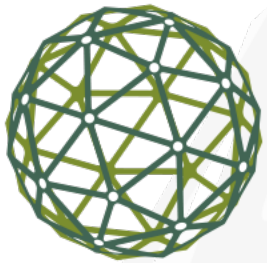


8th International Conference on BIG DATA & Data Science for Official Statistics

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

ARIES
ARtificial Intelligence for Environment & Sustainability

United Nations BigData

EUSKO JAURLARITZA GOBIERNO VASCO

GOBIERNO DE ESPAÑA MINISTERIO DE ASUNTOS EXTERIORES, UNIÓN EUROPEA Y COOPERACIÓN

BC3 BASQUE CENTRE FOR CLIMATE CHANGE Klima Aldaketa Ikergai

EXCELENCIA MARÍA DE MAEZTU 2023-2027

INE Instituto Nacional de Estadística

Eustat EUSKAL ESTATISTIKA ERAKUNDEA INSTITUTO VASCO DE ESTADÍSTICA

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:

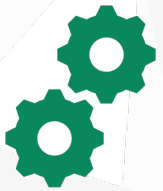
- Environmental modelling has high data needs, which must fulfil necessary standards of quality and consistency;
- Large amounts of data result in long processing times, making compilation a slow exercise;
- Statistical-environmental analyses often make use of biophysical models which require technical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

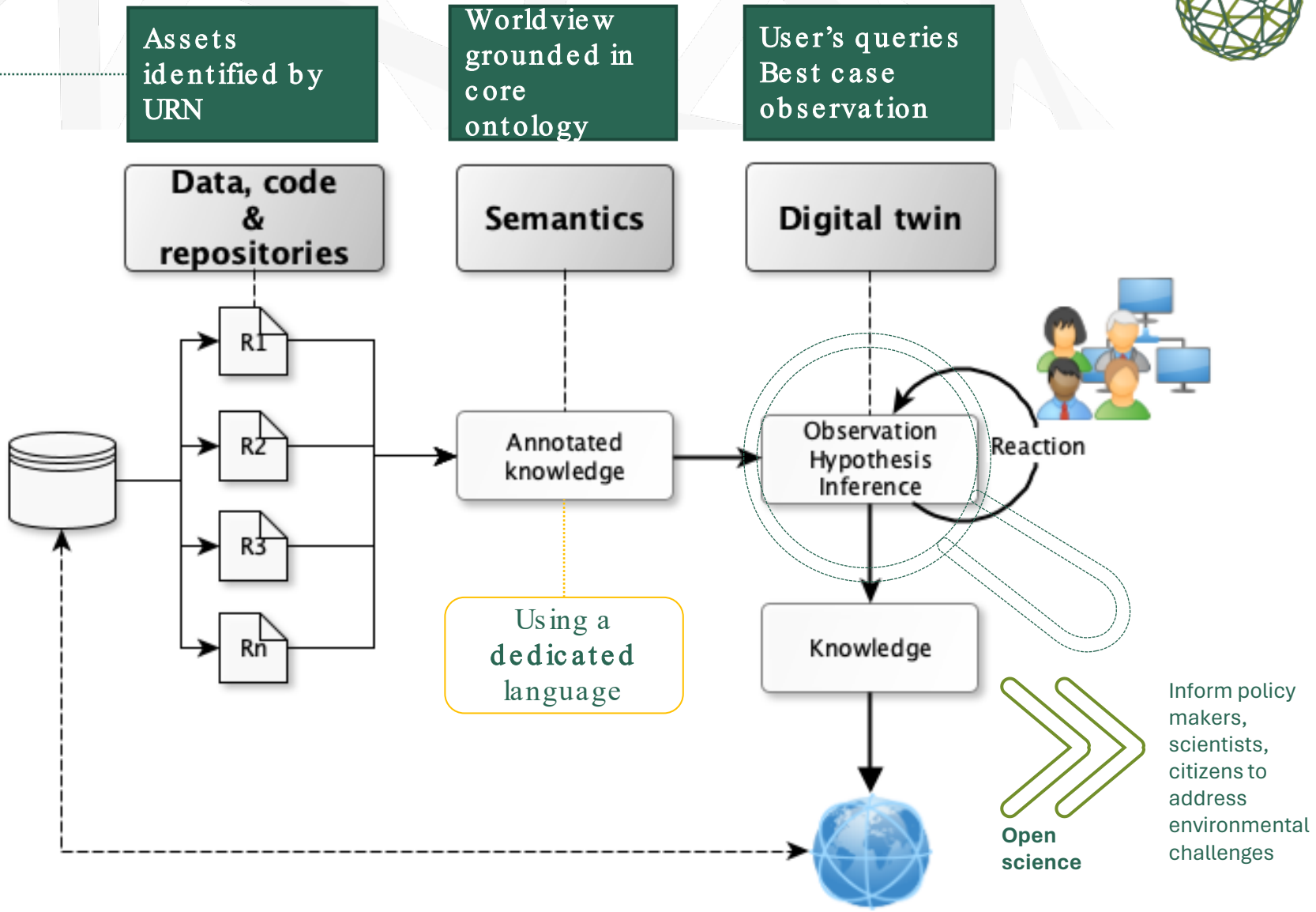
Reasoning algorithms + Decision rules + Multidisciplinary semantics + Open data & models + Open-source software = ARIES: Fast, FAIR multidisciplinary modeling

