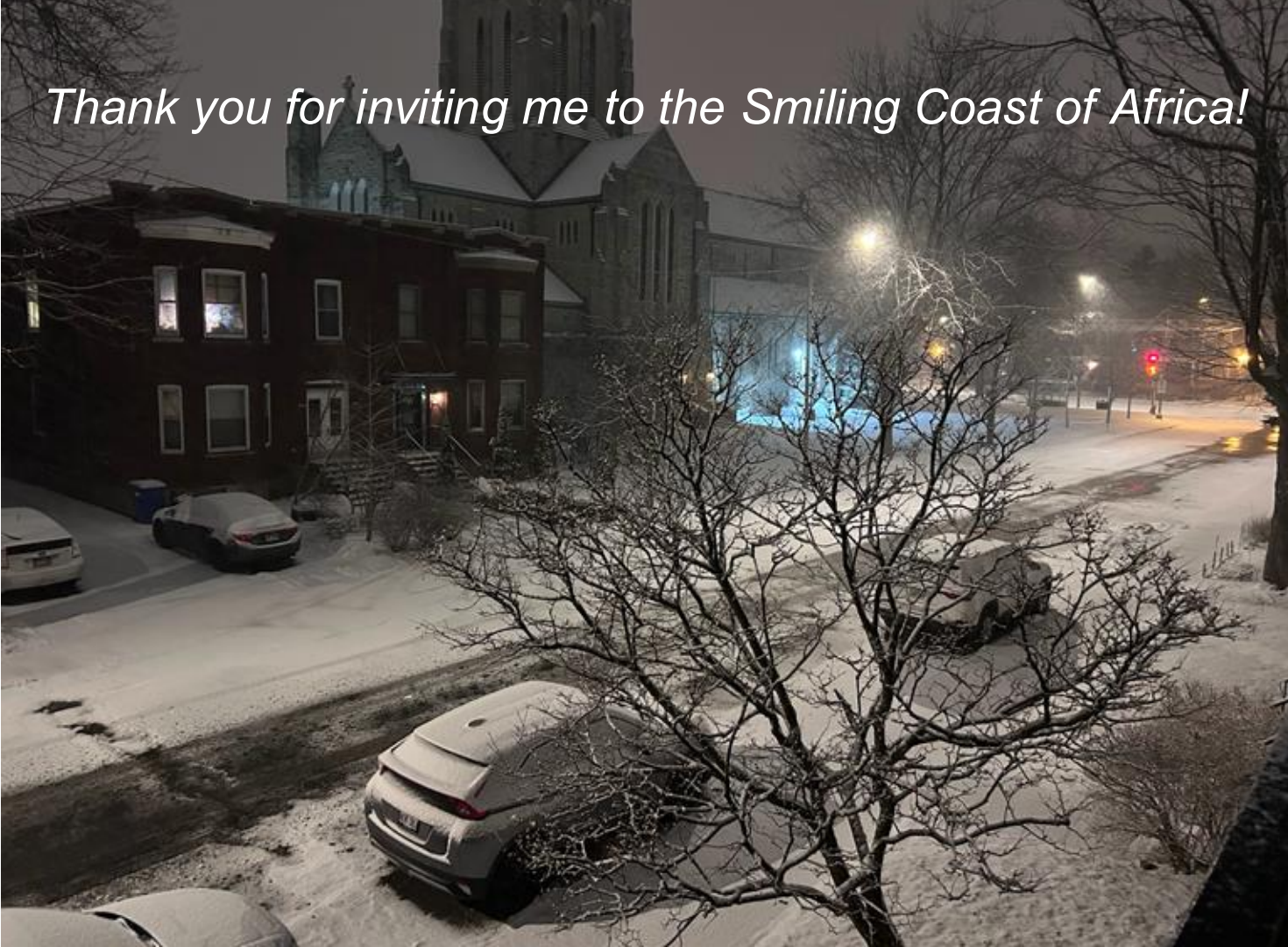


Building Repository Infrastructure for **Next Generation Research**

Kathleen Shearer, Executive Director, [COAR](#)



Thank you for inviting me to the Smiling Coast of Africa!



Over **150** members
and partners from around
the world

About COAR

COAR is an international association that brings together individual repositories and repository networks in order to build capacity, align policies and practices, and act as a global voice for the repository community.

Main areas of work:

- Innovation
- Interoperability
- Visibility
- Capacity building

Times, they are a changin'... REALLY FAST!



Two key factors driving change

1. Artificial intelligence

Science in the age of AI:

How artificial intelligence is changing the nature and method of scientific research

Read the report online

THE
ROYAL
SOCIETY



2. Geopolitical upheaval

Who Owns Africa's Health Data?

