

8th International
Conference on
BIG DATA
& Data Science for Official Statistics

BILBAO 2024

Informing Climate Change and
Sustainable Development Policies
with Integrated Data

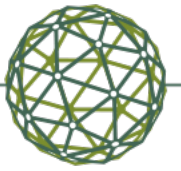
BILBAO. SPAIN **10-14 JUNE 2024** **#UNBigData2024**

Knowledge Commons for Global Sustainability Challenges

Building an integrated knowledge space for the global statistical community

Ferdinando Villa, BC3

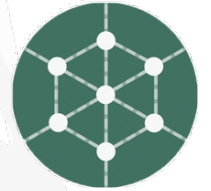




State of the art: a rationale for the ARIES approach

1.

Decision-makers with limited data and technical capacity often lack **access to scientific knowledge**. Many are left behind due to **cost or technology barriers**.



2.

Ever-increasing volumes of data are held in silos – different disciplines, geographies, data types and access rights - making it challenging to connect information and make sense of it.



3.

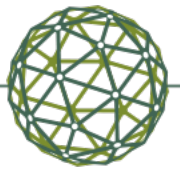
Public **trust** is one of the biggest hurdles faced by AI technologies. People struggle to accept the decisions and answers that AI-powered tools provide as many do not make their inputs, operations, and end goals visible.



4.

The AI technology ecosystem is currently dominated by Big Tech - enclosed assets - for profit perspective. Although much software is open-source, **access to data remains tightly controlled**.





The role of Artificial Intelligence (AI)

Governmental agencies and policy makers often face **high barriers** when generating environmentally-related knowledge:

- Environmental modelling has high data needs, which must fulfil necessary standards of quality and consistency;
- Large amounts of data result in long processing times, making compilation a slow exercise;
- Statistical-environmental analyses often make use of biophysical models which require technical expertise

The statistical community would **benefit** greatly from data and models which are **Findable, Accessible, Interoperable and Reusable (FAIR)**.

Beyond FAIR, the **deep knowledge integration** enabled by AI can have consequences much beyond convenience.



Findable



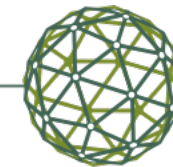
Accessible



Interoperable



Reusable



ARIES: Artificial Intelligence for Environment & Sustainability

- # 1 It is a **modelling technology**, rather than a collection of models or specific program/application,
- # 2 It is an **AI modeller**, based on **machine reasoning**, a less known branch of AI;
- # 3 It defines a variety of data, models and the relationships between them using **consistent and uniform terms**. This allows different data and models to be used together, depending on which data and models are “most appropriate” for the context set by the user;
- # 4 It uses AI to determine the “**most appropriate**” data and models for users’ requests.

