# System integration track


Ferdinando Villa, Stefano Balbi

BC3 & Ikerbasque

PEOPLE-EA, Athens, 23 May 2024

# Blueprint of the integrated architecture

Three main components:

1. ARIES, as semantic front-end of the integrated system, powered by the k.LAB software stack;
2. OpenEO, as data and processing workflows supplying engine (via UDP/UDF catalogues);
3. INCA, as dedicated NCA data preparation routines and models, according to EU standards.

# Architecture

**ARIES-OpenEO-UDP/UDF catalog**
**(public repository to access INCA functionality & advanced OpenEO functionality)**

Access via k.LAB client API in UDF

OpenEO adapter

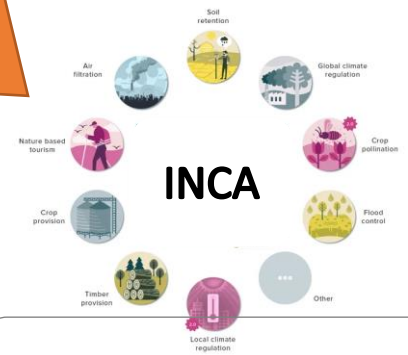
**Data and Processing Engine**

data pre-processing

models & methods

k.LAB
Knowledge Integrated Modelling
SEEA

Access via k.LAB client API




**Semantic front-end**

**NCA standard for EU**

# A broader perspective

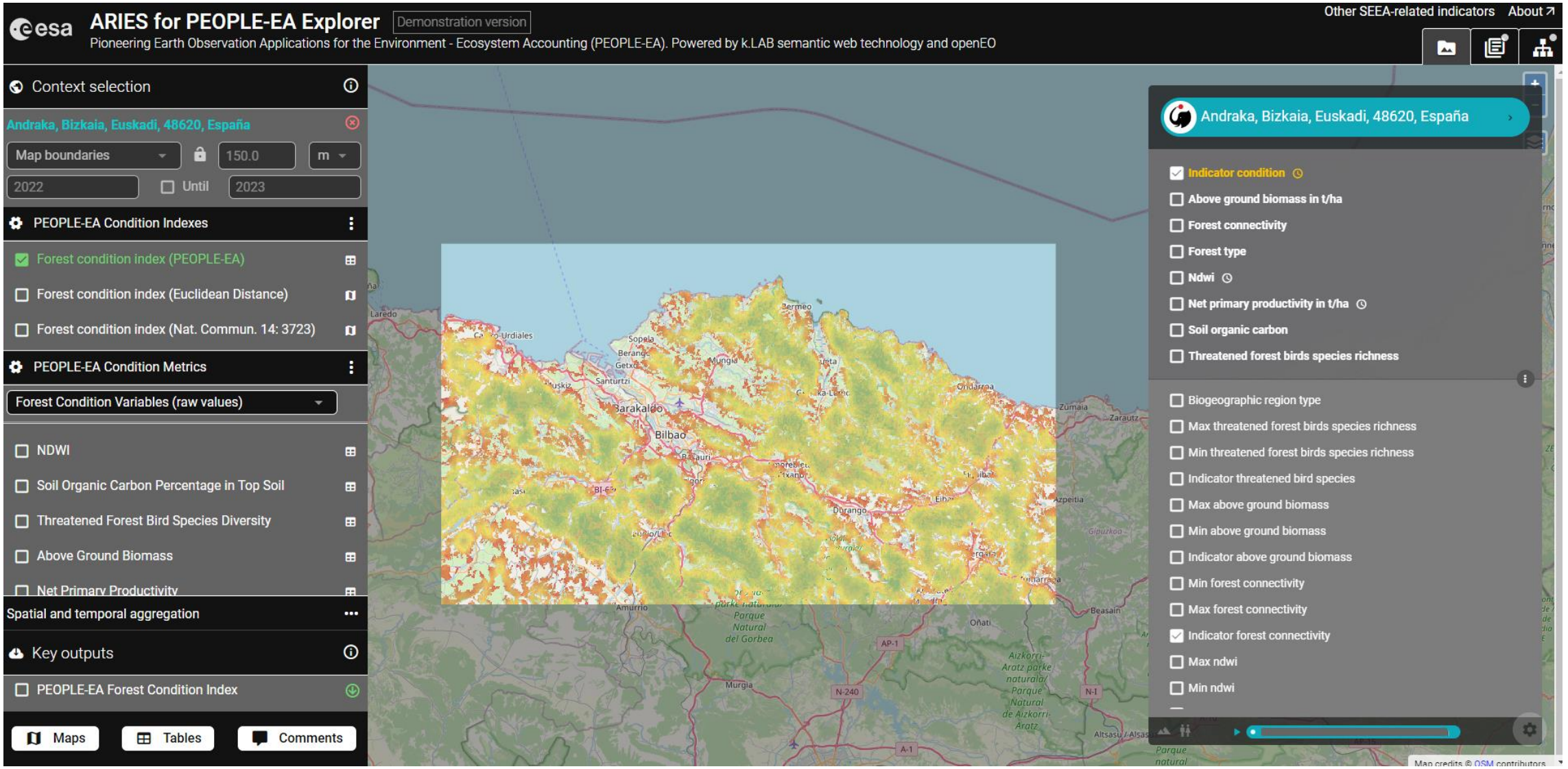
1. Navigating distributed but connected data and models through semantics for better Natural Capital Accounting (NCA) using earth observation (EO).

