Concept Note on Open Science Infrastructure

Introduction

Open Science is a movement that seeks to make scientific methodologies, research, data, teaching materials and dissemination accessible to all levels of an inquiring society. Its goal is to create a scientific process that is more transparent, inclusive, and collaborative. By promoting the principles of openness and collaboration, open science aims to enhance the transparency, reproducibility, accessibility, and impact of scientific research across the globe.

The foundation of open science lies in its core values of transparency, reproducibility, collaboration, collective benefit, and equity. Transparency ensures that the creation and dissemination of scientific knowledge foster a responsible and trustworthy academic environment. Reproducibility allows scientific work to be verifiable, thus enhancing trust in the research process. Collaboration extends beyond academia to include broader societal engagement, enriching research and education. Open science also emphasizes the collective benefit, ensuring that scientific knowledge serves the greater good and addresses real-world challenges equitably.

The implementation of open science requires robust infrastructure to support open data sharing, transparent research processes, and collaborative networks. This infrastructure must be built on principles that ensure data is findable, accessible, interoperable, and reusable (FAIR). Additionally, it must support the engagement of diverse stakeholders in the research process, fostering inclusivity and innovation.

This initiative aims to establish an open science infrastructure (across Southern Africa) to promote equitable access, transparency, and collaboration in scientific research. For Southern Africa, implementing an open science framework presents an opportunity to integrate local knowledge systems with global scientific communities, leveraging collaboration between partner countries such as Botswana, Zambia, Germany, the Netherlands, and Slovenia. This collaborative effort aims to create a sustainable and inclusive scientific ecosystem that addresses regional challenges while contributing to global scientific advancements.

Commented [1]: Whats our scope? Is the infrastructure to be shared?

Objective

To create a sustainable, FAIR-compliant infrastructure (as well as underpinning policies) supporting open science, focusing on data sharing, open access to publications, and fostering interdisciplinary collaboration. AI capabilities will be incorporated to support data curation, discovery, and metadata management. Existing materials such as those from UNESCO should be reused and, if necessary, brought up to date.

Commented [2]: FAIR (Findable, Accessible, Interoperable, Reusable)

Dimensions of Open Science

The infrastructure will address:

- **Open Access**: Enhance unrestricted online access to research publications and data.
- **Open Data**: Enable reuse of datasets with compliance to FAIR principles.
- **Open Source and Open Hardware**: Promote open-source tools for reproducibility and accessibility.
- **Open Collaboration**: Support online platforms and communication channels to enable broad collaboration.
- Open Education: carrying out education freely available, usually using digital technologies.

Key Components of the Open Science Infrastructure

- **Repository and Data Management Platform**: Develop a repository for publications, datasets, Open educational resources (OER) and other digital content, adhering to FAIR principles and including AI-driven metadata generation.
- **Open Access Publishing Support**: Establish funding for publication fees and partnerships with open-access publishers.
- **Data Stewardship and Metadata Standards**: Implement AI tools to automate metadata tagging and ensure compliance with FAIR and CARE (for indigenous data) standards.
- **AI-enabled Discovery and Analytics**: AI tools to support the integration and analysis of complex datasets while ensuring ethical use.
- **Collaboration and Networking Platforms**: Establish online spaces for interdisciplinary collaboration.
- **Capacity Building and Training**: Provide workshops on open science practices and AI literacy for researchers.

- **Outreach and Community Engagement**: Promote open science benefits through events and community-based projects.

Governance and Policy Framework

- **Policy Development**: Formulate policies for open access, data sharing, IP rights, and research evaluation.
- **Incentive Structures**: Establish metrics for open science impact and incentivize researchers.

Infrastructure Development and Sustainability

Long-term investment is required, with ongoing support for technology, training, and partnerships.

Timeline

A phased approach, including milestones for repository establishment, platform integration, and AI deployment.

Expected Outcomes

- Increased accessibility of research outputs
- Enhanced regional collaboration
- Compliance with FAIR principles, supported by AI-driven metadata management
- Growth in interdisciplinary research outputs

Conclusion

The development of robust open science infrastructure is essential to advancing knowledge, fostering innovation, and enhancing the global competitiveness of research in Southern Africa. For institutions of higher learning, establishing such infrastructure is critical to enabling researchers to share, access, and build upon each other's work openly and efficiently. This infrastructure not only supports transparency, collaboration, and reproducibility in research but also addresses regional challenges of equitable access to scientific resources and expertise.

By committing to the development of open science infrastructure, universities and research institutions will empower their communities, attract international partnerships, and align with global standards for open and inclusive scientific practices. The approval of this concept is a step toward building a sustainable, FAIR-compliant infrastructure that will amplify the impact of scientific research across Southern Africa and beyond.

Infrastructure Components and FAIR Indicators

- **Repository and Data Management**: A centralized repository with FAIR compliance
  - **Indicators**:
    - Percentage of research data with standardized metadata for Findability.
    - Accessibility rate (percentage of data accessible within six months of publication).
    - Data interoperability score based on standardized formats.
    - Reusability: frequency of data re-citations.

- **AI-enabled Data Stewardship**: Use AI to automate metadata tagging and enhance findability and accessibility.
  - **Indicators**:
    - Number of datasets with AI-generated metadata tags.
    - AI-based accuracy in categorizing data fields for interoperability.

- **AI Tools and Applications for Discovery**: AI integration for metadata management, discovery, and content analytics.
  - **Indicators**:
    - Use of AI in metadata accuracy and data classification.
    - User satisfaction rating for AI-driven data search tools.

- **Capacity Building**: Training on FAIR principles, AI usage, and data ethics.

Outreach and Community Engagement

- **Engagement Strategy**: Plan for raising awareness among stakeholders.

Evaluation and Monitoring

- **FAIR Principle Compliance**: Use of FAIR indicators for repository evaluation.
- **AI in Monitoring**: Automated monitoring for data sharing practices and FAIR compliance.