ISU 2018: Introduction to ARIES and k.LAB

BASQUE CENTRE FOR CLIMATE CHANGE
Klima Aldaketa Ikergai

International Spring University on Ecosystem Services Modeling

A partnership for shared, distributed, collaborative modelling

SEMANTICS for data and models

- Maintenance of the core conceptualization (subjects, processes…)
- Maintenance and delivery of the shared worldview for cross-domain communication

OPEN SOURCE SOFTWARE

- User-end (modelers and end users)
- Server technology (institutions)
- Developer team and user support

INTEGRATED MODELING INFRASTRUCTURE

- Assembly of models from networked data and model components
- Accurate coupled human-natural system representations

COLLABORATIVE MODELING

- Interoperable data and models
- Direct support of partner projects
- International Spring University since 2013

APPLICATIONS

- Ecosystem services assessment (ARIES)
- Food and other environmental securities
- Integrating hydrology, primary production, nutrients with agent models to best represent SES.

The challenge of data/model integration and reuse

Scientists in the past collected data in notebooks. In the digital age, we want scientific data and models to be FAIR - [Findable, Accessible, Interoperable, and Reusable], to ensure their maximum value.

A fully connected information landscape using open, safe, accurate, "Wikipedia-like" sharing and linking of models can enable data-intensive science for decision making on a scale yet unimagined:

1. **reuse** the abundance of data and specialized knowledge available and needed to analyse social and natural processes (and their interactions)
2. **avoid** the risk of **fragmentation** hidden in the use of ad-hoc (or no) semantics to describe data
3. enable **simple user workflows** in modelling, supporting **direct** questions like: What is the social dynamics of water in basin X? How does switching to crop Y affect rural food security in region Z?

Where are we along this path in 2018?

Using and reusing data: The state of the art

1. Distributed access to datasets over the web (OGC, OpenDAP, …)
2. Linked Open Data paradigm: open standards, each artifact is coupled with a URI pointing to its “meaning”.
3. Problem: the meaning *differs for each observer* - unless semantics is coherent across domains, uses and goals.
4. If it’s not consistent, it’s not FAIR