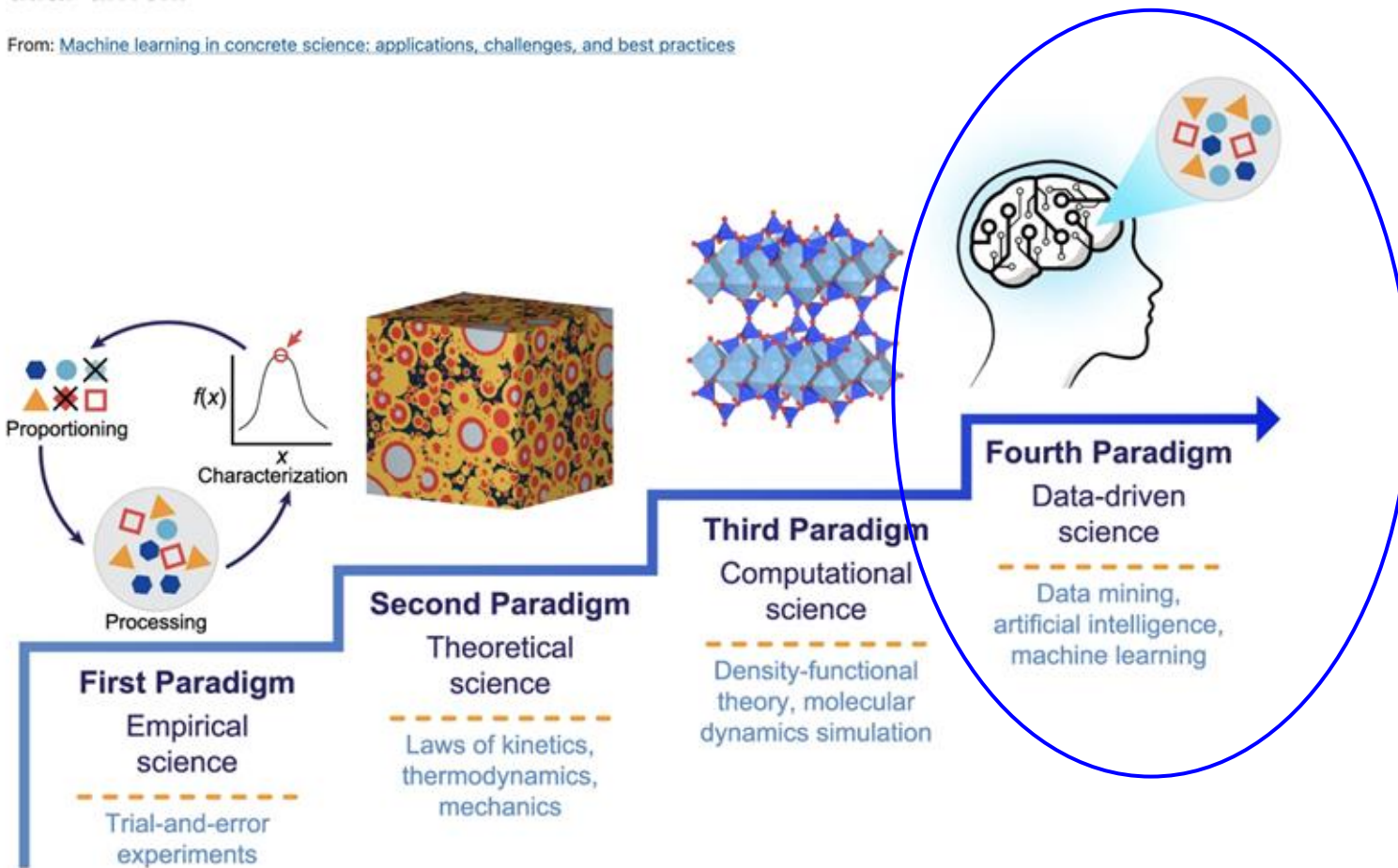
Rethinking US–Africa Health Partnerships

Building Digital Health Systems Without Losing Control: Africa's Dilemma



Fig. 1: Four paradigms of concrete science: empirical, theoretical, computational, and data-driven.

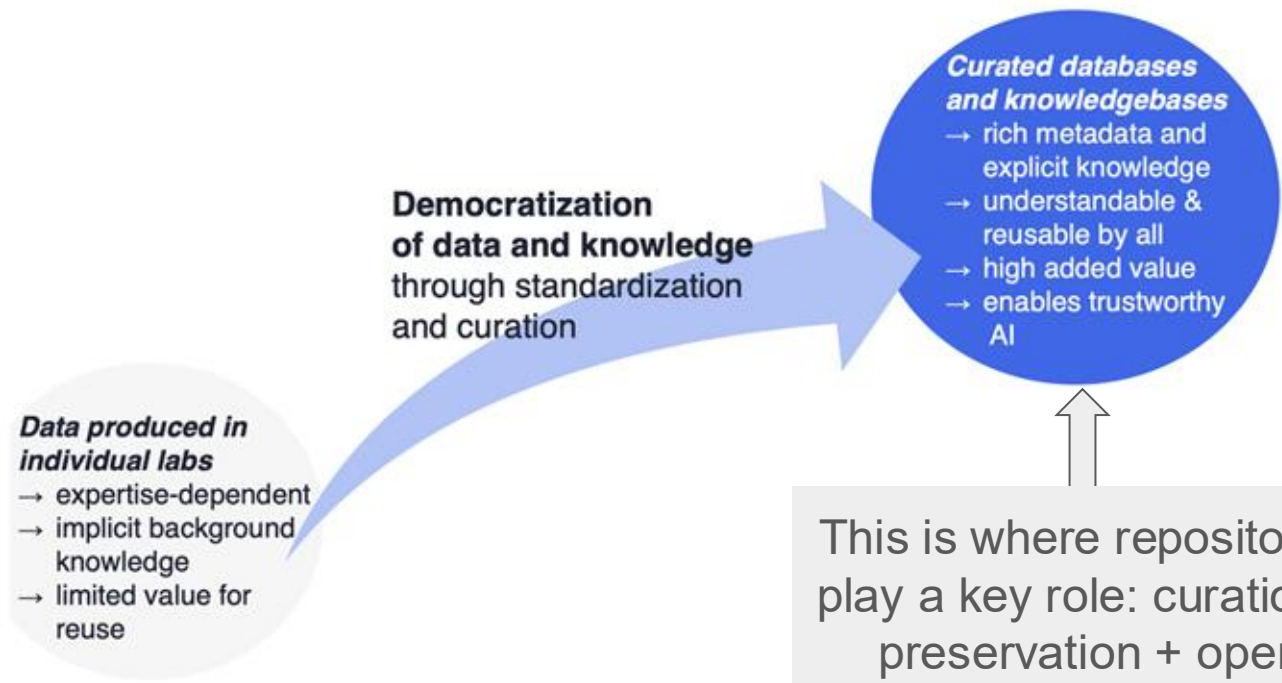
From: [Machine learning in concrete science: applications, challenges, and best practices](#)



