第四届可持续发展大数据国际论坛
The 4th International Forum on Big Data for Sustainable Development Goals

Safeguards for Creating Trusted Research Environment (TRE) on Open Science Infrastructure (OSI) Data-sharing Platforms for SDGs

Prof Joseph Muliaro Wafula PhD, FCCS, FCSK

Lead Scientist AOSP EA Node, Kenya

Chair CODATA, Kenya

Associate Professor JKUAT, Kenya

Open Science Infrastructure

- Open Science Infrastructure plays a pivotal role in accelerating progress towards the SDGs.
  - Breaks down barriers to knowledge and fostering international collaboration,
    - Enhancing Access to Knowledge (SDG 4 - Quality Education)
    - Promoting Innovation (SDG 9 - Industry, Innovation, and Infrastructure)
    - Supporting Health and Well-being (SDG 3 - Good Health and Well-being)
    - Environmental Sustainability (SDG 13 - Climate Action, SDG 14 - Life Below Water, SDG 15 - Life on Land)
    - Reducing Inequalities (SDG 10 - Reduced Inequalities)

Ensures that science contributes effectively to sustainable development, benefiting both local communities and the global population

TREs to facilitate secure and ethical data sharing.

- TREs implement rigorous security measures, including data anonymization, encryption, and controlled access, to safeguard against unauthorized access or breaches.
- TREs enable researchers to access and use data in a controlled, transparent manner that adheres to ethical guidelines.
- TREs ensure that data is used responsibly, respecting participants' rights and mitigating the risk of misuse through enforcing governance, and legal (National/International) frameworks eg General Data Protection Regulation (GDPR) in the EU or the Health Insurance Portability and Accountability Act (HIPAA) in the U.S.
- TREs foster collaboration among researchers by providing secure access to shared datasets, while also ensuring that the integrity and confidentiality of the data are maintained,
- TREs build trust among researchers, data providers, and the public. This trust is crucial for encouraging data sharing and collaboration, which are essential for addressing complex global challenges
- TREs by protecting sensitive data and building trust among stakeholders, they enable collaborative research while safeguarding individual privacy and data integrity

TREs in Data-Sharing Platforms

- TREs allow multiple researchers from different institutions to collaborate securely without the risk of data breaches or unauthorized access
- TREs ensure that sensitive data is handled responsibly, maintaining its integrity and preventing alterations that could affect research outcomes.
- TREs build trust among stakeholders, including the public, research institutions, and data providers
- TREs ensure that data sharing adheres to strict legal and ethical frameworks

Primary Safeguards for Privacy and Data Security

- **Encryption** ensures that data is transformed into an unreadable format while in storage (at rest) and during transmission (in transit).
- **Access Controls and Authentication**
  - **Role-based Access Control (RBAC)**: Users are given access based on their role within an organization
  - **Multi-factor Authentication (MFA)**: Requires users to provide multiple forms of verification (e.g., password and phone authentication)
- **Data Anonymization and Pseudonymization**
  - **Anonymization**: Irreversibly removing personally identifiable information (PII) from datasets to ensure that individuals cannot be identified from the data.
  - **Pseudonymization**: Replacing PII with fictional identifiers (such as random IDs) so that data can still be re-linked to individuals by authorized parties if needed.
- **Audit Trails and Monitoring:**
  - Keeping detailed logs of who accessed the data, when, and for what purpose provides transparency and accountability
- **Data minimization** involves collecting and storing only the data that is necessary for a specific purpose.
- **Regular audits and penetration tests** are used to identify potential vulnerabilities in the system
- **Privacy and Security Policies:** Establishing clear internal policies regarding data use, storage, sharing, and deletion ensures that employees understand their responsibilities regarding data privacy
- **Legal and Regulatory Compliance**

Challenges sharing data across borders with varying legal frameworks

Different Privacy Laws and Regulations

- **Varying Legal Standards**: Different countries have different privacy and data protection laws
- **Data Localization Requirements**: Some countries require that data be stored or processed locally (data sovereignty**)**

Inconsistent Consent Requirements

- Different jurisdictions have different standards for obtaining consent to process and share personal data

Cross-Border Data Transfer Mechanisms

- Organizations need valid transfer mechanisms like Standard Contractual Clauses (SCCs) or Binding Corporate Rules (BCRs) to comply with data protection laws, but these are often complex and difficult to implement.

