

CESSDA's Stepping Stones in Realising the FAIR principles

The Consortium of European Social Science Data Archives (CESSDA) is a distributed social science data infrastructure with currently 15 members (www.cessda.eu). We work on FAIR data and are developing a social science data platform as part of the EOSC. CESSDA is an ESFRI Landmark and in June 2017 it became a European Research Infrastructure Consortium (ERIC).

CESSDA's goal is to bring together the expertise of the CESSDA Service Providers and realise an infrastructure that enables researchers to perform high quality research. It has a small Main Office in Bergen and each Member assigned a Service Provider that functions as a national social data archive and data services organisation. Via CESSDA these Service Providers join resources and expertise to realise this top data infrastructure for social science data, and the national Members discuss and create an overarching CESSDA strategy.

A key challenge is to realise the EOSC – using the FAIR principles, taking into account the distributive infrastructure within social sciences, and the sensitivity of many social science data – requiring safe & secure access.

At CESSDA we follow a stepping stone strategy in realising the social science data cloud using the FAIR principles. We have realised the 'F' and have pathfinder projects on the other principles.

First, we will present an important milestone: the joint CESSDA catalogue to find and search the data. With this catalogue, data have become Findable. We will explain the steps that were needed to realise this catalogue (currently in beta version; operational in December 2017 with 30,000 studies): the technology backbone, connecting to the distributed data archives, set standards for the meta data, build the user-interface, etc. By describing these steps and their do's and don'ts, we can share our expertise and practices with other distributed infrastructures.

Next, we will present our stepping stone strategy on how we work out the other FAIR principles: especially legal aspects of Accessibility, development of data clusters to improve Interoperability, and tools to increase the Reuse of data – including reuse of sensitive data. We will also stress the importance of Skills and will elaborate on our Training activities – taking the data life cycle as starting point. This means providing tools for data owners, but also focus on users, e.g. by creating user communities around data clusters, and how to share expertise in a distributed data infrastructure.

Finally, we will explain on the CESSDA approach to coordinate actions and activities: how we set up these pathfinder projects within a distributed infrastructure will serve as a use case.