

## Societal Challenges

On 20 April 2020, the European Commission together with several partners launched a [European COVID-19 Data Platform](#) to enable the rapid collection and sharing of available research data. Populating this platform with data from the Social Sciences and Humanities aims

to pool and augment behavioural and attitudinal data to enable researchers to investigate the opinions and attitudes of European populations during the crisis, and monitor the occurrence and spread of the virus.

## Technical Challenges

Data variety presents a challenge to the alignment of catalogues, metadata, and protocols with the life sciences and other parts of the platform. Collecting and combining quantitative and qualitative data in many formats (text, audio, video, social media) and in

multiple languages requires multilingual thesauri and ontologies to make data findable and comparable;

Ensuring the security and protection of sensitive data collection and analysis.

## How EOSC can help and add value

*Adding Social Sciences & Humanities data to the COVID-19 Data Platform, providing contextual data and a knowledge development environment*

The fast-track element of the task involves the creation of a catalogue of key datasets for social, economic, psychological analyses, organised by data producer and national service provider in data hubs to replicate the EMBL structure, and supplemented with contextual economic, social, cultural, health, and migration data. The catalogue is supported with background infrastructures such as multilingual thesauri, controlled vocabularies, metadata profiles based on global standards, and data sensitivity tags.

Specific challenges for the social data are their multilinguality and large variety of data types. Qualitative data have high information value, but require specific techniques for annotations, analysis and extraction

of information. Tools and expertise within the [SSHOC](#) project will be used to deal with qualitative and multimedia data and to address multilinguality.

A knowledge development track will make use of online surveys and AI techniques, focusing on the knowledge cycle and data interoperability, including non-hierarchical data, via semantic techniques such as Knowledge Graphs. To optimise reusability, concerted and collaborative actions to group and enrich data are envisioned in cooperation with research communities. Bringing relevant data together will increase efficiency, improve comparability and provide all researchers with the same opportunities for using the data. To facilitate cooperation and to provide seamless access even to sensitive data, secured environments will be set up for bringing data and researchers together.