Reasoning
algorithms

+

Decision
rules

+

Multidisciplinary
semantics

+

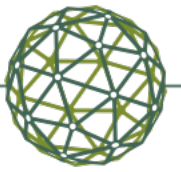
Open data
& models

+

Open-source
software

=

ARIES: Fast, FAIR
multidisciplinary
modeling

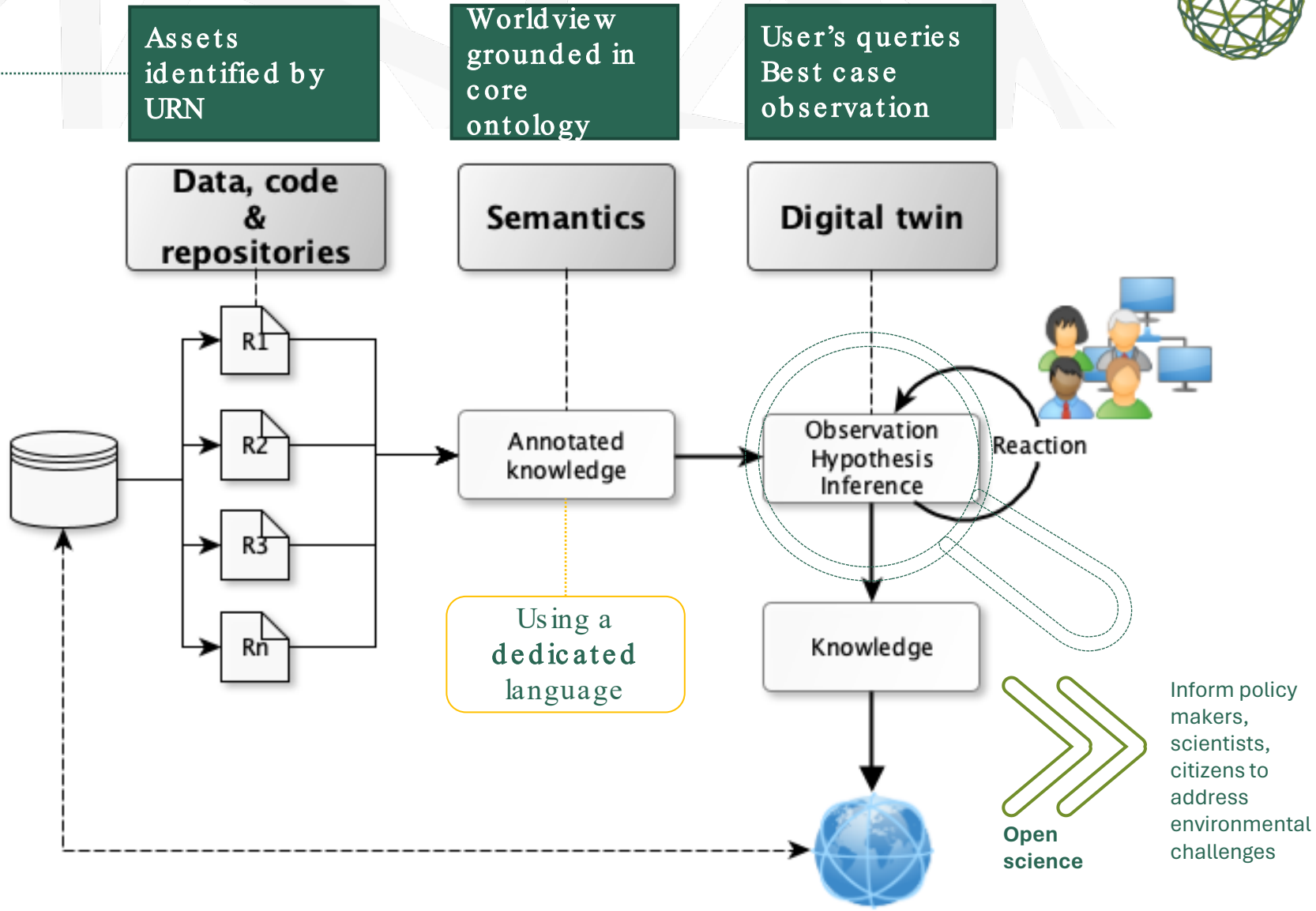


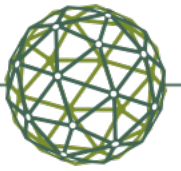
```
imf.aries.risk:imf_kenya_pilot.lst_2018_2023_kenya
```

- Copernicus Program - sentinel satellite data
- NASA/USGS Landsat Program Landsat satellite data



Environmental monitoring, Disaster management, Urban planning – **K-LAB Modelling Paradigm**





Potential use cases for ARIES



Spatial economic
valuation of
ecosystem
services



Conservation
planning



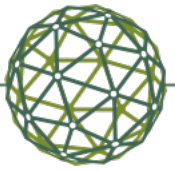
Spatial policy
planning
(derived from
EO data)



Forecasting
changes in
ecosystem
service provisioning



Natural capital
accounting



The ARIES for SEEA Explorer

Available since April 2021

1

An application of the ARIES technology to compile ecosystem accounts that are consistent with the **SEEA Ecosystem Accounting**;

2

It utilizes freely available remote-sensing derived **data and models**;

3

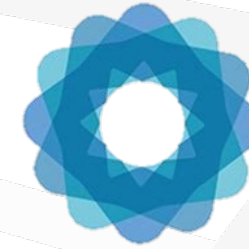
It can generate accounts for **any** user-specified **terrestrial area** in the world;

4

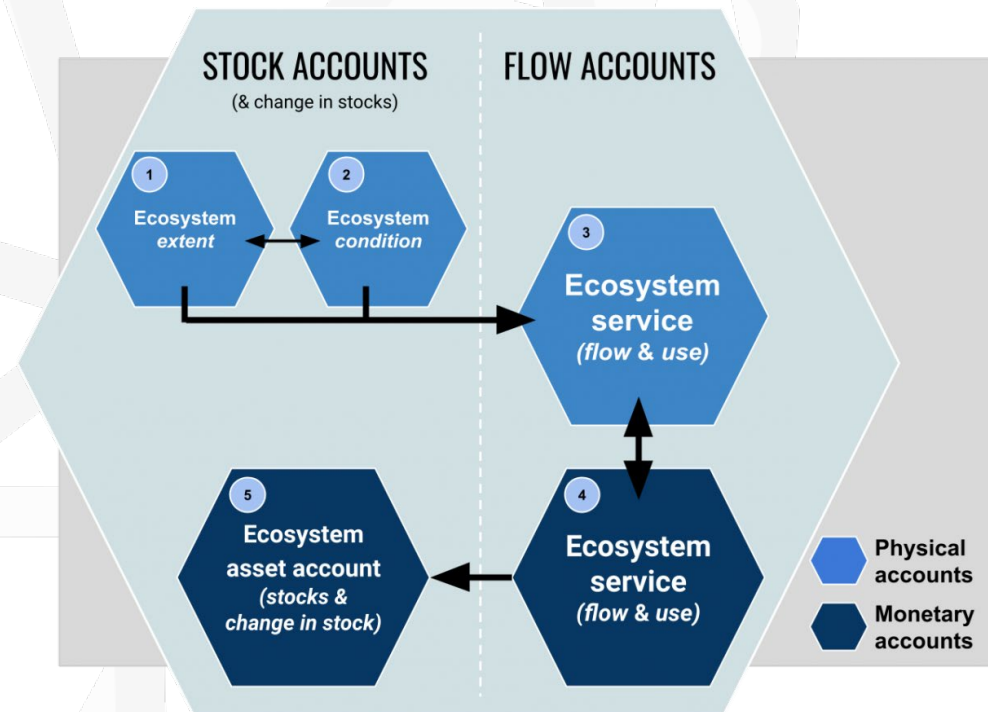
It **rapidly** computes these accounts online, using a web browser;

