles (ELSEC). Receive tailored guidance and practical suggestions. Requests: info@sobigdata.eu. Get also our periodical magazine at <http://sobigdata.eu/magazine>.

INTERCONNECTIONS



COOPERATION WITH OTHER RIS

SoBigData RI is an EOSC service provider enabling the creation of synergies with other RIs. The RI cooperates with the following RIs: CLARIN: SoBigData will provide basic Natural Language Processing methods and services. EHRI will collaborate with SoBigData to develop a use case related to avoiding arrest among European Jewish communities during the Second World War. ODISSEI: in this case, the collaboration combines AI tools, high-performance computing, and the latest methodologies among the two RIs. RESILIENCE (Religious Studies Infrastructure: Tools, Innovation, Experts, Connections and Centers): is a unique, interdisciplinary scientific RI for all Religious Studies. SLICES supports large-scale, experimental research on parallel and distributed computing and cloud and edge-based architectures and services. OpenAIRE aims to promote open science and offers diverse public services supported by a network of experts from national organizations across European countries. EGI comprises national and intergovernmental computing and data centers. These federated centers make EGI one of the largest distributed computing infrastructures. The RI is involved in GreenDIGIT. The project brings 4 major RIs: EGI, SLICES, SoBigData, EBRAINS, to tackle the challenge of environmental impact reduction.