`imf.aries.risk:imf_kenya_pilot.lst_2018_2023_kenya`

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

Atmosphere (CAMS) · Marine (CMEMS) · Land (CLMS)
Climate (C3S) · Emergency (EMS) · Security

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;

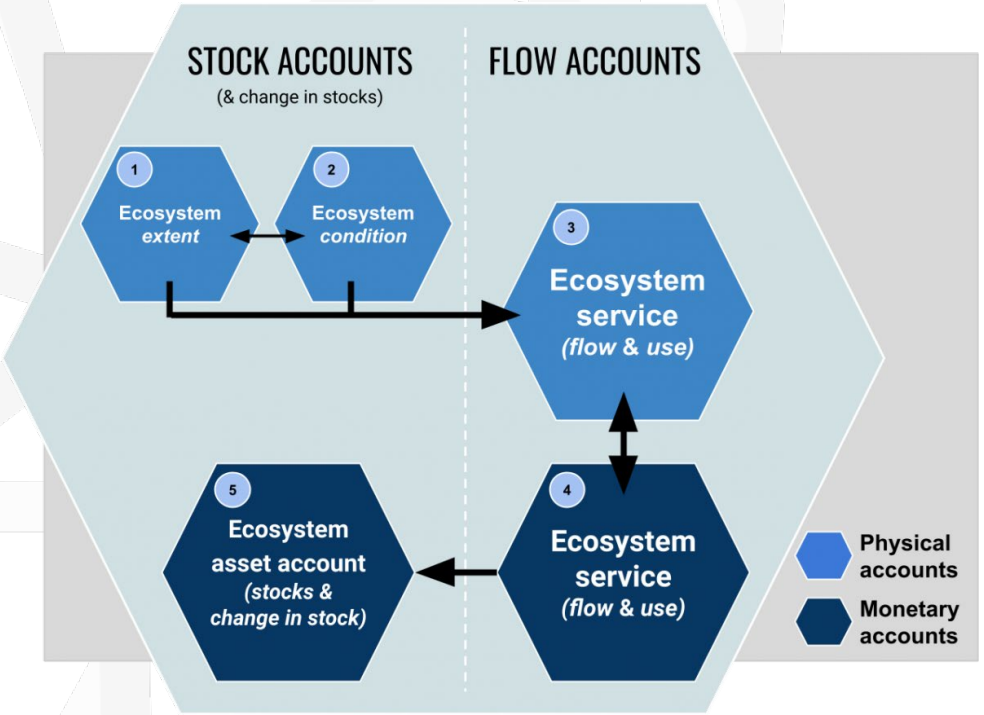
 # 2 It utilizes freely available remote-sensing derived data and models;

 # 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

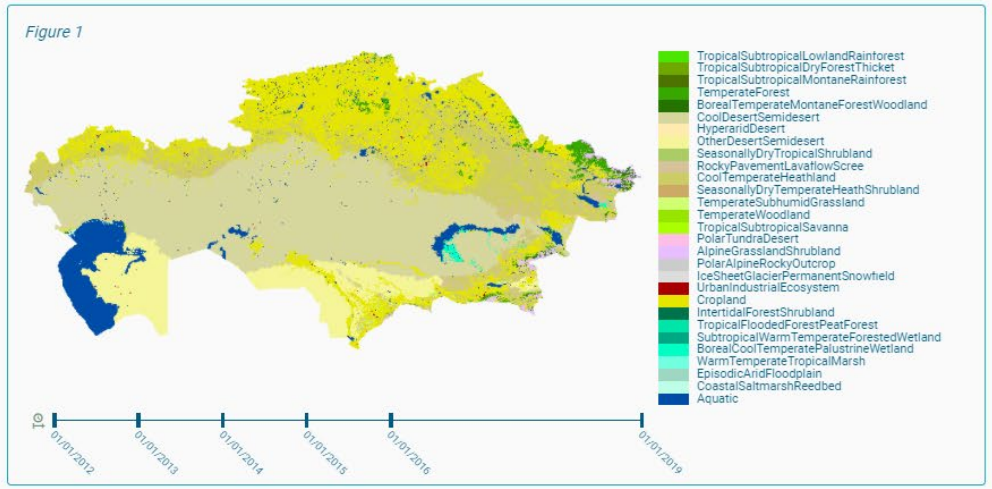
1

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 (km²)	1987.60	747162.95	3849.41	125.80	2742.39	24530.55	1232.82
Extent at start of 2019 (km²)	2037.92	771970.26	4729.99	125.80	2817.87	24354.44	1232.82

