

## **Building Together an Open Science Monitor**

**Authors:** David Osimo, Jon Switters (Lisbon Council); Laia Pujol, Jonathan Wareham (ESADE); Stephane Berghmans (Elsevier); Rodrigo Costas, Ingeborg Mejer, Ziga Turk, Thed Van Leeuwen (CWTS)

The **Open Science Monitor** ([www.ec.europa.eu/info/open-science-monitor/](http://www.ec.europa.eu/info/open-science-monitor/)) was initiated by the European Commission in 2013 with the aim of monitoring and analysing open science trends worldwide in a collaborative and open way.

The indicators cover **Open Access**, **Open Research Data** and **Open Collaboration**, through a mix of bibliometrics, data mining and surveys, using the widest range of data sources. The methodology was updated in 2018 following an open consultation which resulted in over 300 comments. The debate was particularly heated about the usage of proprietary versus open data, and single versus multiple data sources, in monitoring open science. The main modifications to the methodology included the integration of Unpaywall data alongside Scopus, the double checking of results using different sources (e.g. WoS) and the Scopus data is now accessible to researchers in order for them to replicate the methodology. In some cases, proprietary data are still necessary (e.g. Scopus) because of the accuracy and richness of metadata. Hopefully, once the European Open Science Cloud is fully deployed, indicators will be automatically produced from open data sources. Besides the revised methodology, one outcome of the consultation was the creation of a permanent forum of collaboration through a LinkedIn group and an Advisory Board.

The paper presents the methodology used by the monitor, the overview of indicators and an in-depth presentation of the result of a dedicated survey of researchers about practices of research data sharing. The analysis reveals very different status of data availability across different trends. Data are fundamentally robust and available concerning open access, but there are major data gaps concerning open research data and open collaboration. Finally, the lessons learnt from large scale, community collaboration about defining the indicators will be presented.

The topic is of a particular relevance for the Open Science community, and participants of Open Science conference as it touches upon the core of how to track progress on opening up science. Better data are urgently needed for making open science a reality. While open mandates are increasingly adopted by governments, funding bodies and journals, these mandates are difficult to implement. One of the main reasons for this implementation gap is the data gap: for instance, when it comes to research data sharing, we do not have anything close to the quality of bibliometric data on open access to publication because there are no widely adopted standards for data citation. While there are pioneering initiatives that potentially could solve this problem, such as the FAIR data metrics, OpenAIRE and the European Open Science Cloud, they still have to scale and their results will arrive several years from now.

The project aims to share the results and lessons learnt, and gather further feedback on how to improve the indicators, the data sources and the collaboration process itself.

**(490 words)**