Royal Society - “Relevant stakeholders should ensure access to high-quality datasets and interoperable data infrastructures across sectors and regions.

Fig. 1

From: [AI and the democratization of knowledge](#)



The need for democratization of data and knowledge. For data to become broadly valuable, it must be transformed into a form that can be correctly interpreted and used by a broad community comprising both non-experts and AI.

Welcome to the International Repositories Directory (IRD)

This is the International Repositories Directory, managed by the [Confederation of Open Access Repositories \(COAR\)](#). The directory aims to be an **authoritative source** of information about repositories, providing the community with an accurate and timely record of the current repository landscape. It contains information such as the types and versions of platforms in use, scope of repository collections, geographic location, institutional affiliations, functionalities supported, and so on.



Records in IRD

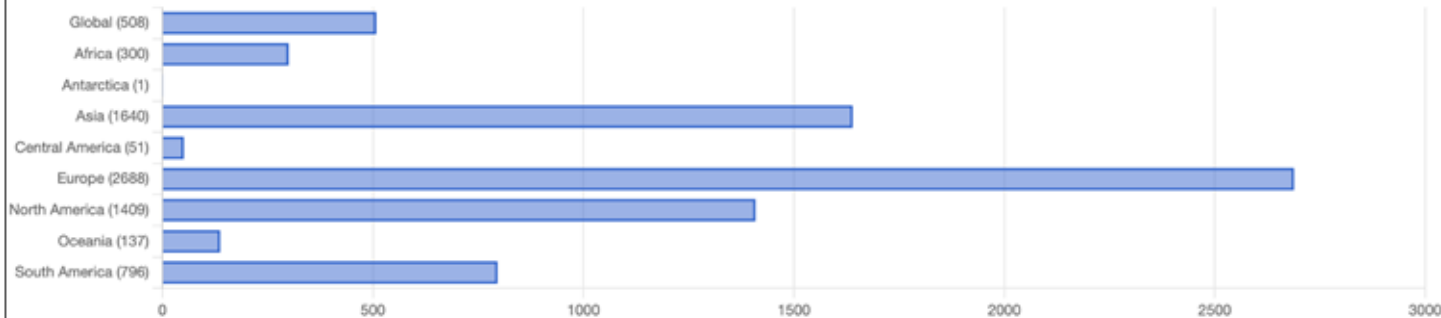
Total Records
7530

Verified Records
1149

Records Awaiting
Review
4111

Records Under Review
2270

Repositories (count) by Continent



<https://ird.coar-repositories.org/>

There are thousands of repositories across the world designed to provide open access to research outputs. These repositories collectively contain hundreds of millions of resources representing articles, datasets, images, code, etc and their content is continually growing.

Data-driven research is being
supercharged
by AI

(or, “data is the new oil”)

Detecting rare diseases in electronic health records using machine learning and knowledge engineering: Case study of acute hepatic porphyria

