ARIES

ARtificial Intelligence for Environment & Sustainability

ARIES for SEEA: a web-based application for automatic yet fully-transparent compilation of SEEA Ecosystem Accounts

September 26, 2022

Making Science Matter in Policy-Making Where Nature Counts.

ARIES
ARtificial Intelligence for Environment & Sustainability

Index

- The complexity of environmental modelling
- FAIR principles
- Web-semantic modelling
- What is ARIES for SEEA?
- How are fast, transparent and yet customizable accounts generated?
- How can I access ARIES for SEEA?

ARIES

ARtificial Intelligence for Environment & Sustainability

Environmental modelling: finding Simplicity in Complexity

The complexity of environmental modelling

Data access and manipulation

- Difficulties to access the data
- Time to process the data
- Data storage capacity
- Access to results

The complexity of environmental modelling

Data access and manipulation

Multidisciplinarity

Forest

Environment

Biology

Geography

Ecology

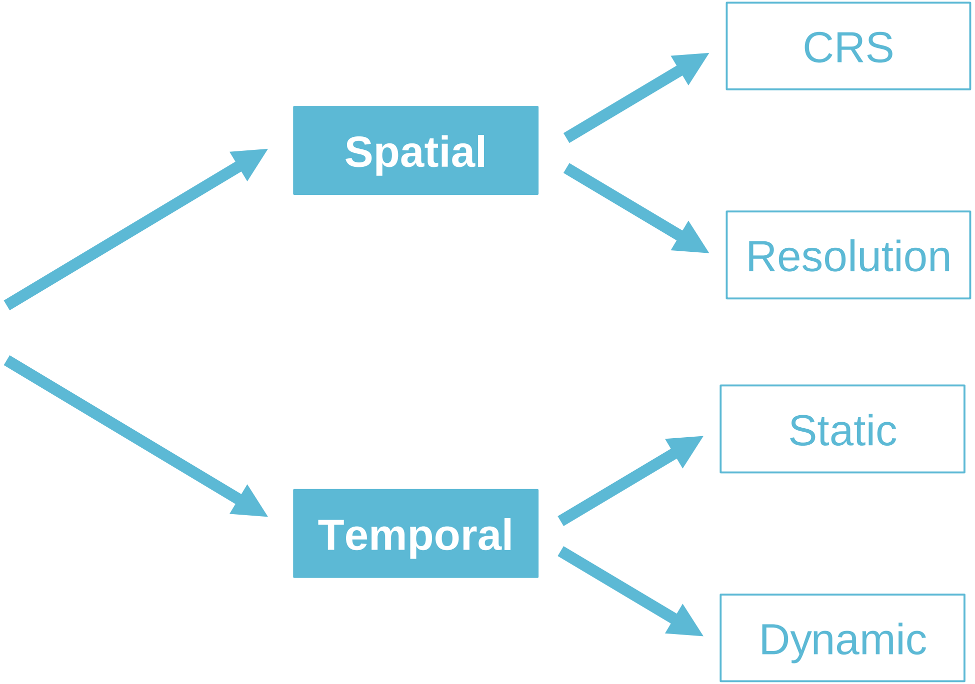
Social Agents

The complexity of modelling

Data access and manipulation

Multidisciplinarity

Different spatial and time scales


Spatial
CRS
Resolution
Temporal
Static
Dynamic

The complexity of modelling

Data access and manipulation

Multidisciplinarity

Different spatial and time scales

Blackbox models

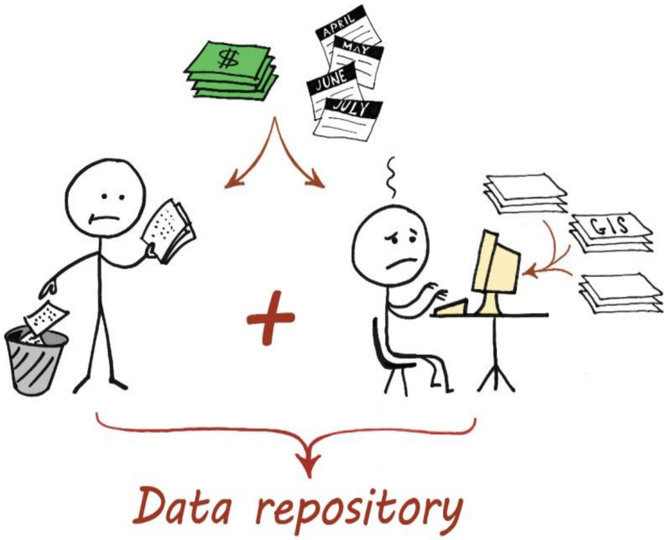

Input
?
Output

the FAIR principles:

One of the most relevant modelling challenges

FAIR principles

Findable	Accessible	Interoperable	Reusable
Open, free and accessible data	Unique identifier that allows to find data quickly and efficiently	Data are organized by thematic areas, but also related interconnected across them	Data (and models) can be applied to other spatial and temporal contexts



Status quo

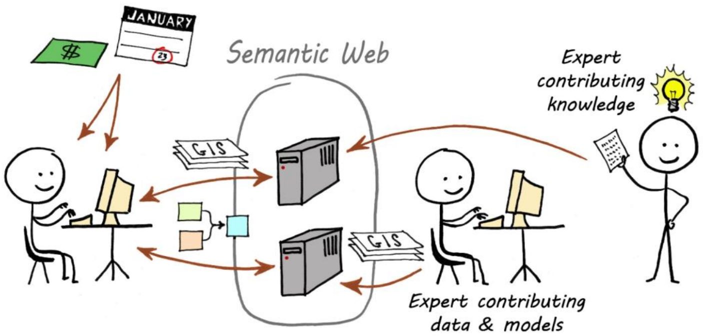
A methodology is often **difficult to implement:**

- Time to obtain and process the data;
- Access and process the data and models;
- Limited computational capacity;
- Limited storage capacity.

Double-efforts are very common.

Everyday **accessibility hindrances**:

- Final users cannot access **results** (many time she's not a technician);
- Technical users cannot access the **models** to re-use, adapt or improve them.



Optimizing the processes through interoperability and the web-semantic

Totally transparent access to:
- Data;
- Methodology;
- Processing steps (the combination of the above);
- Results.

Great storing and computational **capacity**.

Interoperability within data and models.

Access for all users: expert, technical and non-technical.

Collaborative platform to optimize efforts made by different experts in different areas.

“The paradigm behind ARIES is to use technology to build integration and to break up into granular, customizable components, to representing ecosystem services as they are.”

ARIES LEADER FERDINANDO VILLA

At the virtual Expert Forum on SEEA Experimental Ecosystem Accounting (SEEA EEA) 2020

