

Facts

Project No: 665940

Program: H2020 | CSA | GARRI-3-2014

Duration: 09/2015 - 08/2017

Is European research interested in TDM?

Yes it is. This is manifested both by the number of scientific articles published in journals and conferences by European scholars and the number of TDM related projects and infrastructures (in fact projects which have resulted in the development of infrastructures). These two axes are considered as reflecting the strategic decisions of the EU and the policies adopted as regards the selection of specific research domains for funding, as well as the scientific trends attested in the publications.

How long has Europe been interested in TDM?

Europe has been interested in TDM since the first appearance of the term Data Mining, back to the early nineties. It was then, when the EU started funding and conducting research on TDM within the framework of **FP3 (1990-1994)**. This interest has grown over the years as indicated by the continuously increasing number of projects and infrastructures related to TDM in the subsequent frameworks. Specifically, within the **FP7 (2007-2013)** and **Horizon 2020 frameworks (2014-2020)**, **910 out of the 30456 projects focus on TDM**, an equivalent of 3% of the total number of projects.

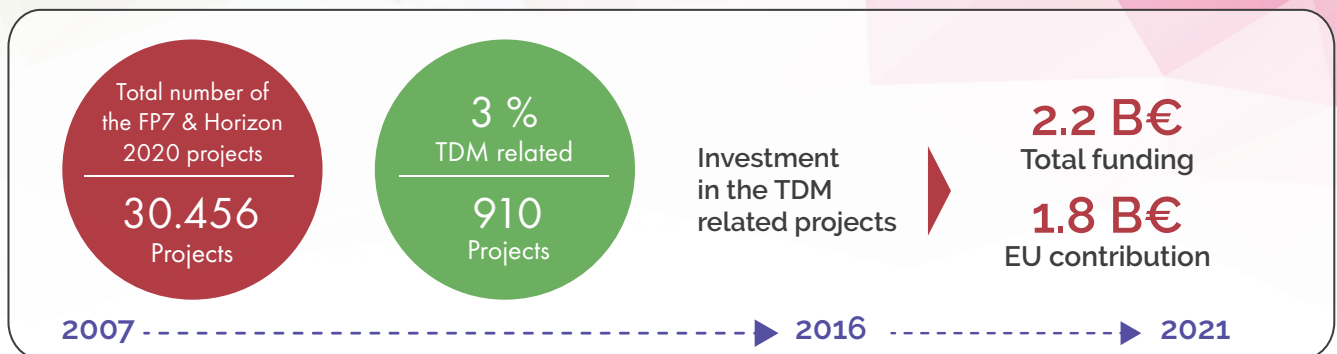


In which way is a project or infrastructure defined as TDM related?

All projects and infrastructures referring to the exact terms Text and Data Mining, Data Mining and Text Mining as well as alternative terms such as Big Data, Data Analysis, Machine Learning etc. are perceived as focusing on TDM.

What is the objective of a TDM project or infrastructure?

Their objective is to provide the tools and technologies for TDM or/and to create big collections of data. These two dimensions are viewed independently or, more frequently, in combination, since tools are developed to process specific types of data.



What is the size of the investment made by the EU in TDM projects and infrastructures?

Projects related to TDM have received a total funding of **2.2 B€** with an EU contribution of **1.8 B€**. Infrastructures, on the other hand, have received 4% of the total project investment by the EU, fact which indicates the strategic priority that the EU has given to them. Specifically, the total investment made in infrastructures within the frameworks of FP7 and Horizon 2020 amounts to approximately **2,8 B€**, of which the EU funding rises up to **2 B€**.

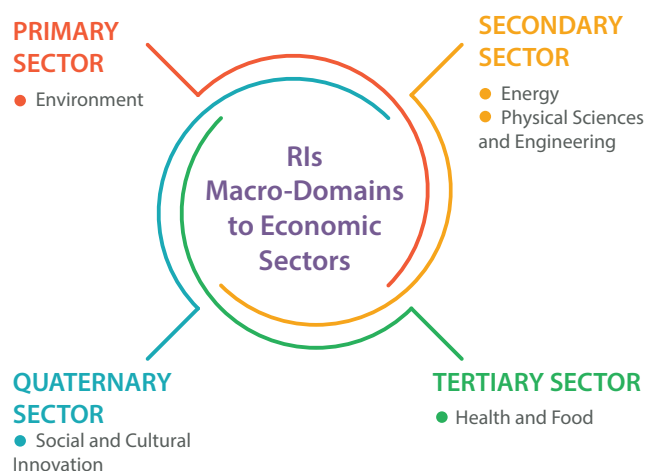
Which are the economic sectors targeted by the TDM projects and infrastructures?

The TDM projects target all the economic sectors. The **tertiary sector**, covering health care, social care and IT services amongst others, has attracted most attention and investment, with **313 projects** totalling an EU funding of **852,4 M€**. The quaternary sector, focusing on research, development and education, follows with **273 projects** which were granted **319,9 M€** from EU (total cost: **402 M€**). Note that the investment made in the secondary sector, comprising all types of industry and energy, for almost the same number of projects (253) is double the investment made in the quaternary sector: **806,4 M€**, of which **633,8 M€** correspond to EU funding. The infrastructures cover five research macro domains pertaining to the respective economic sectors: Environment (primary sector), Energy, Physical Sciences & Engineering (secondary sector), Health & Food (tertiary sector) and Social & Cultural Innovation (quaternary sector).

The quinary sector (Funding and High-level decision) is not independently addressed, given that high level decision making is applied to all macro-domains and, consequently, the corresponding infrastructures have been subsumed in these.

Is TDM research affected by legal limitations or regulations?

Despite the obvious fact that data and tools have more value when shared, used, re-used and re-purposed, openness although needed is far from being established. The challenge of securing the benefits of data sharing **without compromising the rights of all parties involved** (legal, ethical/personal, financial) needs an urgent answer. In parallel, **personal, private or sensitive data** need to be highly respected; any such data (e.g. medical records, trial participations in new treatment protocols, etc.) should be made available after processing through careful anonymization processes.



Mapping of RIs macro-domains to economic sectors

Discover more
VISIT OUR COLLECTION

STORIES 

PROJECTS 

ORGANISATIONS 

TOOLS 

CHALLENGES 