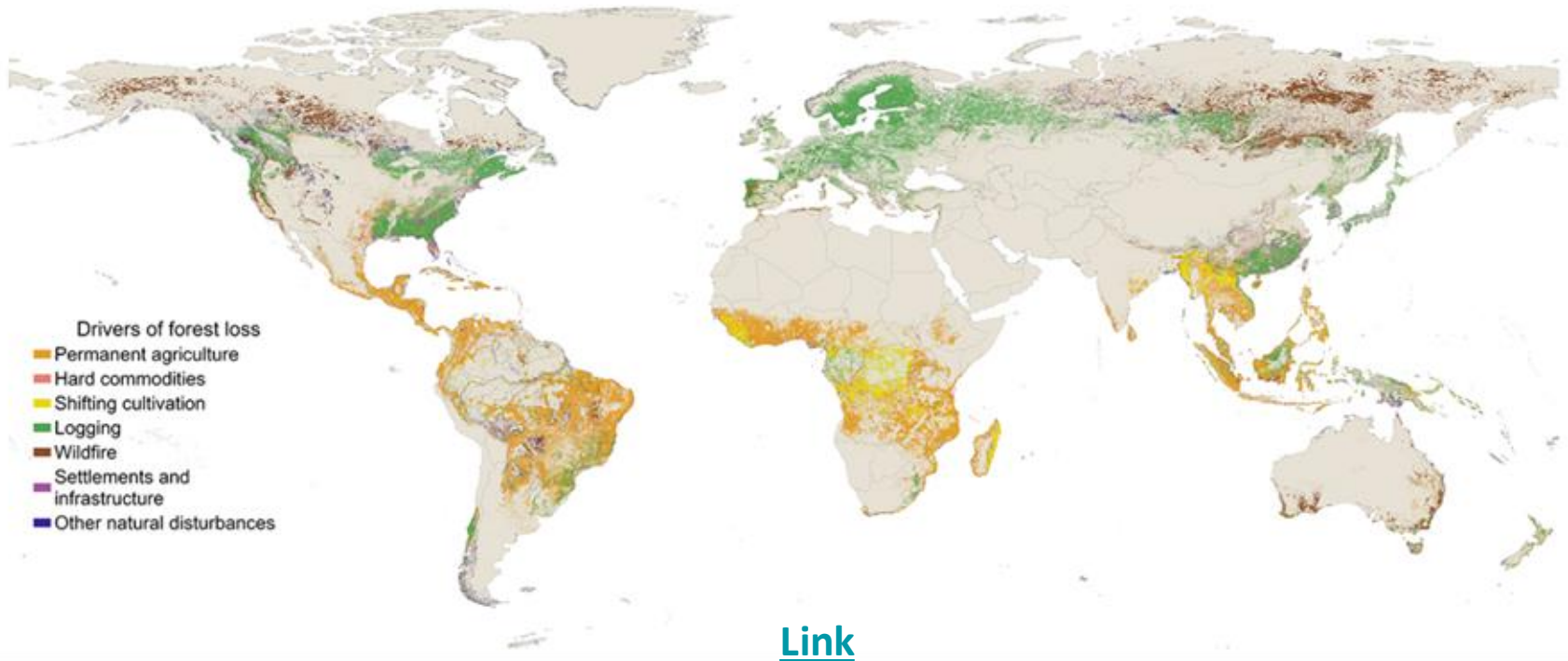
Aaron M. Cohen , Steven Chamberlin, Thomas Deloughery, Michelle Nguyen, Steven Bedrick, Stephen Meninger, John J. Ko, Jigar J. Amin, Alex J. Wei, William Hersh

Published: July 2, 2020 • <https://doi.org/10.1371/journal.pone.0235574> [Link](#)

Article	Authors	Metrics	Comments	Media Coverage	Peer Review
					

The combination of large data sets, machine learning techniques, and clinical knowledge engineering can be a powerful tool to identify patients with undiagnosed rare diseases. This level of precision in identifying potential cases of AHP from EHR data is much higher than would be expected by the prevalence of the disease.

We interpreted nearly 7000 reference samples to train a global neural network model that classifies the driver of tree cover loss with an overall accuracy of 90.5%. Our results show that permanent agriculture was the leading driver of forest loss globally, representing 35% of loss from 2001 to 2022.



[Link](#)



BRIDGING THE PAST AND FUTURE WITH
TEXTUAL ANALYSIS AND AI

Historical-Critical Textual Analysis and AI: Bridging the Past and Future



Tim Williams, PhD

AI algorithms can assist by analyzing vast corpora of historical documents, drawing connections across a broader swath of the record.

AI helps historians reconstruct the context in which texts emerged. By considering events, social norms, and cultural nuances, we gain a richer understanding.

How AI can help to save endangered species

Scientists are using artificial intelligence to fight biodiversity loss by analysing vast amounts of data, monitoring ecosystems and spotting trends over time.

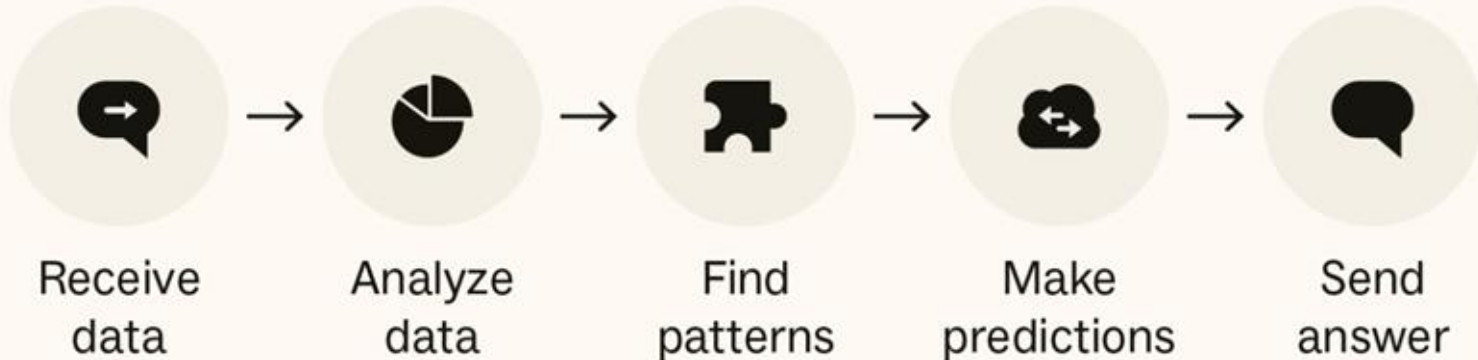
By [Tosin Thompson](#)



AI was used to rapidly analyse thousands of datasets to identify which species are most at risk globally

AI and machine learning are really good at:

- Analysing massive amounts of data to uncover patterns
- Developing predictive models that forecast future outcomes based on historical data
- Undertaking literature reviews of huge number of articles in many different languages





Navigating the Uneasy Interdependence of AI and Open Science



Mitigating the impact of AI bots

We have a number of concerns about the current AI-mediated ecosystem, but most urgently we want to prevent a situation where OA repositories start blocking machines because of aggressive bots traffic.