ARIES

ARtificial Intelligence for Environment & Sustainability

Web-semantic modelling

Semantics are concepts

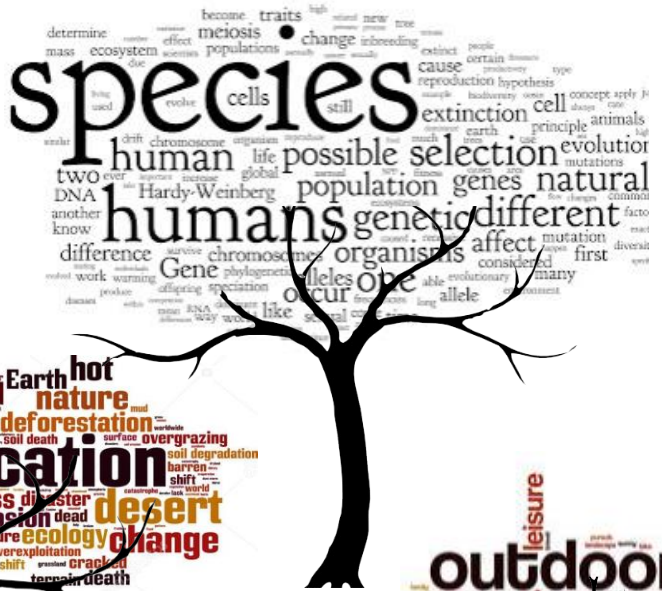
spe
cies

FARM

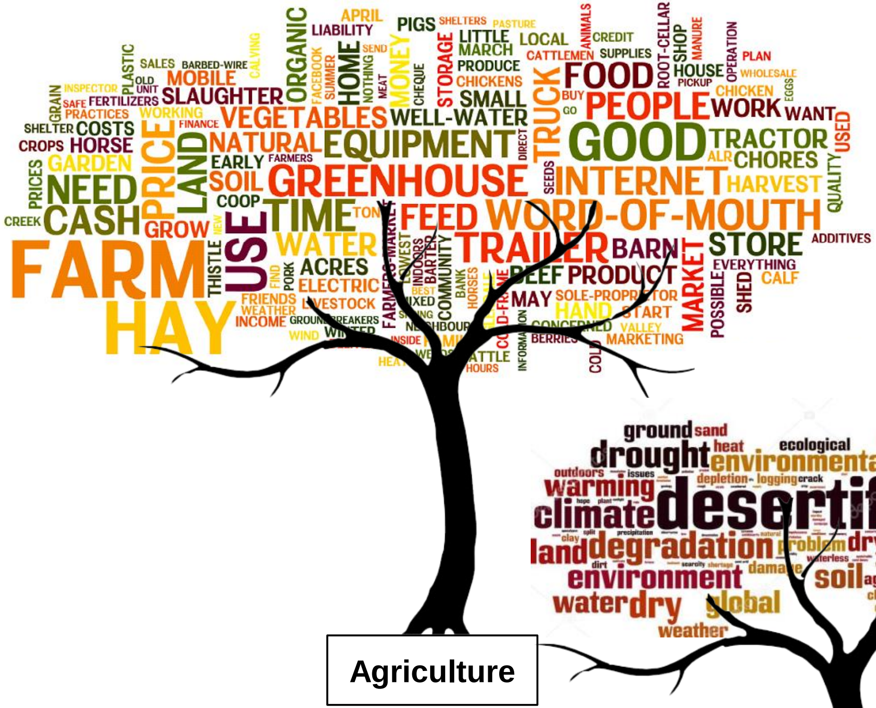
global

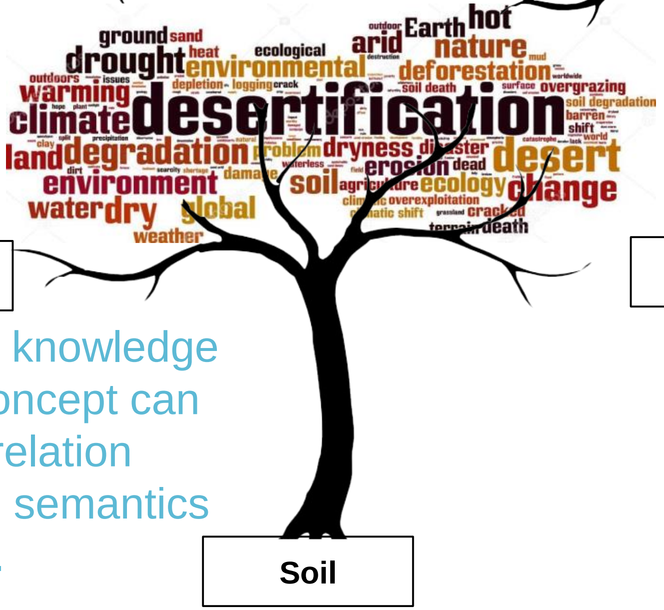
outdoor

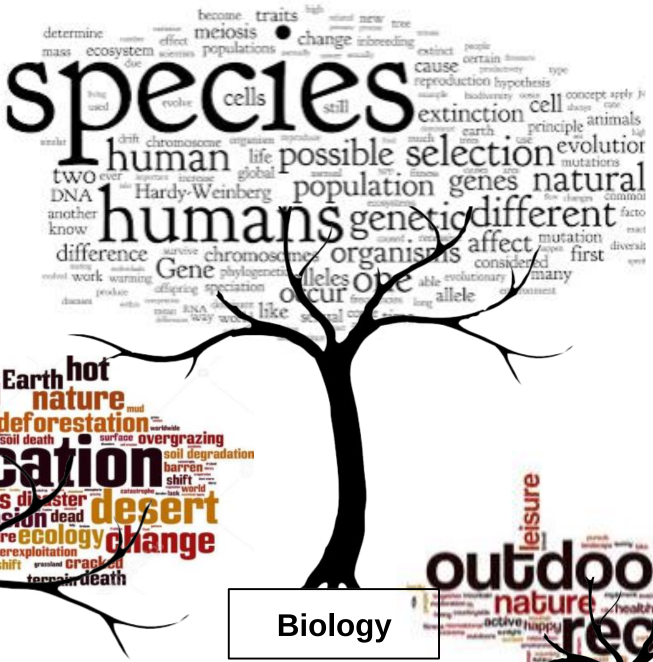
What do they mean and how are they interconnected with each other?

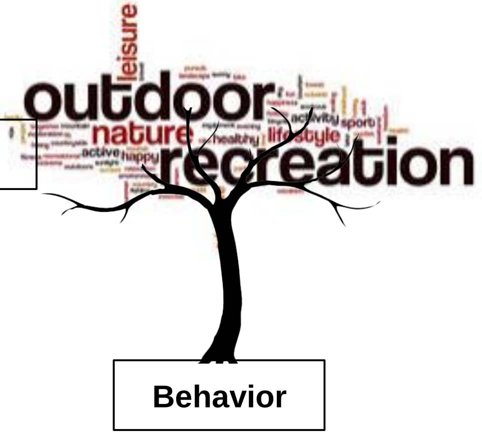


The definition and the relation with other concepts are defined in the **ontology**









Depending on the area of knowledge applied, each semantic concept can change its definition and relation (ontology), for this reason semantics are grouped by **domains**.

Growth

im

abstract process Growth
"Any endogeneous transformation happening in a system its internal structure."
is IntransitiveProcess
has children
(**Growth**
"Positive or negative growth makes a system acquire a larger size or functional throughput"
creates Collapse);

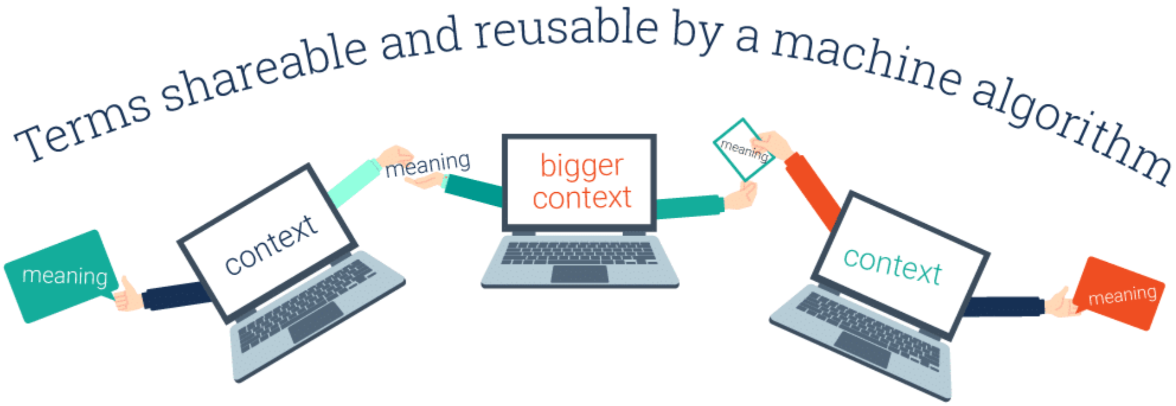
Ecology

process Growth
"Growth in ecology refers to population growth."
is im:Growth **within** Population