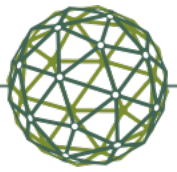
5

It generates a comprehensive **report, fully documenting the data, models, coefficients and methods** used.



System of
Environmental
Economic
Accounting





Some ARIES for SEEA output examples

1

A combination of statistical and spatial analysis summarized in **Tables(1)** and **Maps(2)**. Tables are SEEA compliant.

Table 1. Ecosystem extent account - net balance

1

Table 1. Occurring ecosystem types (selected level 3 Ecosystem Functional Groups of the IUCN Global Ecosystem Typology 2.0)

	Coastal saltmarsh reedbed	Cropland	Urban industrial ecosystem	Polar tundra desert	Polar alpine rocky outcrop	Alpine grassland shrubland	Ice sheet glacier permanent snowfi
Extent at start of 2012 (km ²)	1987.60	747162.95	3849.41	125.80	2742.39	24530.55	1232.82
Extent at start of 2019 (km ²)	2037.92	771970.26	4729.99	125.80	2817.87	24354.44	1232.82

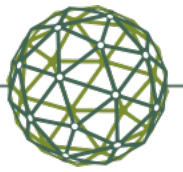
Net char

Figure 1. Ecosystem type

2

Figure 1

© 01/01/2012 01/01/2013 01/01/2014 01/01/2015 01/01/2016 01/01/2019



Some ARIES for SEEA output examples

2

Full transparency for replicability and traceability through **Reports(1)**, a **Resource Section(2)** & a **Dataflow Diagram(3)**.

1. Introduction

Disclaimer

The designations employed and the presentation of material on this map and any map used in this application do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

This application allows users to select different background backgrounds such as Open Street Map are available for accounts presented in this report are based on the con

SEEA-EEA

The System of Environmental-Economic Accounting values of services provided by ecosystems changes of ecosystems and the value of the services they produce

The U.N., in collaboration with the ARIES team, has mapping, and valuation for various accounts including compiled for any region on Earth using global data, wit

1.1. Ecosystem Extent

The Ecosystem Extent Account is the first SEEA EA¹ used in all other accounts, so are fundamental to SEEA

Ecosystems are defined as units whose functioning Ecosystems in this context should not be confused with