Digital content is the foundation for AI in research

1. For training AI models
2. For analysing data using AI models

< Go back

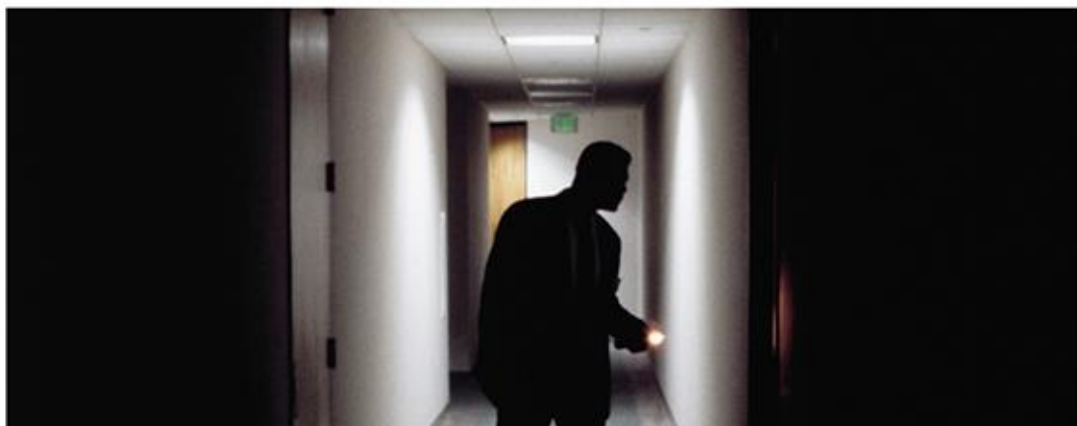
RESEARCH FORTNIGHT

26 AUG 2025 | 🔒

But... AI can't learn from what researchers don't share

By Michele Darrow and Mark Basham

Share

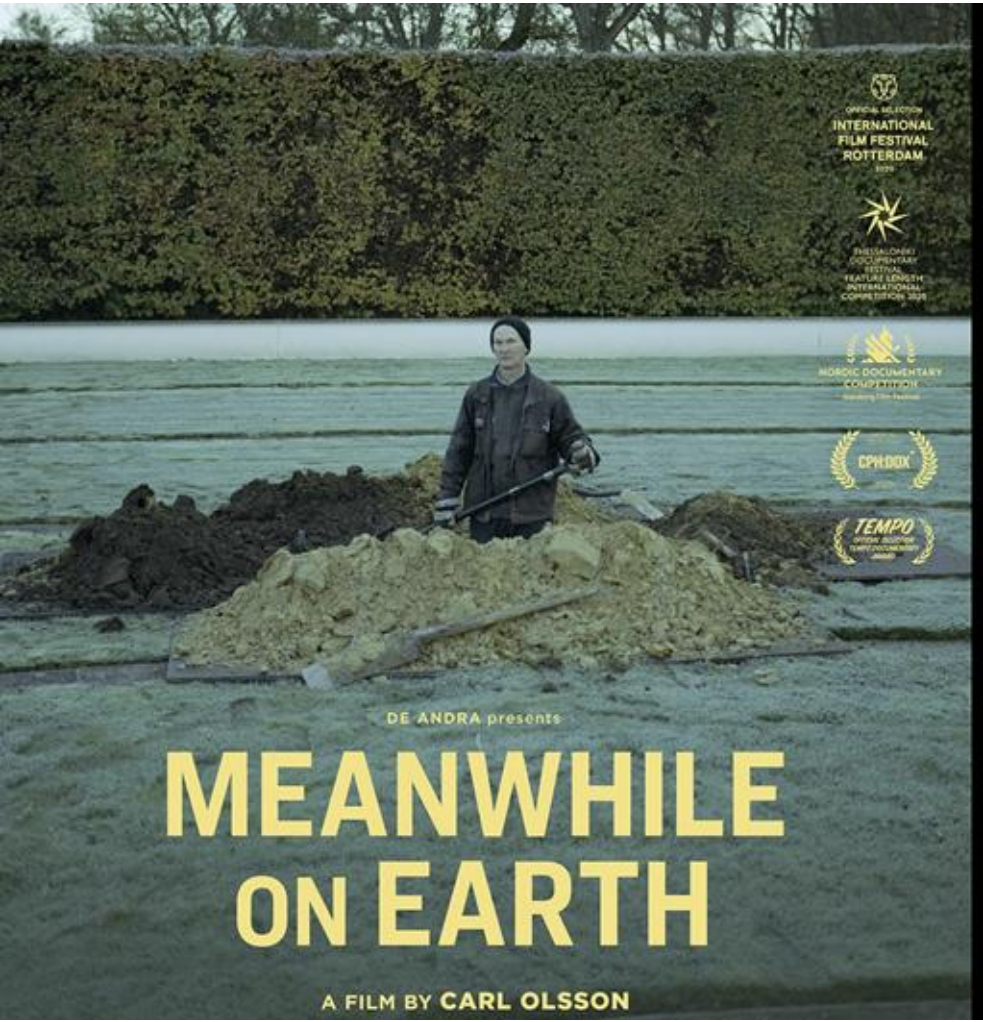




Data gaps hampering Africa's development agenda

JOSEPH MAINA

NOVEMBER 28, 2022



OFFICIAL SELECTION
INTERNATIONAL
FILM FESTIVAL
ROTTERDAM
2020

EMILIA ONERI
DOCUMENTARY
TESTING
FEATURE LENGTH
INTERNATIONAL
COMPETITION 2020

MOBIC DOCUMENTARY
COMPETITION
Featuring Film Festival

CPH:DOX
2020

TEMPO
OFFICIAL SELECTION
DOCUMENTARY
COMPETITION 2020

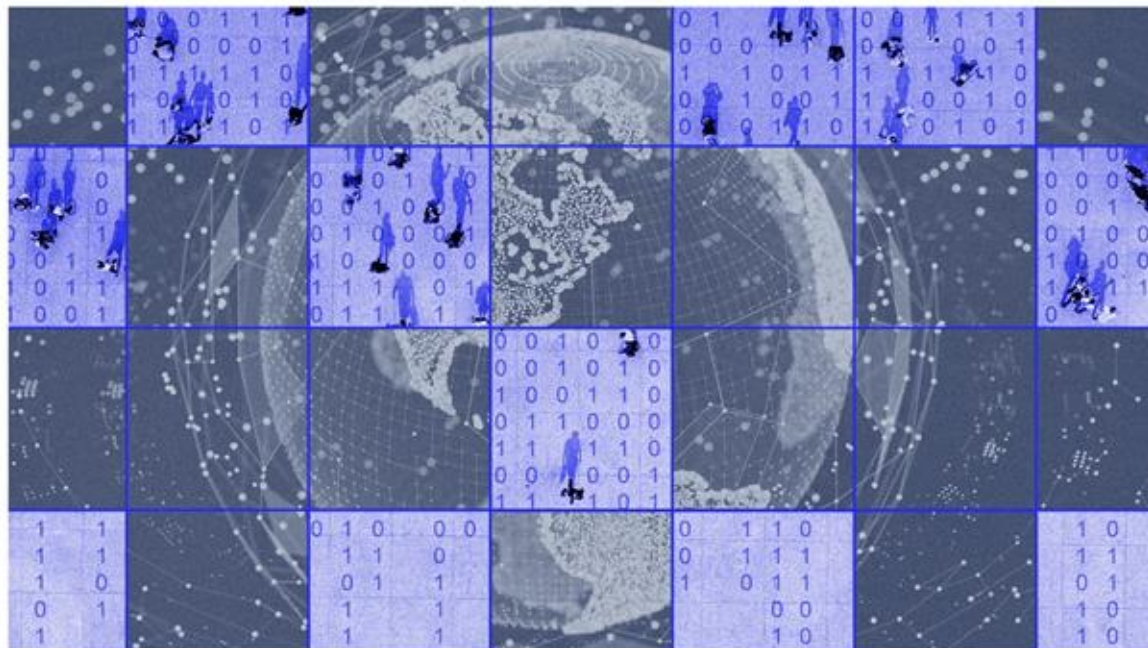
DE ANDRA presents

MEANWHILE ON EARTH

A FILM BY CARL OLSSON

Why Big Tech is threatened by a global push for data sovereignty

Countries are forcing tech giants to store citizen data locally, challenging the standard business model of harvesting data abroad while keeping profits at home.



A man with glasses, wearing a purple blazer over a black shirt, is speaking at a podium. He is holding a microphone in his right hand and looking upwards with a thoughtful expression. The background is dark with several bright spotlights. Large, bold white text is overlaid on the image.

ATTENTION: OCTOBER 2025

GOVERN OR BE GOVERNED

There are certain “flavours” of digital
sovereignty emerging in the research realm



Think Tank
European Parliament

Data Sovereignty in Research: Global Dependencies, Risks, and the European Response

Workshops [STOA](#)

SPAAK 7 C 50



“The dominance of **external and commercial databases** poses **strategic vulnerabilities** for academic freedom, innovation, and research continuity in Europe.” ([link](#))



Quem controla seus dados? Ciência Aberta, Colonialismo de Dados e Soberania na era da Inteligência Artificial e do Big Data