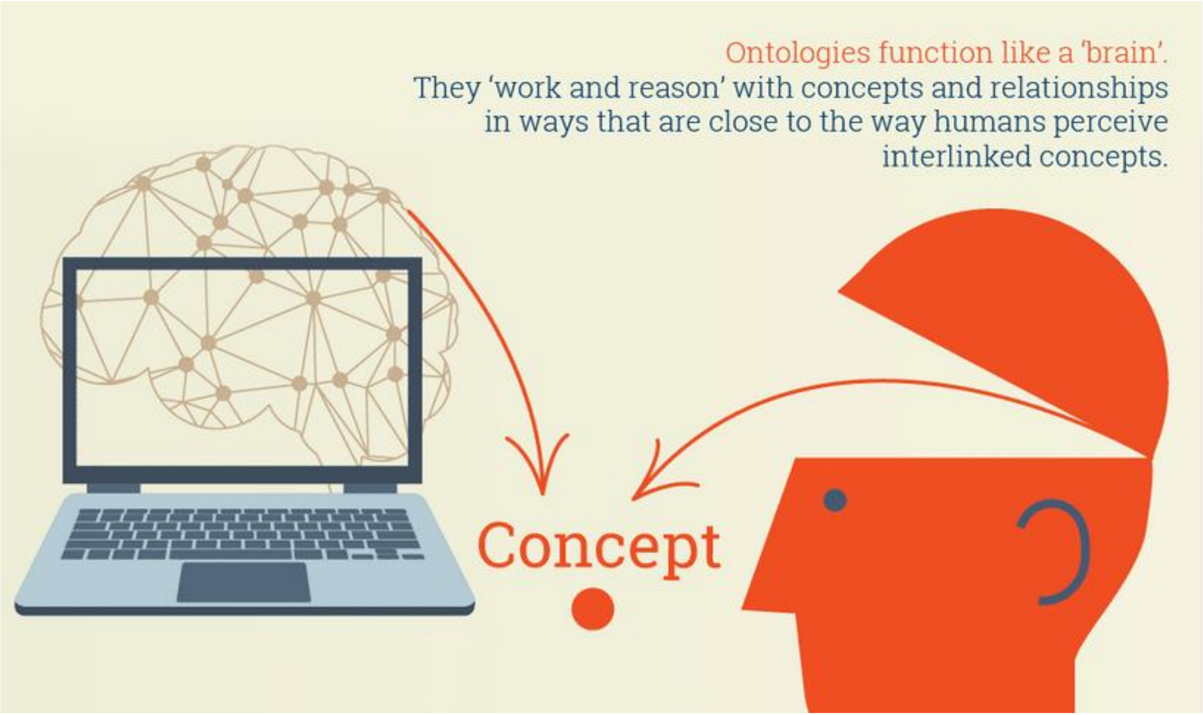
Biology

process Growth
"Biological growth only happen during life and affects the biomass of an individual."
is im:Growth
affects Biomass **within** Individual
applies to Life;

Well-defined concepts, are written in our language but also machine-readable. This allows is the ability of the artificial intelligence to think automatically or semi-automatically **(the machine reasoning).**
Moreover, this enables the information to be exchanged with other machines and also adjusted to different contexts.
Since it's an **open and collaborative** system, this is **constantly improving**



N. Guarino, D. Oberle and S. Staab, "What Is an Ontology?" In: S. Staab and R. Studer, Eds., Handbook on Ontologies, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 1-17.



N. Guarino, D. Oberle and S. Staab, "What Is an Ontology?" In: S. Staab and R. Studer, Eds., Handbook on Ontologies, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 1-17.

ARIES
ARtificial Intelligence for Environment & Sustainability

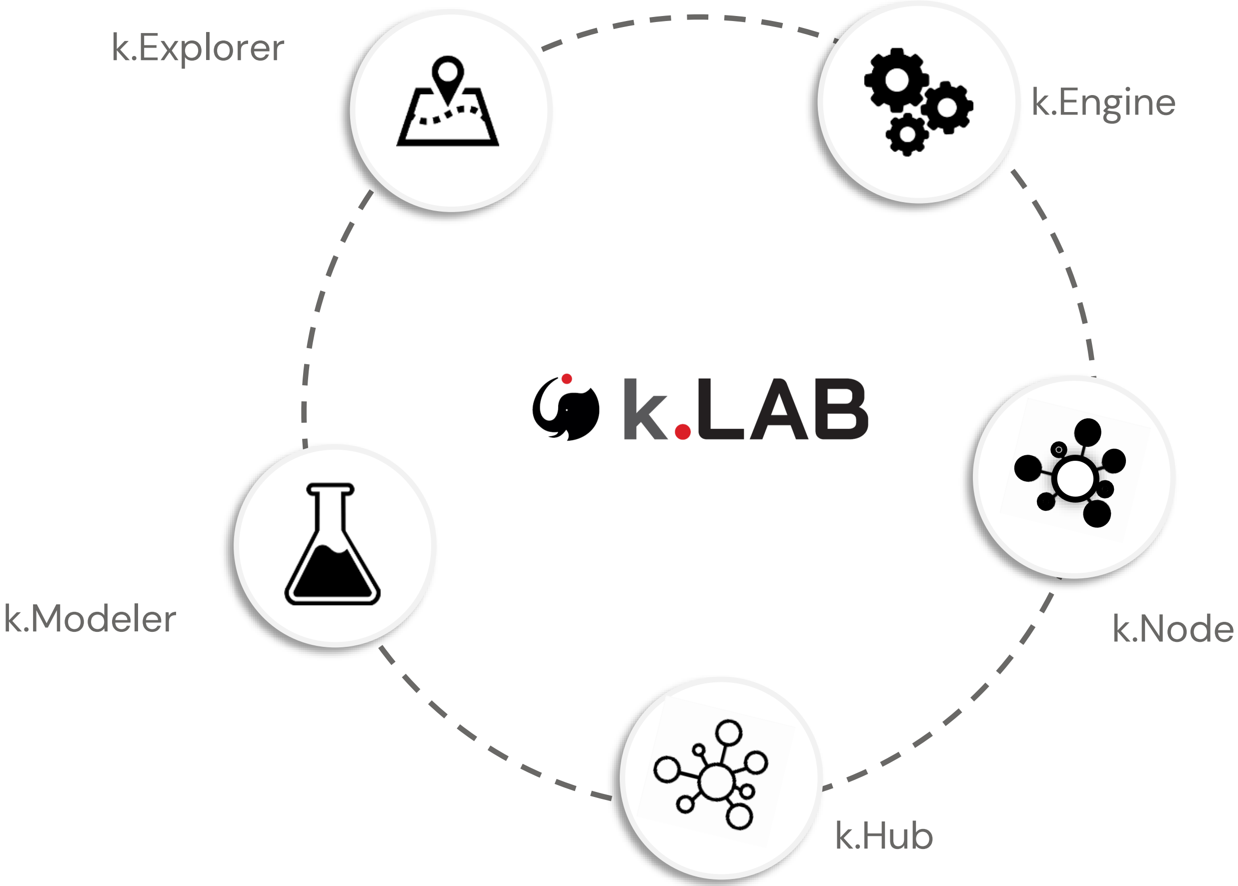
k.LAB

The knoweldge laboratory

A semantic web for sustainability: revolutionizing how we write, find, link and reuse data and models.

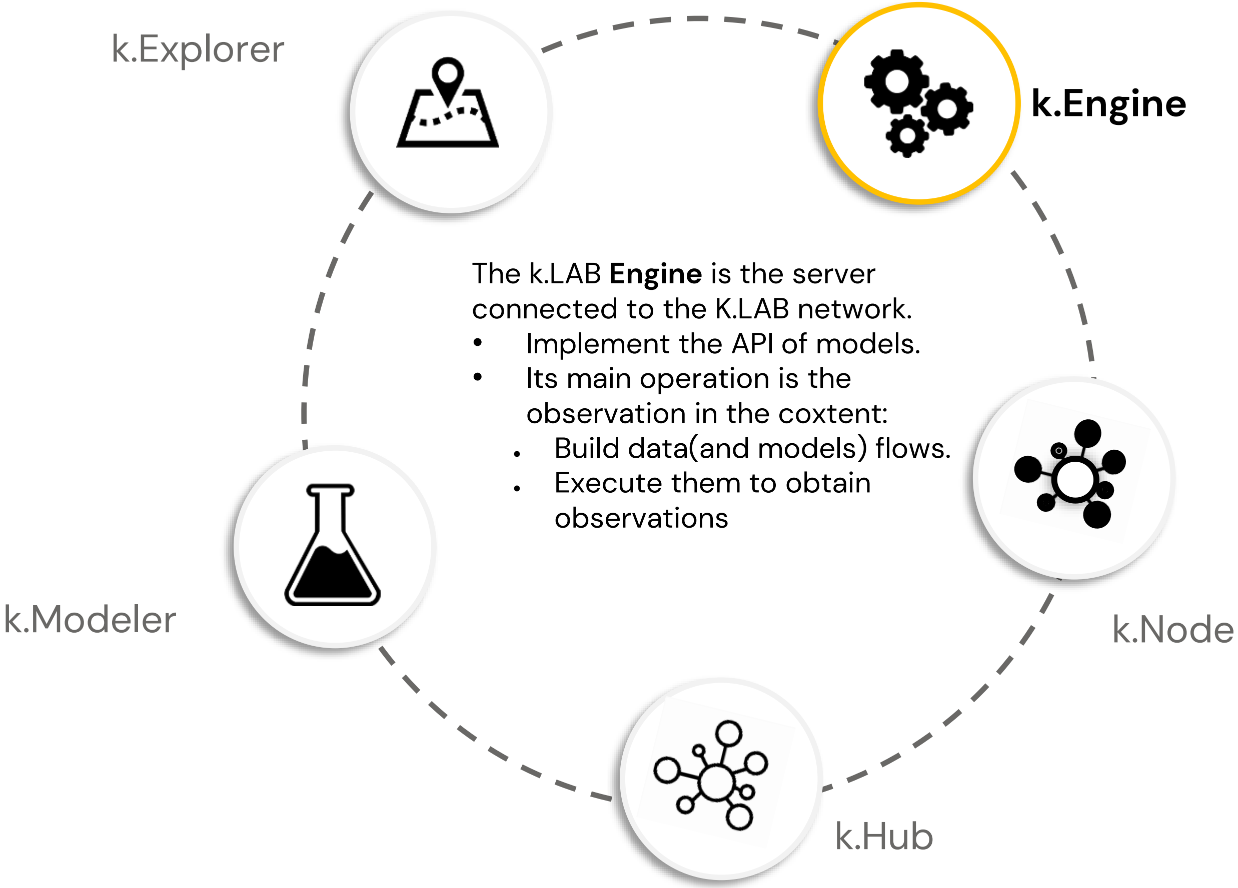
The k.LAB platform supports observations made in the semantic web