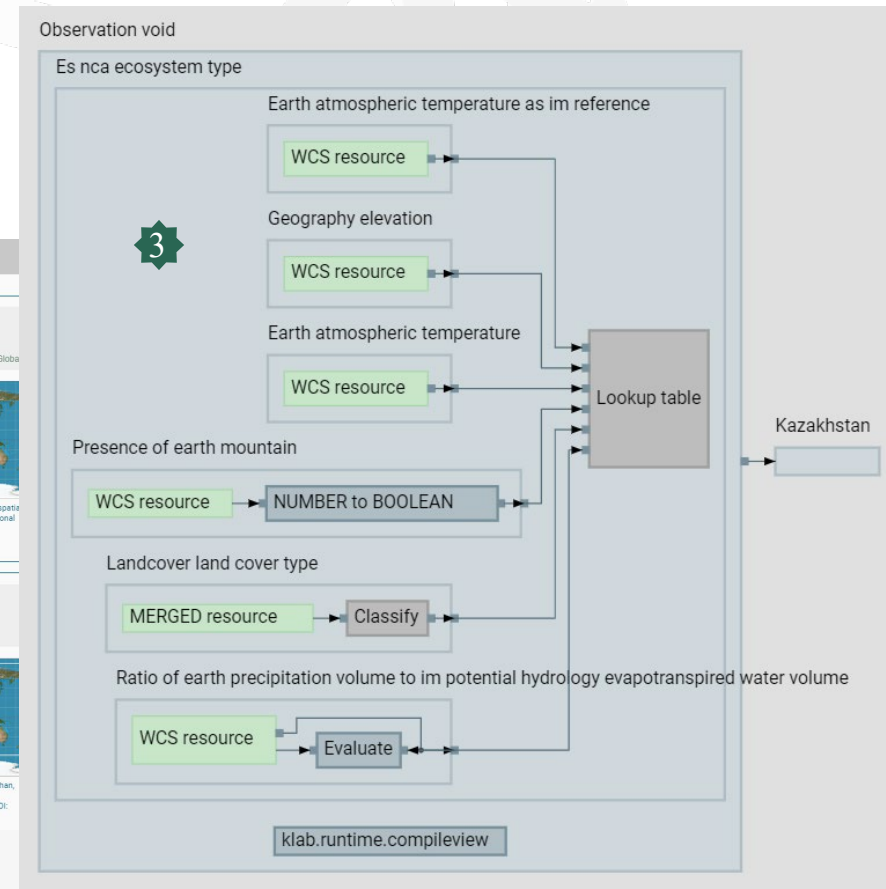
IUCN's Global Ecosystem Typology is the standard applications relied exclusively on land cover data².

A full ecosystem extent account includes gross change of the same ecosystem type and for each accounting forest clearing for agricultural land), natural expansion (adjustments to initial estimates resulting from improved ecosystem services in the formulation of the SEEA EA

July average temperature, WorldClim 2.1
WorldClim
Mean temperature (degrees C), 1970-2000 average for July at 30-second spatial resolution

Global Mountain Explorer (GME)
USGS

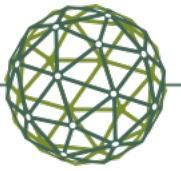
This resource was developed by the U.S. Geological Survey (USGS), in partnership with Esp. the Center for Development and Environment of the University of Bern (CDE), the Global Mountain Biodiversity Assessment (GMA), and the Mountain Research Initiative (MRI). The work is part of a Group on Earth Observations (GEO) initiative called GEO ONDIME, GEO's Global Network for Observations and Information in Mountain Environments The K3 resource was developed using a finer spatial resolution (250 m) DEM and feature-based extraction algorithms with variable NAW sizes used to extract a set of global Hammond landforms with 16 landform types, of which four were mountain classes. E. H. Hammond was a pioneer of landform mapping and described three parameters for distinguishing different types of plains, hills, mountains, and tablelands. The three classification parameters are slope, relative relief, and profile, where the profile parameter assesses the amount of relatively flat terrain in upland locations to delineate tablelands. The 250 m global Hammond landforms product was based on an automated extraction of classes in a GIS environment, and the K3 mountains product was an export of the four mountain classes into a global mountains datalayer.



Thank you!



#UNBigData2024



Inequalities



Access to scientific knowledge



Technology barriers



Fragmented information



Data in silos

Lack of transparency



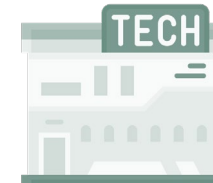
AI distrust

Lack of shared narrative

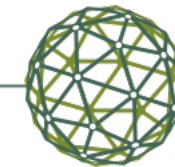


Crisis of collective intelligence

Lack of data control



Ownership



Integrated Modeling

is a practice meant to maximize the value of scientific information by ensuring its modularity, reusability, interoperability and traceability throughout the scientific process.



1. Linkage

Combine independently produced scientific products into workflows that would be too complex for individual humans to conceive, validate and navigate.



2. Integration

Integrate different modelling paradigms from simple (e.g., deterministic and probabilistic models) to complex approaches (e.g., agent-based and networks) depending on context and scale.



3. Rescaling

Rescale smartly across scales, from local to global, promoting adaptive solutions that are automatically customized to the scale of observation.



4. Adaptive contextualization

Adaptively incorporate the best-available knowledge, from curated global public datasets to “big data” to user-provided data.



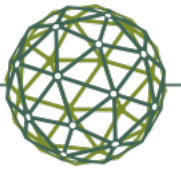
5. Delivery

Adopt shared, non-ambiguous **semantics** in the implementation, documentation and dissemination of products.



6. Tracking

Track quality, **uncertainty** and provenance throughout modelling workflows.



Relevance of semantics and ontologies for ARIES

- **Unambiguous identification** of concepts, data, models and results.
- **Description** of real environmental scenarios.
- **Interoperability** of data and models.
- **Transparency** with users in the methodologies applied and results obtained.

Transformation of Knowledge to Knowledge – an insight into modelling in

