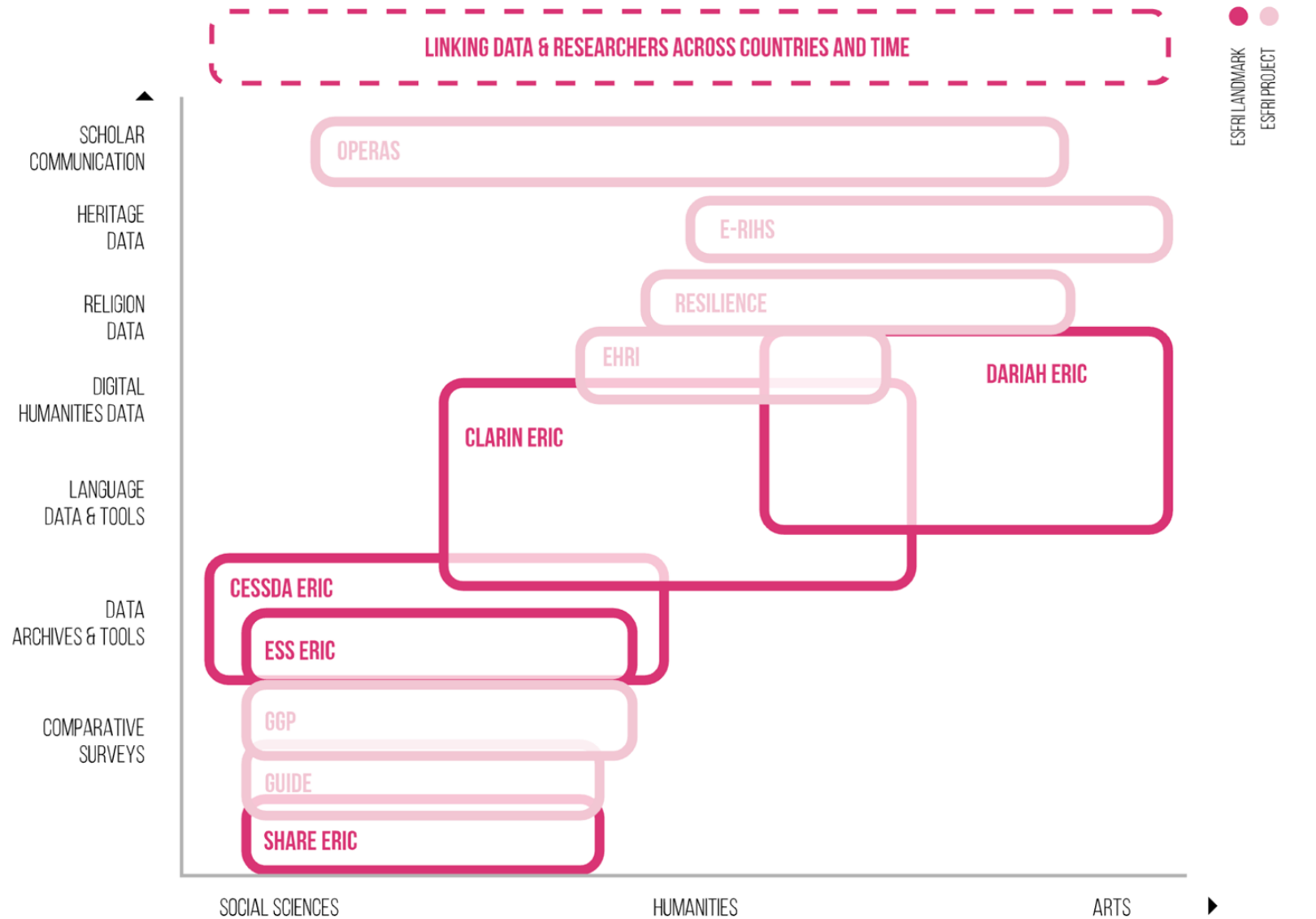# SSHOC and OSCARS 2

Bonnie Wolff-Boenisch (CESSDA) & Darja Fišer (CLARIN)

23 January, 2026

OSCARS

- SSHOC as a consortium was ´too´ successful
  - 121 SSHOC proposal and 24 cross-disciplinary domain

- In certain domains we reached the longer tail of research - but there is a need to nurture disciplines less exposed to existing ERICs, and EOSC in particular.
  1. SSH domain is highly fragmented & heterogeneous
     a. with a lot of different methodological approaches
     b. different data types (videos, images, sounds, interviews, survey data, sensitive data…) + language barriers

     = convergence of data formats, virtual research environments and methodological approaches would be highly beneficial - a set of standards to connect different data sets and further data reuse

     = Science clusters/competence center can play a very important role **as knowledge transfer and one-stop shop.**

Key topics - 2nd Call of OSCARS (2025)

OSCARS

Rank	Category	Count
1	FAIR Data Infrastructure & Services	18
2	Cultural Heritage & Archaeology	22
3	AI, NLP & Text Analytics	19
4	Linguistic & Literary Corpora	16
5	Oral History & Memory	11
6	Survey & Social Methods	12
7	Historical & Economic Research	9
8	Education & Learning	8
9	Policy & Social Impact	6
10	Interdisciplinary & Emerging	10
	TOTAL	**121**

AI Integration Across SSH

~40% of proposals incorporate AI/ML, but show maturity in using AI as enabler rather than end goal. Strong emphasis on explainability and responsible deployment.