Include its own semantic language, **k.IM**, used to program in this context



The k.LAB platform supports observations made in the semantic web

Include its own semantic language, **k.IM**, used to program in this context

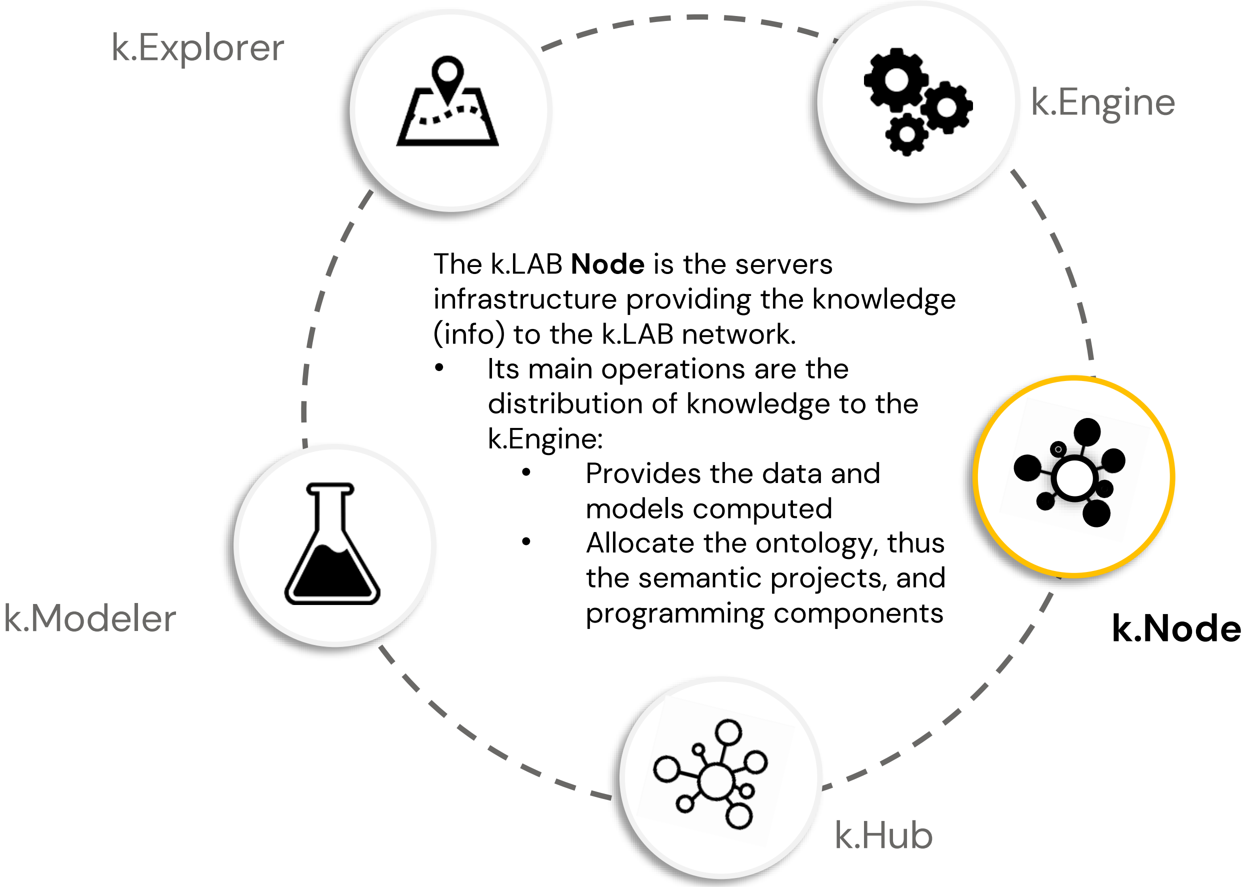


The k.LAB **Engine** is the server connected to the K.LAB network.

- Implement the API of models.
- Its main operation is the observation in the coxtent:
  - Build data(and models) flows.
  - Execute them to obtain observations

The k.LAB platform supports observations made in the semantic web

Include its own semantic language, **k.IM**, used to program in this context

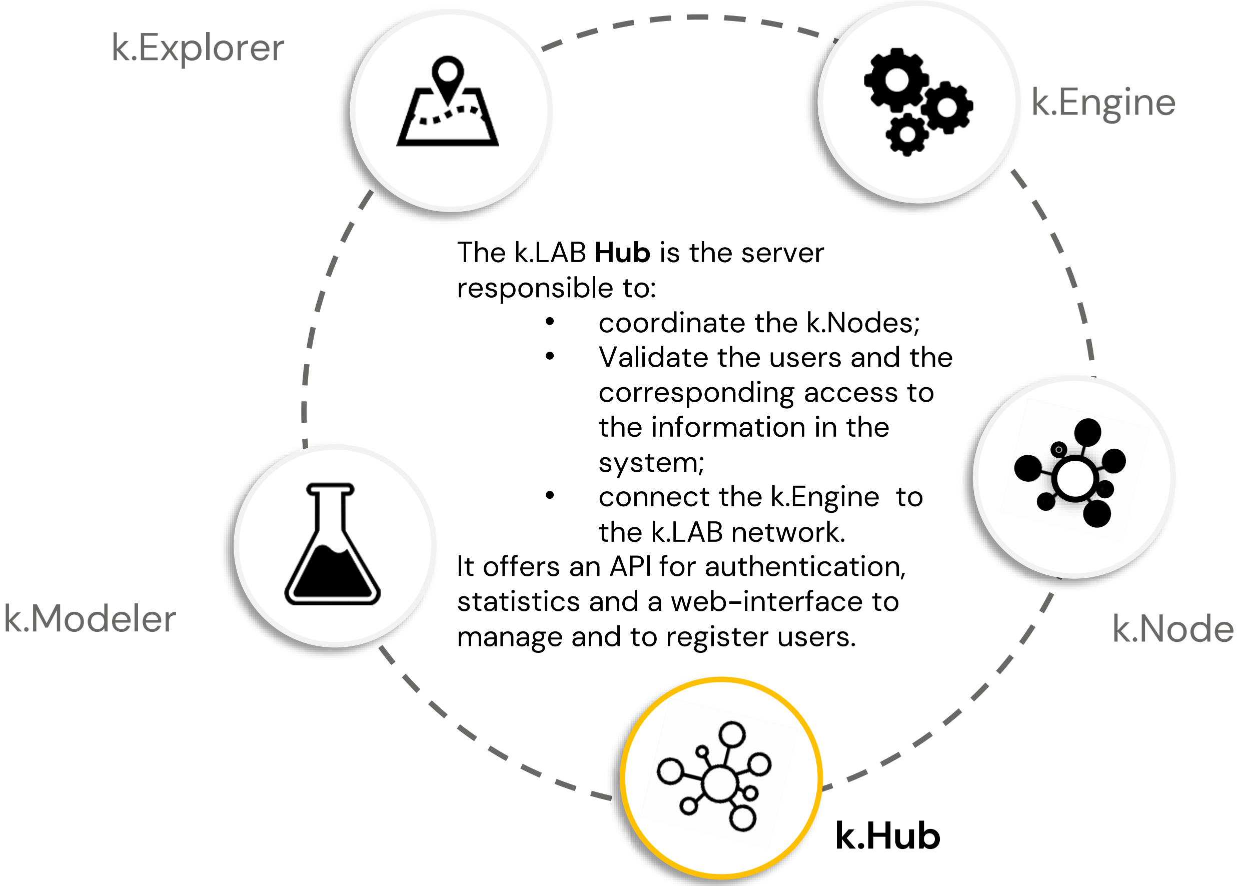


The k.LAB **Node** is the servers infrastructure providing the knowledge (info) to the k.LAB network.

