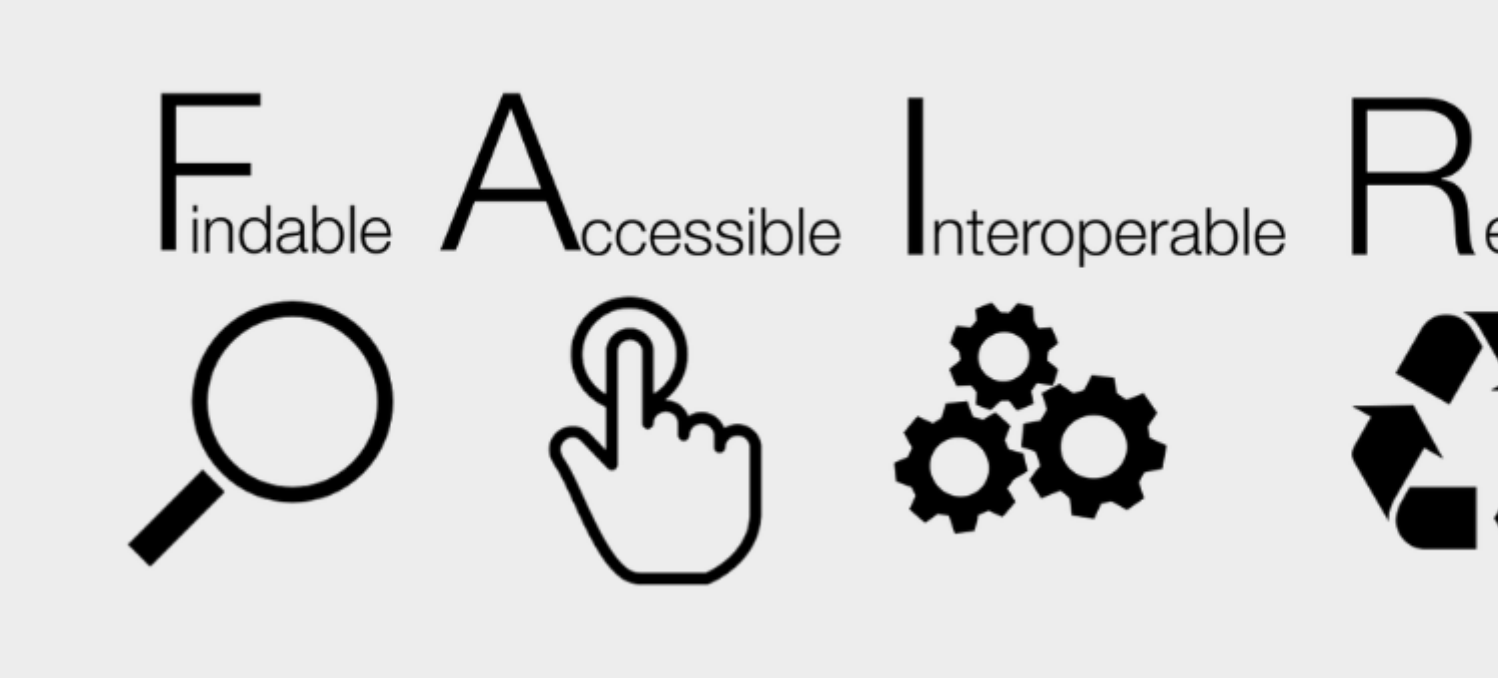
Fair Data Stewardship, Metadata Standards, Semantics, Sensitive Data



We ensure FAIR, AI-ready data for social sciences and humanities, focusing on quality, interoperability, and automation within the CSSH repository at BSC Dataverse.

Summary

This research line focuses on ensuring data is managed and shared according to the FAIR principles—making them **Findable, Accessible, Interoperable, and Reusable**- and Open Science values. The goal is to enhance the quality of data describing the context with substantive knowledge and employing metadata standards to enable interoperability and reuse across different disciplines and data systems. Ultimately, this research line aims to build processes, protocols and tools to ensure AI ready data for the social sciences and humanities.

Additionally, this line addresses the handling of **sensitive data**, to facilitate access to social data for research without compromising utility, while ensuring compliance with legal and ethical requirements.

Currently, we focus on setting up the CSSH data repository within the BSC Dataverse repository and building the necessary tools and processes to manage and share AI ready and FAIR machine and human-actionable social science and humanities data. We aim to automate some aspects of large-scale data curation and harmonization using an AI and computational approach.

Objectives

- Ensure data follows FAIR principles and Open Science values.
- Improve data quality with metadata standards for better reuse.
- Develop tools for AI-ready data in social sciences and humanities.
- Securely manage sensitive data while ensuring compliance.
- Build the CSSH repository within BSC Dataverse for sharing.
- Automate data curation and harmonization with AI.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

---