FAIR Principles as Foundation

FAIR compliance is expected baseline. Focus shifted from 'making data FAIR' to 'building FAIR-native infrastructure'.

EOSC Integration Priority

60%+ mention explicit EOSC integration. Projects design connectors and EOSC-compliant metadata from outset.

Democratization & Accessibility

Strong theme of lowering barriers. Multiple proposals target non-technical users and design GUI (graphical user interfaces).

Sensitive Data Handling

Growing sophistication in managing sensitive data: anonymization, consent management, federated analysis.

Qualitative Methods & Qualitative research infrastructure

- **Environmental Sciences Dominance:** Largest category (34 proposals) reflects urgent climate/biodiversity research infrastructure needs
- **Health-Environment Nexus:** Multiple proposals bridge health and environmental monitoring, anticipating One Health integration
- **Heritage Preservation Urgency:** High scores for archaeology/heritage (83.7 avg) reflect recognition of digitization urgency
- **Policy and social impact:** Several proposals address underserved socio-economic topics (public health, microdata, administrative data) and regions (Bosnia-Herzegovina, Ukraine)
- **Citizen Science Integration:** Multiple proposals incorporate citizen science as core component, not afterthought

- **Infrastructure Over Tools:** Moving from standalone tools to interconnected ecosystems

- **Workflow-Centric Design:** Complete workflows rather than single-function tools

- **Gaps in the SSH (see next slides):**

In OSCARS 1
- The `**very open´- Call** approach will not be possible for the SSH

In OSCARS 2
- Changes in the mechanics of the calls:
  - Very specific calls
  - One page proposals?
  - With focus on impact and how open science can reach socio-economic impact

- **SSHOC - cross-disciplinary projects**
  - **AI, GenAI - 'AI Slop' the aim is to identify tools, narratives to verify and ensuring trustworthy data**

  - The problems:
    - how can we trust news, social media, data injected into databases (image analysis)?
    - Conventions on how to authenticate e.g., manuscripts or datasets
    - How do we build resilient systems?
    - **Digital preservation and data preservation (including data governance, and sensitive data for digital resilience and sustaining and contributing to reuse;**
    - **research security (cybersecurity)??? - may be already covered**

- ○ **E.g., Social Media Data**
- ○ Data everyone wants to use but everyone is collecting from scratch and in their own ways -> we want encourage projects that builds datasets in a way that enable reuse by lots of different communities
  - ■ TBD if this subject can be covered through a SSHOC cross-disciplinary or a domain specific call (see later)

- Technical
  - ○ EOSC-relevant plugins/extensions for repositories (e.g. for metadata curation)
  - ○ Data harmonisation tools; meta data standards* - part of the call text
  - ○ Sensitive Data - TBD

- Public health or SDGs-related (environmental) domains - SSH plus Life Sciences or plus Environmental Sciences
  - Requirement - **topic however needs to be carved out prior with the other corresponding clusters**

- Domain-specific (organised under subcalls)*

    - **Gaps in the Social Sciences:**
        - Untapped data: microfirm data (aka financial data) and administrative data; qualitative data
            - Strategic Partnerships in CLUSTER 2: Culture, Creativity, Inclusive Societies
            - Social Transformations and Resilience (CESSDA is involved)

    - **Gaps in the Humanities**
        - Hot data (along with corresponding processing tools and methods): spoken, multimodal, images -> we should encourage tackling complex, out of the ordinary datasets
    - SSH
        - **Sensitive but highly impactful data**: medical records, public administration, court records -> we should encourage projects on "problematic" data as they have a big impact potential
        - *Non-standardized data annotation schemas (e.g. error annotation) -> we should encourage projects that use annotation schemas that can be extended to other languages & enable comparative analyses*
        - *Processing tools: most language-specific -> we should encourage paradigms that are*

- Further development of the **Competence Center as knowledge transfer and one-stop shop (for RDM & FAIR & EOSC matters etc)**

  - For SSHOC training on specific tools and data integration
    - convergence of data formats, virtual research environments and methodological approaches would be highly beneficial
    - AI solutions to improve offer (replacing helpdesks with chatbots, semi-automated depositing services, improved findability of resources)

- Training/education - offering webinar trainings for all proponents (including the one not successful) towards SSH ERICs plus EOSC

Beyond OSCARS 2 (2032)
- Designated ressources (on the long-term) from SSH RIS? Or a temporal CC with the specific purpose (as outlined in OSCARS 1 ane 2) and then dissolved or integrated into EOSC