- Its main operations are the distribution of knowledge to the k.Engine:
  - Provides the data and models computed
  - Allocate the ontology, thus the semantic projects, and programming components

The k.LAB platform supports observations made in the semantic web

Include its own semantic language, **k.IM**, used to program in this context



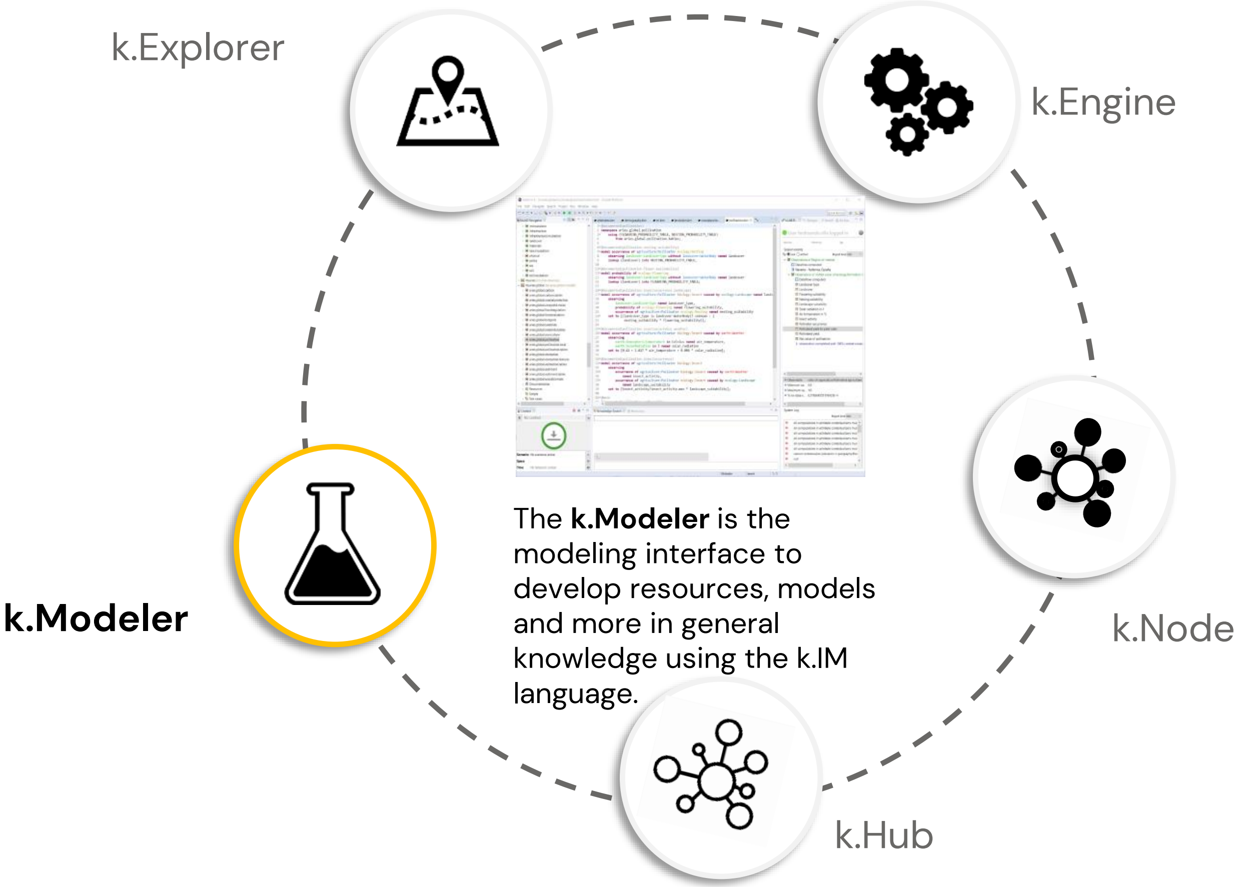
The k.LAB **Hub** is the server responsible to:

- coordinate the k.Nodes;
- Validate the users and the corresponding access to the information in the system;
- connect the k.Engine to the k.LAB network.

It offers an API for authentication, statistics and a web-interface to manage and to register users.

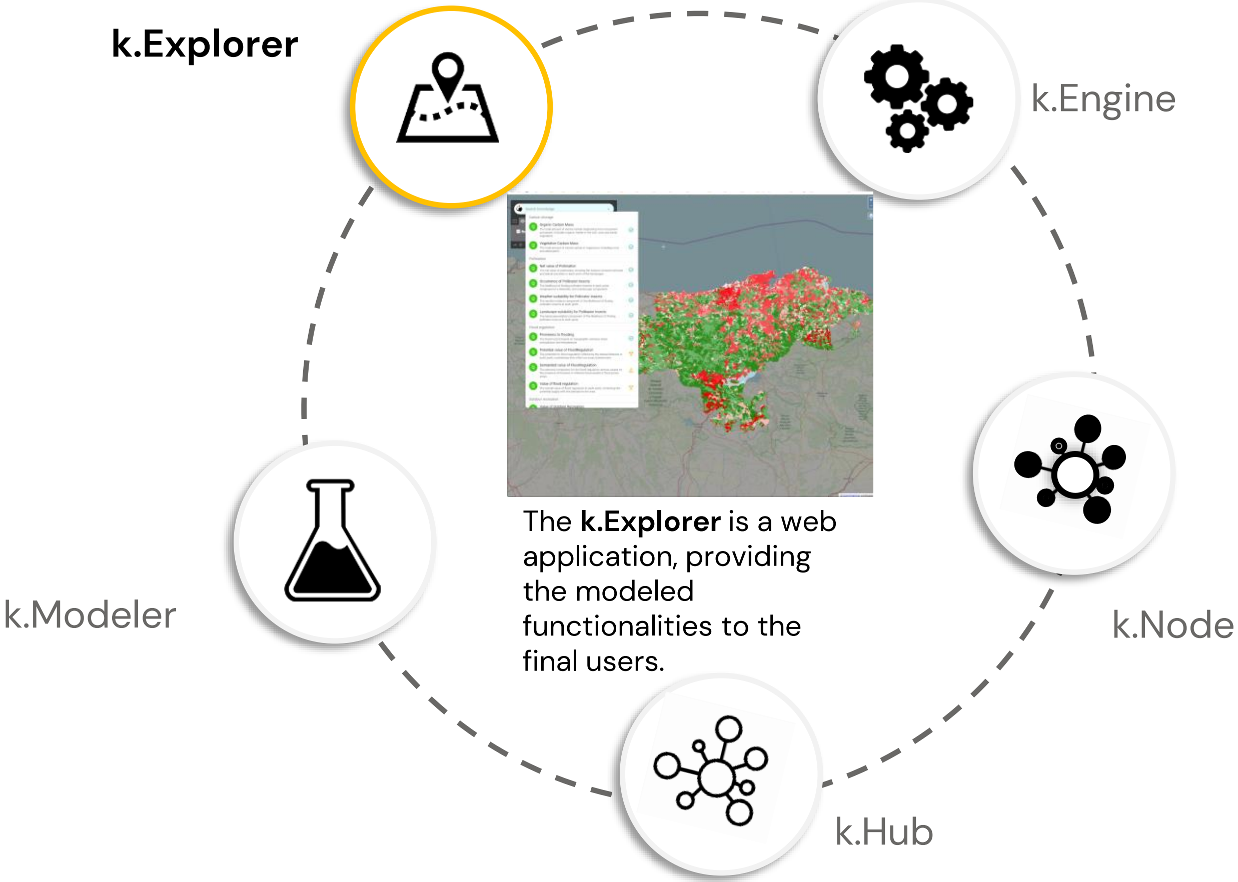
The k.LAB platform supports observations made in the semantic web

Include its own semantic language, **k.IM**, used to program in this context



The k.LAB platform supports observations made in the semantic web

Include its own semantic language, **k.IM**, used to program in this context



The **k.Explorer** is a web application, providing the modeled functionalities to the final users.

How do you model in k.LAB?