ISBN digital: 9788572214742

ISBN impresso: Publicação apenas digital

DOI: 10.31560/pimentacultural/978-85-7221-474-2

Autor: Fabiano Couto Corrêa da Silva

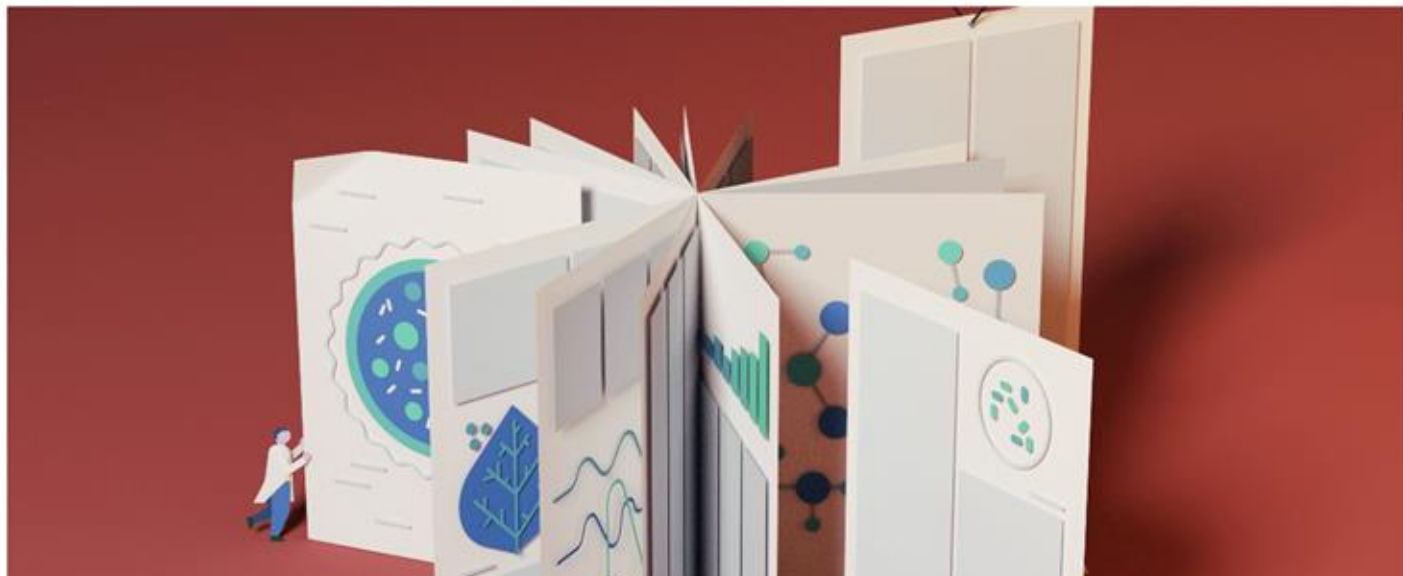
[Download do ebook](#)

“When researchers in the Global South deposit their data in repositories controlled by foreign companies, use proprietary analytical platforms, or depend on cloud infrastructures located in other countries, they are, to some extent, **ceding sovereignty over a strategic resource.**” ([link](#))

China's research clout leads to growth in homegrown science publishing

Self-reliance and a wish to set the agenda might explain concerted push to produce domestic journals.

By [Brian Owens](#)





LOST SCIENCE

**US science after a year
of Trump**

African Public Health Conference Calls for 'Health Sovereignty' and Prioritizing Primary Care Ahead of G20

Africa CDC 27/10/2025 · Matthew Hattingh

Share this:



[Link](#)



“Control over health data influences who sets research agendas, who benefits from innovation, and who bears the risks when systems fail or data is misused”. ([link](#))

But... digital sovereignty does not mean protectionism



We need to also remain open - Research is global and many of our most critical issues are international

Based on these two **sometimes divergent** drivers (AI, openness and digital sovereignty) how should Africa proceed with developing its infrastructure?

Is there a model?

EOSC in the Countries

Click on the map to access each of EOSC-A's 40 country pages including:

- Overview of national ecosystem and EOSC contact points
- National news, events, and Open Science policies
- National contributions to EOSC and EU project involvement
- EOSC-A membership roll
- Photos, videos, interviews and more!

[Explore the EOSC-A country pages](#)



PubScholar
Public Academic Resources ,Serving Open Science Innovation

News [more >](#)
PubScholar collaborative journal
"World Science and Technology
Research and Development" h...
2026-01-13

主题 Discover what interests you...

Retrieved

Advanced Search

99,650,764	1,017,531	4,016,837	105,542,177	6,167	1,165,654	123,759
Journal articles	Dissertation	Preprint paper	Patent literature	News news	Scientific data	Book monographs

AI Retrieval
An intelligent retrieval system that integrates AI large models

Intelligent Review
Generate intelligent review reports to improve your research efficiency and insights

AI Conversations
Dig deep into the knowledge behind the data through the intelligent interaction of AI Q&A

Literature reading
Utilize AI large models to help you quickly capture the core points of the document

The PubScholar public welfare academic platform is a platform built by the Chinese Academy of Sciences, as the main force of the national strategic scientific and technological force, **to fulfill the responsibility of the "national team" for academic resource protection**, and to meet the basic guarantee needs of academic resources for scientific and technological innovation in the national scientific and technological community and the whole society. ([link](#))



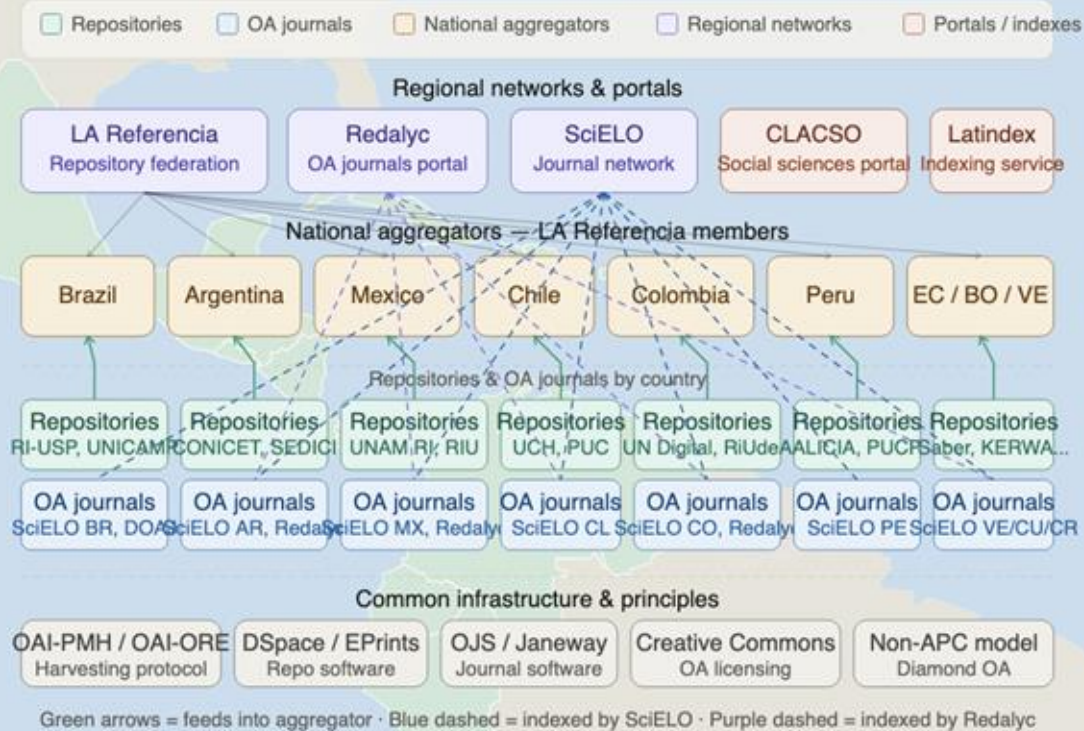
The withdrawal of the US from UNESCO: What does this mean for Open Science?