2. A semantic web of scientific observations for NCA using EO.

# Implemented access possibilities

1. Use of ARIES resources within INCA computational workflows via a k.LAB Python client.
2. Use of INCA models in OpenEO via direct usage of INCA UDP/UDF in Python
3. Use of OpenEO processing in INCA workflows (data pre-processing)
4. Obtaining INCA data and/or using INCA models in ARIES through OpenEO via the OpenEO-Adapter built for ARIES: k.LAB calls INCA UDP in OpenEO
5. OpenEO data processing in ARIES: ARIES calls OpenEO catalogues & processes directly.
6. Circle test resources from the different platforms: Python calls ARIES (via k.LAB client API), which calls INCA models through OpenEO (via OpenEO-Adapter & UDP/UDF catalogue).

# What you see


esa
ARIES for PEOPLE-EA Explorer
Demonstration version
Pioneering Earth Observation Applications for the Environment - Ecosystem Accounting (PEOPLE-EA). Powered by k.LAB semantic web technology and openEO
Other SEEA-related indicators
About
Context selection
Andraka, Bizkaia, Euskadi, 48620, España
Map boundaries
150.0
m
2022
Until
2023
PEOPLE-EA Condition Indexes
Forest condition index (PEOPLE-EA)
Forest condition index (Euclidean Distance)
Forest condition index (Nat. Commun. 14: 3723)
PEOPLE-EA Condition Metrics
Forest Condition Variables (raw values)
NDWI
Soil Organic Carbon Percentage in Top Soil
Threatened Forest Bird Species Diversity
Above Ground Biomass
Net Primary Productivity
Spatial and temporal aggregation
Key outputs
PEOPLE-EA Forest Condition Index
Maps
Tables
Comments
Andraka, Bizkaia, Euskadi, 48620, España
Indicator condition
Above ground biomass in t/ha
Forest connectivity
Forest type
Ndwi
Net primary productivity in t/ha
Soil organic carbon
Threatened forest birds species richness
Biogeographic region type
Max threatened forest birds species richness
Min threatened forest birds species richness
Indicator threatened bird species
Max above ground biomass
Min above ground biomass
Indicator above ground biomass
Min forest connectivity
Max forest connectivity
Indicator forest connectivity
Max ndwi
Min ndwi
Bilbao
Barakaldo
Durango
Laredo
Parque Natural del Gorbea
Aizkorri-Aratz parke naturala/ Parque Natural de Aizkorri-Aratz
Map credits © OSM contributors

# Behind the curtains

1. EO retrieval and processing
   - ESA Sentinel L2A, e.g. NDWI
   - NASA Landsat 7 + 8, e.g. LAI, FAPAR
2. Resources (data and models) used on the fly on the web
   - Not only static datasets served as web services
   - But also computational workflows (e.g. making of account ready…)
3. Distributed system with two different computational architectures:
   - OpenEO - Terrascope
   - ARIES for SEEA Sector Hub of the UNGP – k.LAB tech
     - Link to online data services, including OGC-compliant and STAC catalogues
4. Open Source and openly documented
   - Online open repositories
5. API Networked Economy (virtual credits): single authentication point

# The role of semantics

- Split between resources (data and models) and semantics is neat
- Resources exist online and have a URN, which is by definition unique
  - Can be peer reviewed
  - Have a DOI
  - Have a web page with metadata like in Zenodo