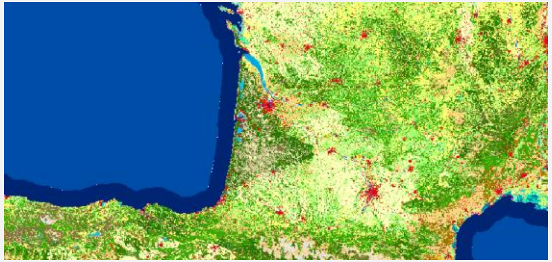
Define the context and pose a query

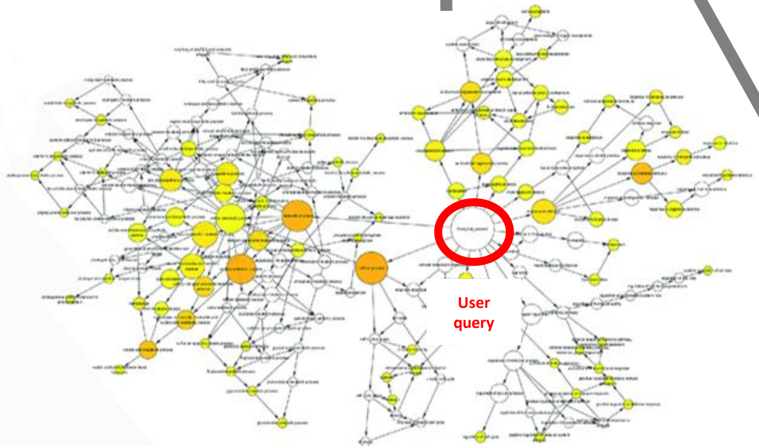
Context and query

Models semantically annotated

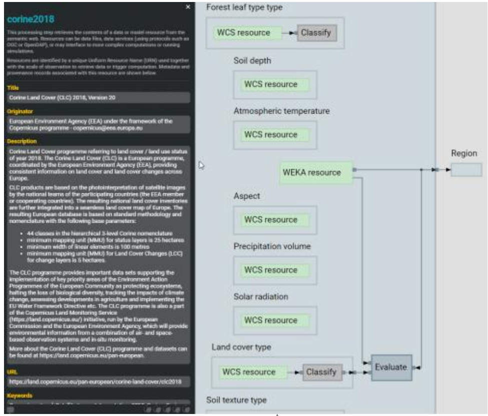
```
model occurrence of agriculture:Pollinator
biology:Insect caused by earth:Weather
        observing
earth:AtmosphericTemperature in Celsius named
air temperature,
earth:SolarRadiation in J named solar_radiation
        set to [0.62 + 1.027 * air temperature
+ 0.006 * solar_radiation];
```

Data semantically annotated





Semantic relations between data and models > Computations



k.LAB Contextualization report

Computed at Mon Oct 14 15:37:51 CEST 2019

1 Introduction

Global supply-demand ecosystem service models for ARIES

The baseline modeling of ecosystem services (ES) supply and demand is conducted in ARIES using a suite of logical statements, data, and models that are used when not enough information is available to build more detailed, dynamic flow models. The models built by ARIES using such statements have, in general, similar resolution and conceptual detail as those available in other ES assessment toolkits such as InVEST or ESTIMAP. They can run anywhere in the world with no user input (using global data and parameters), while offering the option to easily customize models with context-specific data and parameters. This approach enables rapid ES quantification, as models are automatically adapted to the application context and run using the best available data for the context. The models use publicly available global- and continental-scale data as defaults. Advanced users can modify data input requirements, model parameters, or entire model structures to capitalize on high-resolution data and context-specific model formulations.

##Biomass models

The crop yield model and the wood biomass model are machine learning models that replicate outcomes of other models using a less demandig set of predictors which are available at as global public dat layers. Machine learning is enabled in ARIES through the Weka library of algorithms (https://www.cs.waikato.ac.nz/ml/weka/). These two models use the BayesNet classifier which learns Bayesian Networks. These models are still experimental as the were learned in geographical context between the Atlantic and the Mediterranean coast in Europe, so their extrapolation beyond that should be carefully considered.

2 Methods

The wood biomass supply model simulates the ouput of the EFI (European Forestry Institute) model of wood production in European forests developed by Verkerk et al. (2015) Reference 1. Estimates are based on a Bayesian Network model which uses the following variables as predictors:

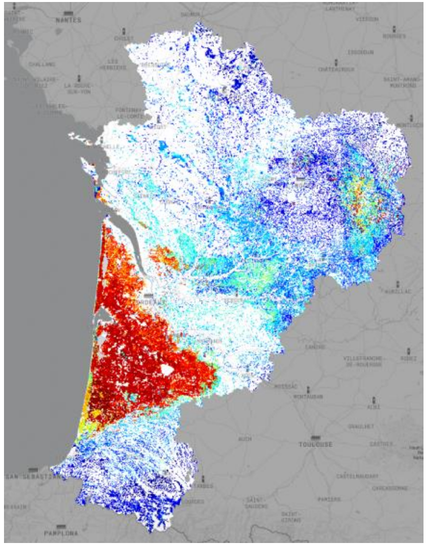
-Ecological parameters: (1) Leaf Type, (2) Enhanced VegetationIndex, (3) Normalized Difference Vegetation Index;

-Climatic conditions: (4) Precipitation volume, (5) Atmospheric temperature, (6) Solar radiation;

-Topographic characteristics: (7) Elevation, (8) Slope, (9) Aspect;

-Soil characteristics: (10) Soil depth, (11) Soil texture;

3 Results



ARIES

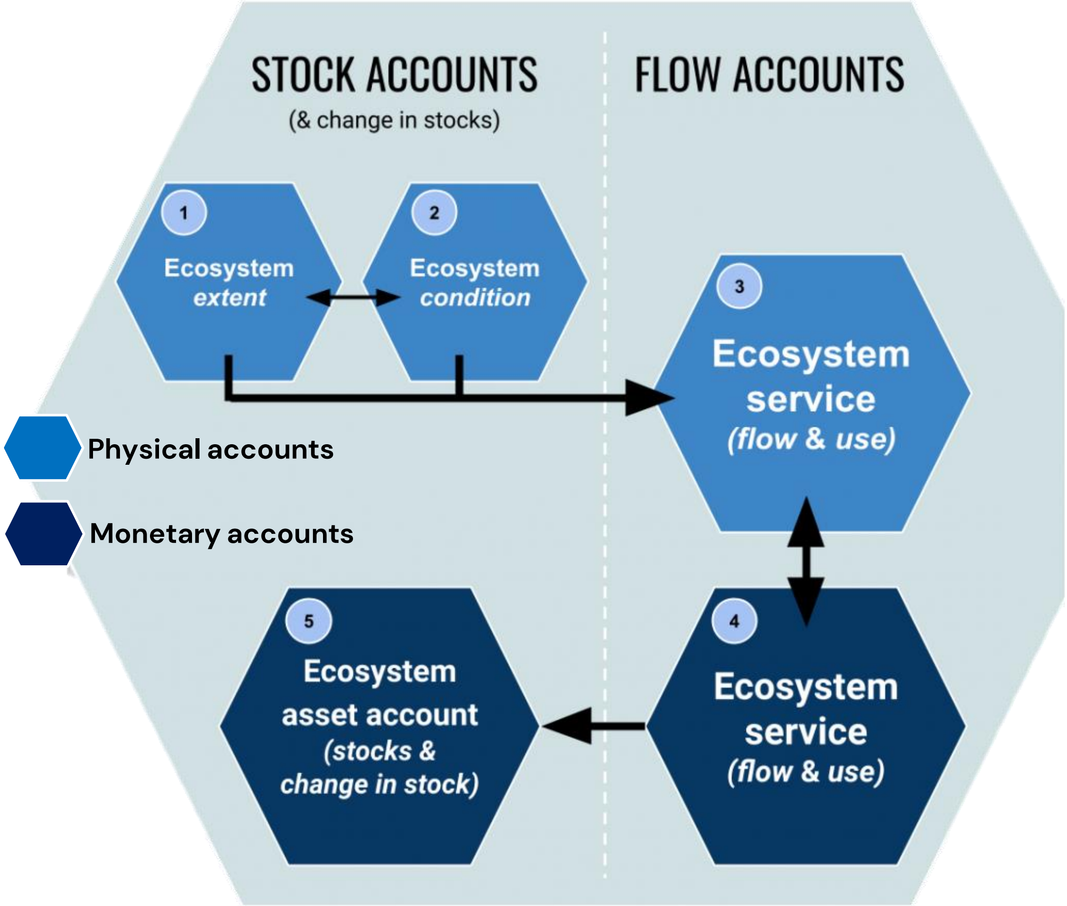
ARtificial Intelligence for Environment & Sustainability

ARIES for SEEA:

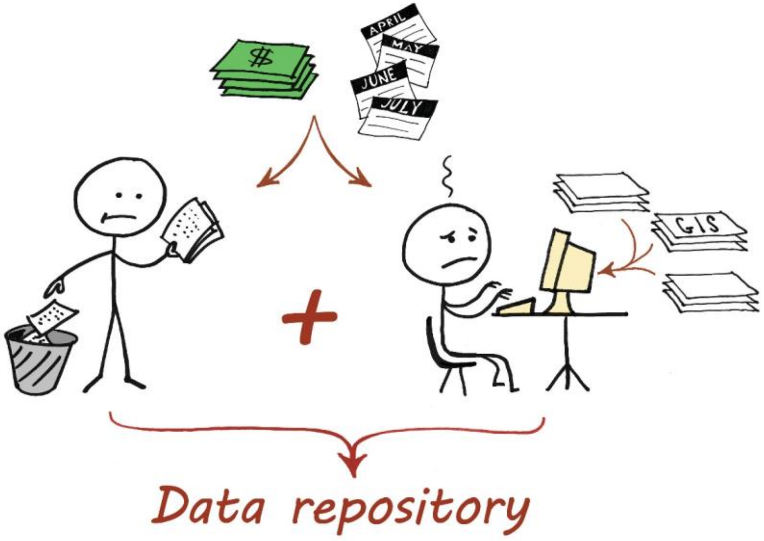
AI Solution:

ARIES for SEEA

Data & Models architecture for knowledge integration

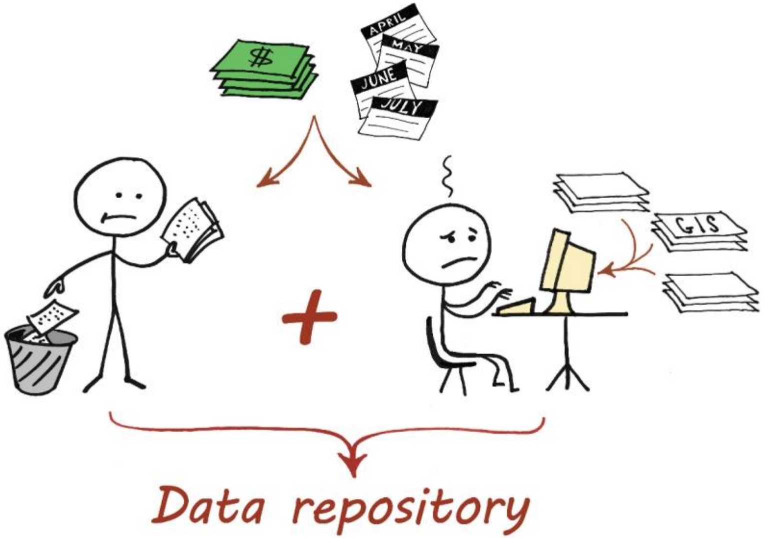


Does ecosystem accounting always need to be painstakingly slow?



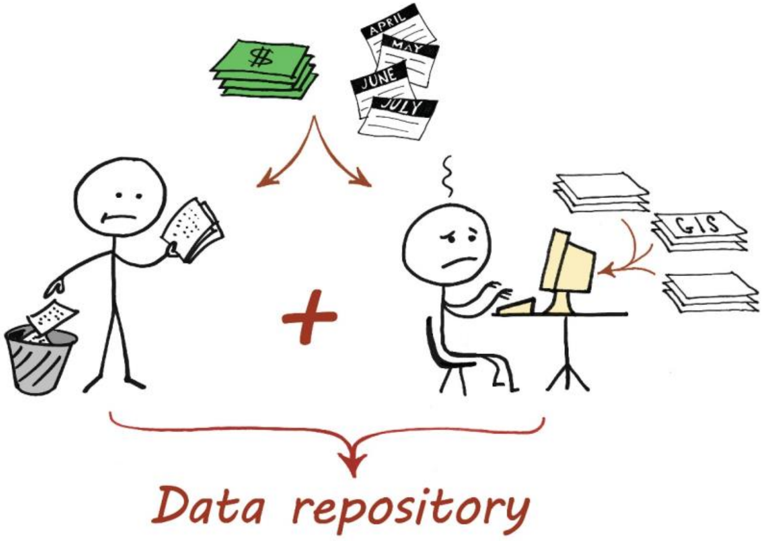
Status quo

Does ecosystem accounting always need to be painstakingly slow?

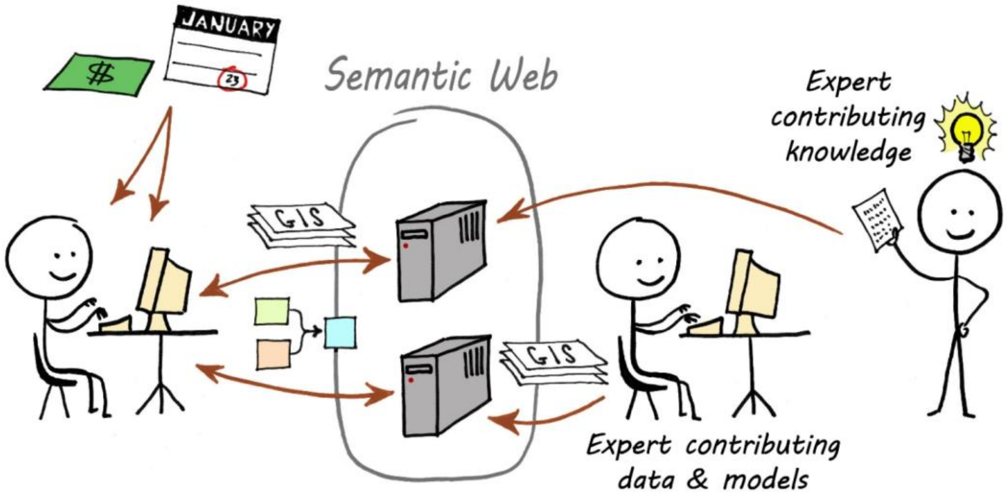


Status quo

Does ecosystem accounting always need to be painstakingly slow?



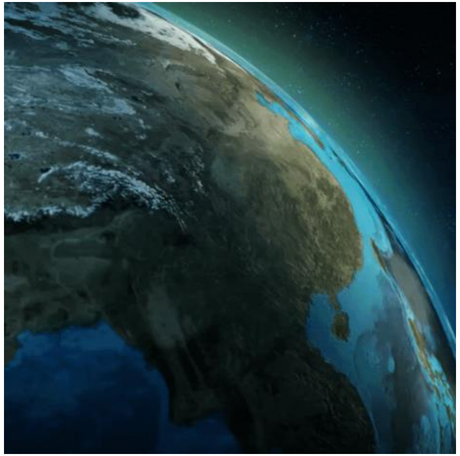
Status quo



Linked, web-based collaborative modeling

ARIES for SEEA
for rapid, standardized account creation

- Global, customizable models approach enables:
  ⟹ **SEEA EA compilation**[1]
- **Faster & easier** to learn than other biophysical modeling approaches
- **Automate production** of accounting tables, maps & reports
- Support adoption of SEEA EA providing an **easy-to-use application**
- Infrastructure for the SEEA community to **share & reuse** interoperable data & models



1: anywhere on the earth with global data,
but models can be improved using local data (and models) where available.

Key building blocks for interoperability

1. SEMANTICS: a flexible, shareable, easy-to-learn **language** to describe scientific observations

2. OPEN, LINKABLE DATA: enabling access & publishing of semantically annotated data

3. OPEN, LINKABLE MODELS: open, accurate, "Wikipedia-like" sharing and linking of models

The key is a technology built for INTEROPERABILITY, developed on FAIR[1] data/models principles

Semantic web modelling

ARIES is connected to a library in which each component of a model is break down into in **modular** components, which allows and greatly simplify the use (and reuse) or the substitution of a certain input in the model composition when a more appropriate element becomes available in the system (i.e. local vs global data)

AI: machine reasoning

As each component of a model or data is connected to a concept through its semantic, the best combination of data and models available in the system is **automatically integrated** to answer the query posed by user (methodological, spatial and temporal dimensions are all considered).

Which algorithms are used to prioritize resources?

A de-centralised system with real democratic access to the knowledge produced

The system is built to can interconnect information hosted on a network of individual nodes (k.IM), based on open-source software language and always free to use for non-profit purposes. The language used to model and integrate data is **universal** because readable by machines as well as humans; being this very similar to English, it is also very **intuitive** for non-expert modelers. As more data and models are integrated in the system, the quality and the variety of those increases, not with a simple additional but a **multiplicative effects.** Each expert can contribute with new knowledge, and leverage on the knowledge created by others. This allows to find better answer to the questions posed to the system.
ARIES also generates very detailed report(s) and a visual diagram of the data and model(s) used to guarantee **full transparency and traceability** of each individual result.