November 05, 2025 • Long-read and Opinion & Commentary • 8 min read

United States - To decentralized and lacking vision

the withdrawal of the US from UNESCO and US legislation being increasingly at odds with Open Science values raises concerns regarding Open Science infrastructure in the US. While much is still unclear, our authors argue that the implications for Open Science could be profound.

Latin American Open Science Ecosystem



Regional level - global discoverability and visibility

National level - public access and research assessment

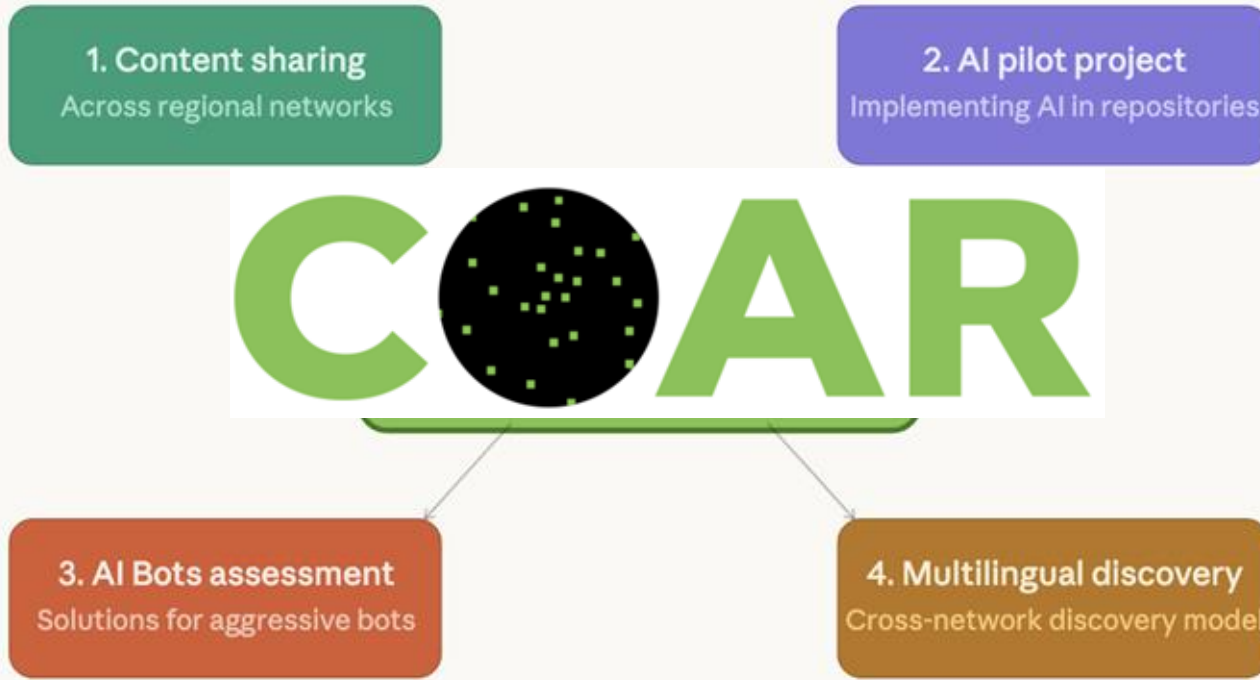
Institutional level - local autonomy, curation and support

International standards and common practices and norms

Recommendations for supporting data-driven science in Africa, for Africa

1. Maintain African data/articles in African infrastructures
2. Use open source repository platforms that are internationally interoperable
3. Identify priority areas for collecting data - you can't do everything
4. Curate the metadata - this adds critical context (machines are getting better at using unstructured data, well curated data contributes to much better outcomes)
5. Work on an interoperable African discoverability layer (like LA Referencia) so African content is visible and available beyond Africa

COAR is a community that works on issues to help repositories and repository networks to navigate the future



Interoperability is the Key to Global Collaboration



**ACT
LOCAL
THINK
GLOBAL**

Thank You!

www.coar-repositories.org

Kathleen Shearer, Executive Director, COAR

 Confederation of
Open Access Repositories