Data Security Concerns

- Different countries have different security standards, and ensuring compliance with the highest level of security across multiple jurisdictions is a significant challenge

Cultural and Ethical Concerns

- Beyond legal issues, cultural and ethical differences in how data privacy is perceived can pose challenges. For instance, in some countries, data privacy is viewed more as a fundamental human right (e.g., in Europe), whereas in others, the emphasis is on business and innovation

Solutions to sharing data across borders with varying legal frameworks

Adopting Global Frameworks and Standards

- Organizations can adopt global standards such as **ISO/IEC 27001** for information security management or **ISO/IEC 27701** for privacy information management

Data Anonymization and Pseudonymization

- By anonymizing or pseudonymizing data before sharing it across borders, organizations can reduce the legal risks of cross-border data transfers

Use of Standard Contractual Clauses (SCCs) and Binding Corporate Rules (BCRs)

- SCCs and BCRs are legal tools that facilitate the lawful transfer of data between countries with different legal frameworks

Establishing Data Sharing Agreements

- Organizations can develop comprehensive data-sharing agreements that define the legal, security, and privacy terms for cross-border data exchanges

Data Localization Solutions

- Involves storing and processing data locally in compliance with the jurisdiction's laws while sharing non-sensitive or anonymized data globally.

Leveraging Secure Data Transfer Technologies

- Using secure data transfer technologies such as **end-to-end encryption**, **secure VPNs**, and **trusted research environments (TREs)** ensures that data is protected while being transferred across borders

Engaging with Regulators and Legal Counsel

- Proactively working with legal experts and regulators in different jurisdictions ensures that organizations understand and comply with local data protection laws

TREs on OSI platforms enhanceing collaboration between researchers, policymakers, and the public

- **Data Security**: TREs ensure that sensitive research data is securely stored and accessed, allowing researchers from different institutions or countries to collaborate without compromising data privacy. By enforcing stringent data protection measures, such as encryption, anonymization, and role-based access controls, TREs encourage data sharing across boundaries.
- **Policymaker Involvement**: TREs on OSI platforms can provide policymakers with direct access to research findings in a secure environment, allowing them to make data-driven decisions. In turn, policymakers can give feedback or contribute to shaping research questions that address current public policy needs.
- **Public Involvement**: TREs can also support public engagement by allowing access to anonymized data for citizen science projects or making research outcomes transparent to the public
- **Bridging Silos**: OSI platforms that integrate TREs allow researchers from different fields (e.g., health, economics, environmental science) to work together on interdisciplinary challenges. This fosters innovation and comprehensive solutions to complex problems like the **Sustainable Development Goals (SDGs)**.

Impact of TREs on accelerating progress toward the SDGs in Africa

- In Africa, where data availability can be fragmented, TREs can consolidate data across regions and sectors (e.g., health, agriculture, education) to generate actionable insights
- TREs enable international researchers and institutions to collaborate securely, overcoming legal and logistical challenges around cross-border data sharing. In Africa, this can lead to shared expertise on global challenges such as climate change (SDG 13), poverty reduction (SDG 1), and access to clean water (SDG 6).
- In many African countries, concerns about data sovereignty and ethical use of local data are significant. TREs ensure that data from African populations is managed ethically and used to benefit local communities rather than external actors. This is crucial for advancing SDGs in a manner that respects local rights and needs.
- TREs can serve as hubs for training African researchers and policymakers on advanced data analytics, machine learning, and other critical skills necessary to drive SDG progress
- In Africa, where data gaps have traditionally hindered policy development, TREs provide the infrastructure to base policies on empirical evidence.
- Data on rural communities collected within a TRE could inform projects aimed at improving rural infrastructure (SDG 9) and access to quality healthcare (SDG 3).

Conclusion:

TREs are crucial tools for accelerating progress toward the SDGs in Africa by enabling secure data sharing, promoting collaboration, supporting ethical data use, building capacity, and facilitating evidence-based policymaking

Bibliography

1. Collaborative Platforms Open Training Handbook
2. RScienceEGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCI of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679
3. The FAIR Guiding Principles for scientific data management and stewardship https://www.nature.com/articles/sdata201618
4. UNESCO Recommendation on Open Science https://unesdoc.unesco.org/ark:/48223/pf0000379949
5. Trusted Research Environment https://www.dcc.ac.uk/search/content?search_api_views_fulltext=Trusted%20Research%20Environment%20articles
6. Recommendation of the Council on OECD Legal Instruments Responsible Innovation in Neurotechnology file:///C:/Users/mulia/Downloads/OECD-LEGAL-0457-en.pdf
7. Transborder Data Flows and Data Privacy https://academic.oup.com/book/5440?login=false

Thank you!

muliaro@icsit.jkuat.ac.ke