1: **F**indable, **A**ccessible, **I**nteroperable and **R**eusable data and models

Roles of key stakeholders

- **Data providers** (NSOs, science agencies, academic scientists): agree on & provide data using common formats & hosting protocols (e.g., OGC[1] standards for spatial explicit data, SDMX[2] for tabular data)

- **Modelers** (science agencies, academic scientists): use modeling practices that will make models more easily linked & documented (more modular, less monolithic); use community consensus semantics

- **NSOs & other institutions** (NSOs, space/mapping agencies, GEO initiatives, large academic collaborations): endorse and maintain interoperable data & models over the long term.

- Publication Ethic: giving credit where credit is due - **Credit** for data and models provider(s) in ARIES

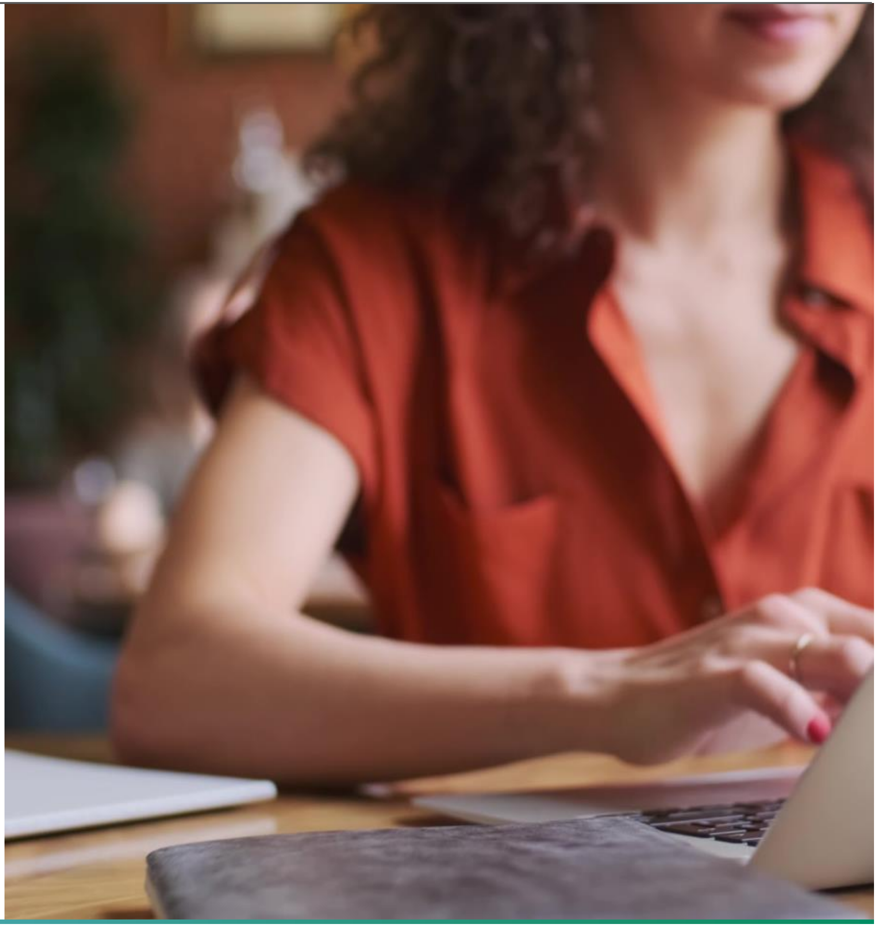
1. Open Geospatial Consortium – more info
2. Statistical Data and Metadata eXchange – more info

Access the application

1. The first step is to **register** in the Integrated Modelling hub
2. Once created a profile, **access the link** to launch the application from your browser (or download the Control Center – the software for modelers, and install it on your engine)
3. Use the intuitive user-interface to **compile account(s) everywhere on earth**

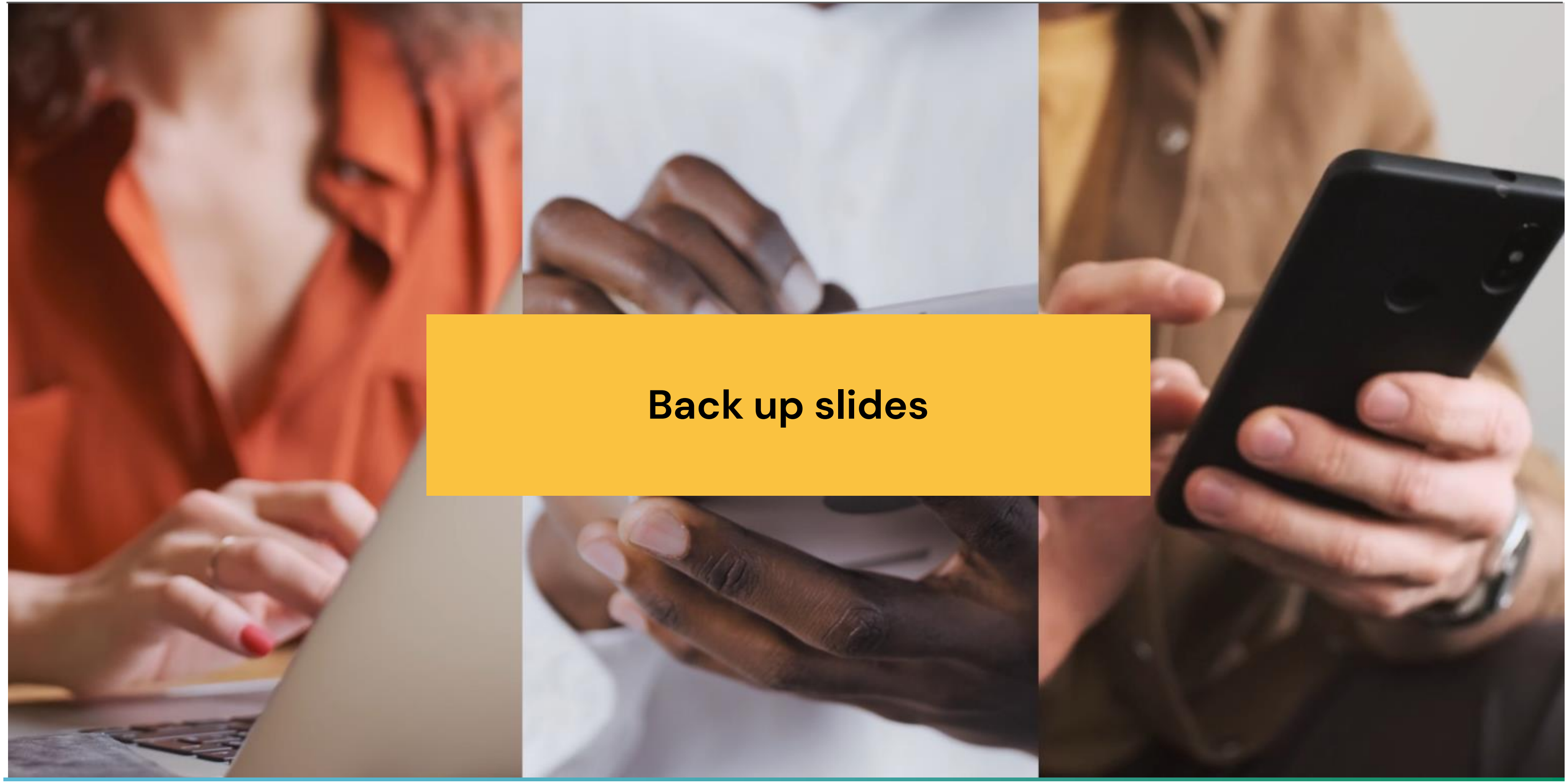
Useful links to explore

1. ARIES for SEEA explorer
2. Registration in the IM hub
3. Technical note
4. YouTube channel
5. Write us for support at support@integratedmodelling.org or for if you're interested to join our modelling journey at aries@integratedmodelling.org



ARIES
ARtificial Intelligence for Environment & Sustainability

Thank you!



Back up slides

Machine reasoning: How do can a machine pick the "best" data/model under which circumstances?

Initial prioritization, adjustable by advanced users:

1. Lexical scope (how "close" are the data/model to the namespace, project, within k.LAB repositories);
2. Trait concordance (shared attributes with concept requested);
3. Scale coverage (data with more complete coverage chosen preferentially);
4. Scale specificity (local models chosen over national, over global);
5. Inherency (models specified for location/scale-specific use chosen over generalized models);
6. Subjective concordance (user-specified metadata & weightings);
7. Evidence (data models chosen over computed models)
8. Reliability (human input that affects the reliability of a source of information)