- Semantics is the orchestrator that connects the pieces

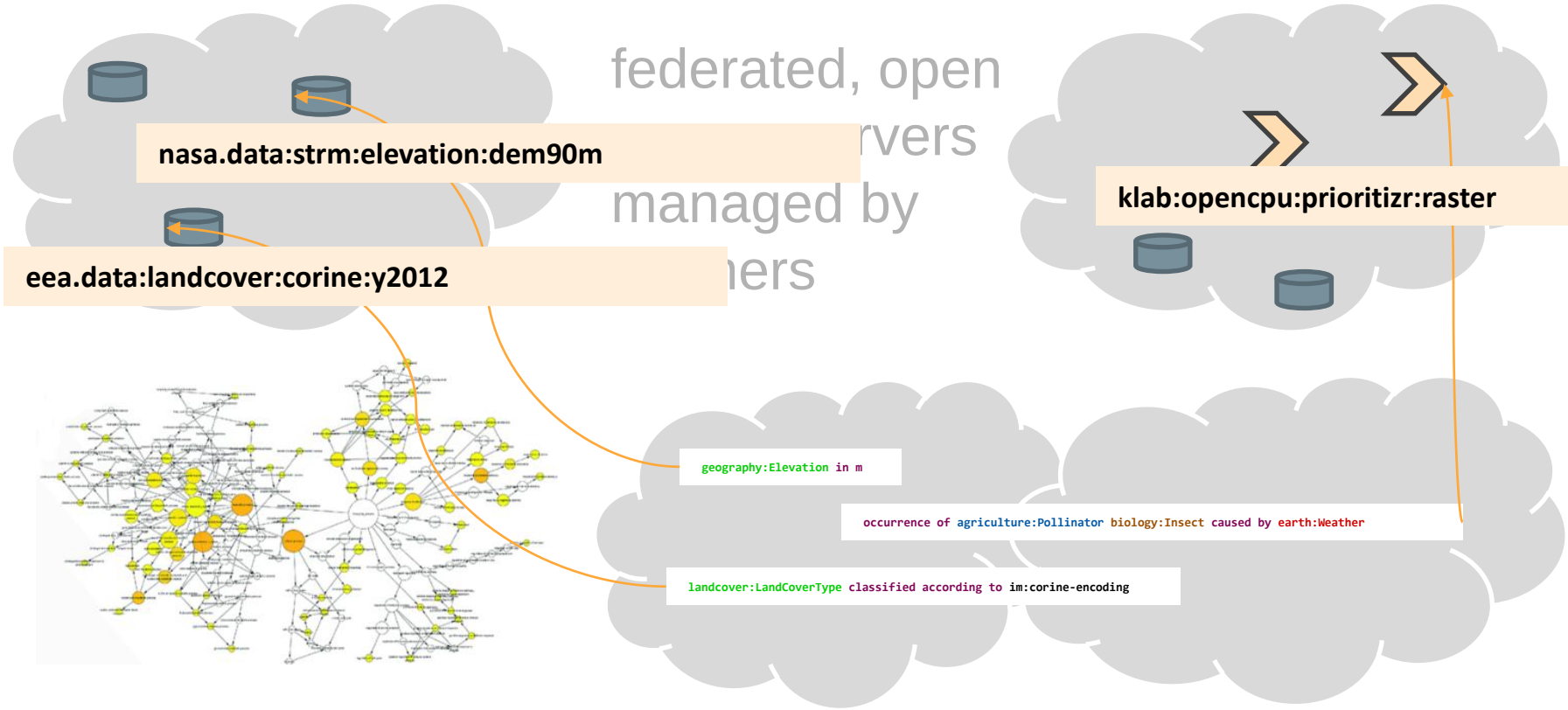
```
/**
 * Global UDP parameters can be set in the URN
 */
model 'local:ferdinando.villa:im.openeo.sandbox:vito.corine.arablemask#output_warp=true'
    as presence of landcover:ArableLand earth:Region;
```

“observing” the **presence of ArableLand Region** concept triggers the execution of the UDP/UDF on the OpenEO server

# A new vision for science

- From “open science” to “deep(ly integrated) science”
- Resources live online independent of their semantic orchestration
- Distributed  and autonomously produced scientific products can be peer reviewed and maintained on the web
- Serve multiple purposes
- Maximum traceability and accountability
- APIs networking between different standards and architectures