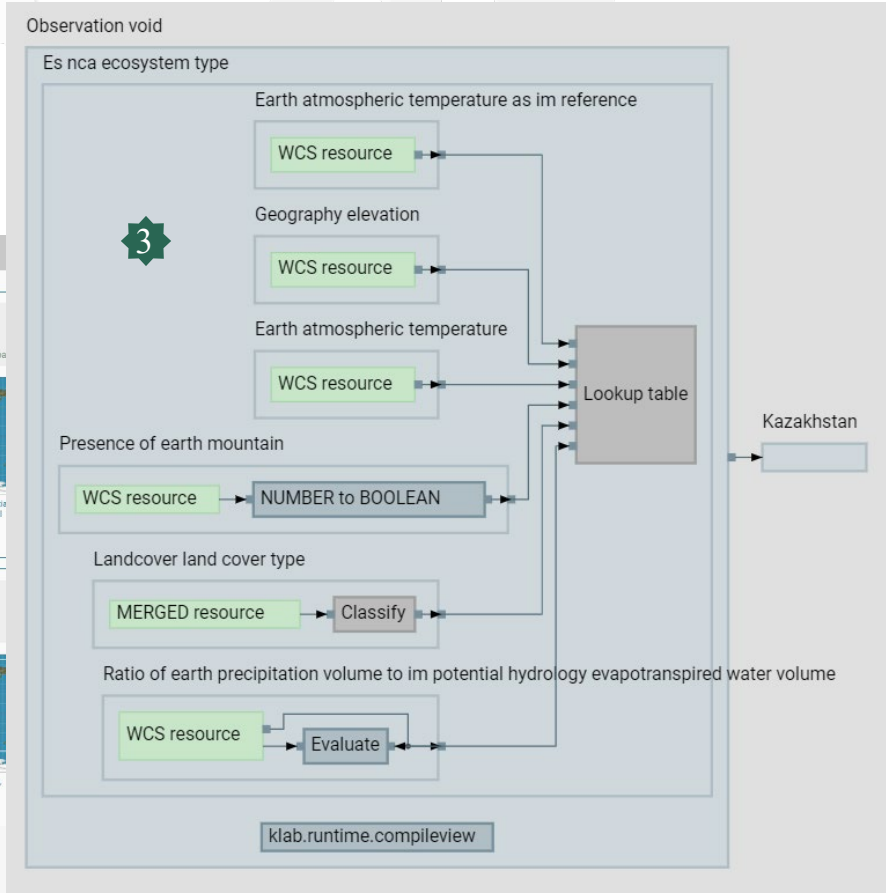
Figure 1. Ecosystem type

2



Some ARIES for SEEA output examples

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



Thank you!

ARIES
ARtificial Intelligence for Environment & Sustainability

United Nations BigData

EUSKO JAURLARITZA GOBIERNO VASCO

Bizkaia foru aldundia diputación foral

GOBIERNO DE ESPAÑA
MINISTERIO DE ASUNTOS EXTERIORES, UNIÓN EUROPEA Y COOPERACIÓN

BC3 BASQUE CENTRE FOR CLIMATE CHANGE Klima Aldaketa Ikergai

EXCELENCIA MARÍA DE MAEZTU 2023-2027

INE Instituto Nacional de Estadística

Eustat EUSKAL ESTATISTIKA ERAKUNDEA INSTITUTO VASCO DE ESTADÍSTICA

#UNBigData2024

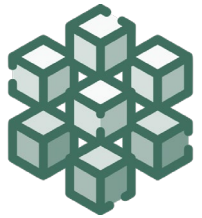
Inequalities

Access to scientific knowledge

Technology barriers

Cost

Fragmented information



Data in silos

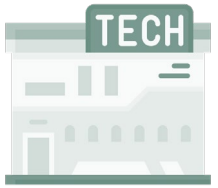
Lack of transparency

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.

Balbi et al. The global environmental agenda urgently needs a semantic web of knowledge. Environmental Evidence 11, 5 (2022). doi:10.1186/s13750-022-00258

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