# ARIES: a semantics-driven, AI-assisted model and data federation




## RESOURCE layer

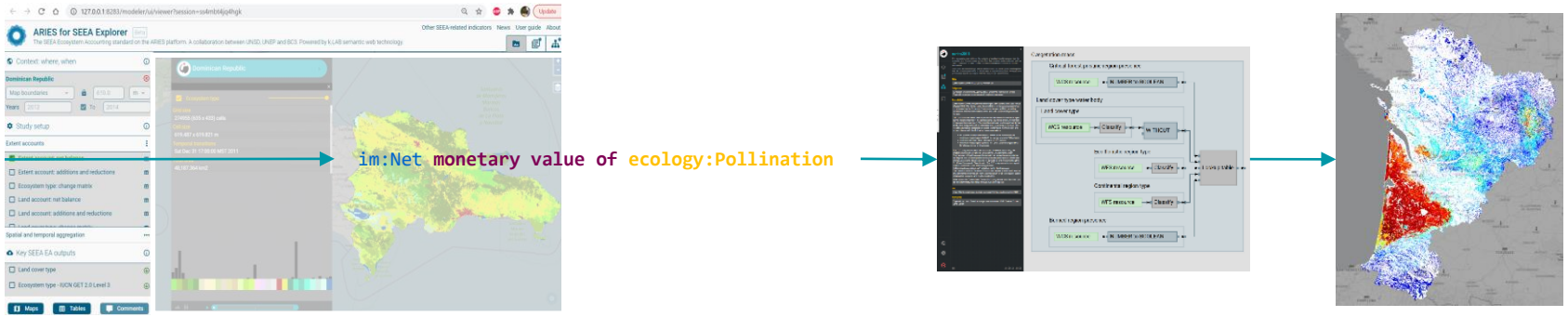
- **Assets** identified by **URNs**
- Include "conventional" data, tadata for
  external services and computational platforms

## SEMANTIC layer

- **Worldview**: shared **concepts** and relationships, communally curated
- **Semantic assets**: associate resource URNs to their meaning in terms of the worldview

## Digital twins (reactivity) layer

- **User queries** ("observe *concept* in *context*") asked through **API** or **applications**
- AI assembles the best-case algorithm to produce reactive **observations**
- **Behaviors** can be specified and triggered



# Breaking new ground

- From open to integrated science: using artificial intelligence for sustainability challenges
  - Collaborative knowledge: machine (AI) + human knowledge = human centric AI
  - Semantics development is the work of knowledge engineers
  - All knowledge is humanly curated

**#Integration: Bridging communities, data & models for intercomparison and reuse**

**#MakeNatureCount: Global understanding of the role of nature for better policies**

**#ScienceBased: Scientific knowledge made accessible and usable by many**

# Technical ToDos and Wishlist

- Extension of OpenEO functionality
  - Extend OpenEO UDP/UDF schema with metadata conventions
  - Support parallel requests
- Transfer of all INCA models to UDP/UDF catalog
  - Currently blocked by missing OpenEO functionalities
- Stability and efficiency improvements
- K.LAB federated system

# Governance Challenges

From the website: “The application is currently only available with restricted access to the Project Team and the Early Adopters, but is targeted to be open and publicly available after the project.”

What do open and public mean?

1. Requires talking about: ESA resources credit system, real costs and/or an alternative setup (openEO endpoint in ARIES for SEEA sector hub?)
2. Marcel can report on the current cost of operations (less than 2,000 €?)
3. Stefano can report on actual usage via online app.

# Demo

- Video: https://www.youtube.com/watch?v=fvChjWO5IN8
- Documentation: https://confluence.integratedmodelling.org/display/AFP

- Point and click UI
- All the complexities are kept under the hood
- All the black boxes are made transparent

- Steps:

1. Choose Spatial and Temporal context and scales
2. Click on the desired account (sections on the left panel)

# Thank you


Ferdinando Villa, Stefano Balbi

BC3 & Ikerbasque

PEOPLE-EA, Athens, 23 May 2024

esa

vito remote sensing

bc3 BASQUE CENTRE FOR CLIMATE CHANGE Klima Aldaketa Ikergai Sustainability, that's it!

Ecounting

UNIVERSITY OF PATRAS ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

Making Science Matter in Policy Making

Where Nature Counts