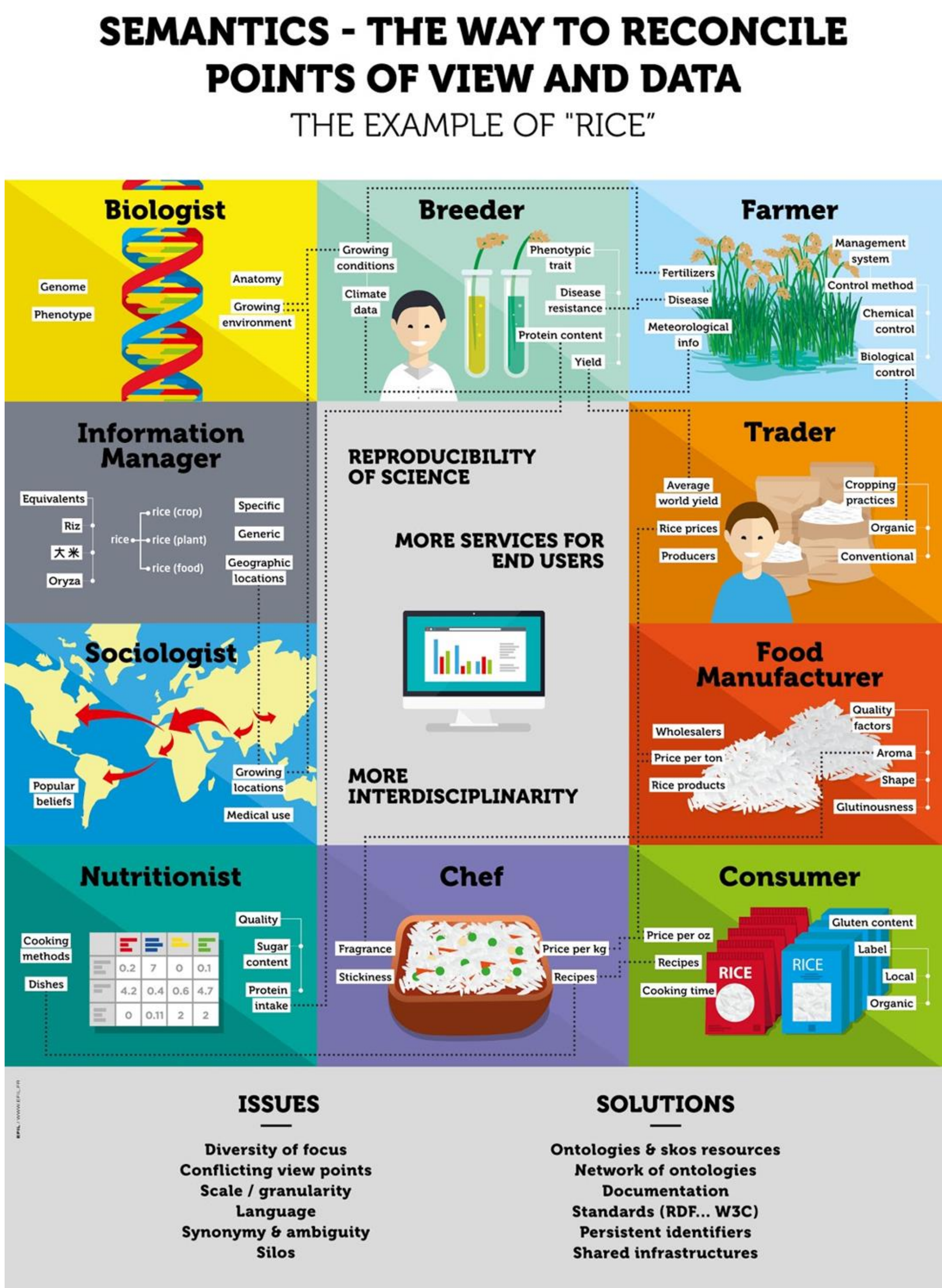


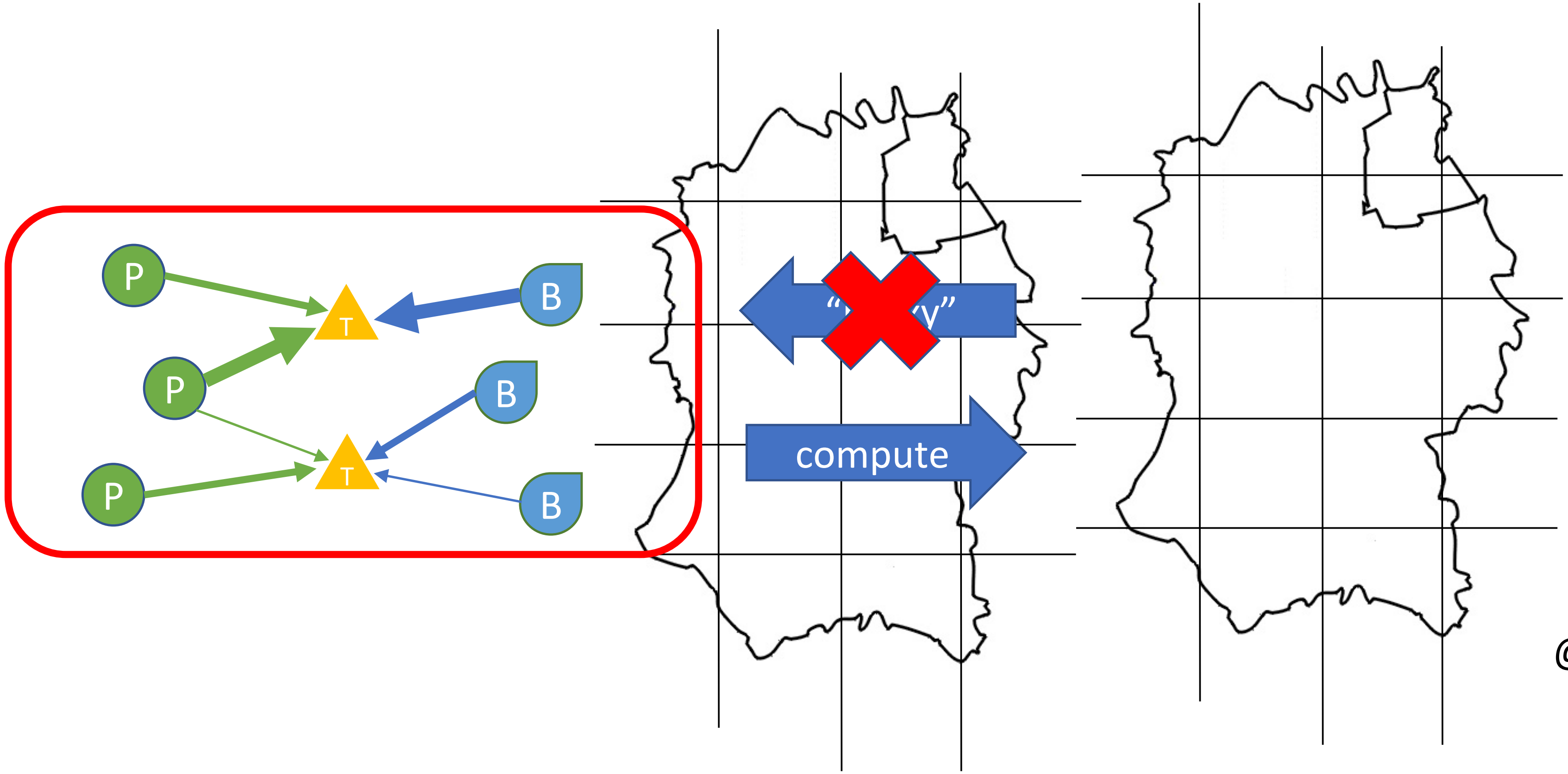
Image credits: INRA, AgriSemantics RDA working group

Reusing models

- Modeling paradigms represent different “metaphors” adopted during model design:
  - process-based vs. agent-based
  - stochastic/probabilistic vs.deterministic models
  - spatial vs. non-spatial, raster/vector, continuous vs. discrete time, etc.
- It remains **difficult to mix and match models incarnating different paradigms** across the lifecycle of an application.
- Often, complex problems are handled with one paradigm that fits some components but must be “tricked” to handle the rest.
- As a result models are still brittle **monoliths**, hard to disassemble and reassemble.
- Integrating architectures (OpenMI &C.) only handle the technical aspects of integration, addressing only a subset of the problem.

A case in point: accounting for human-natural interactions

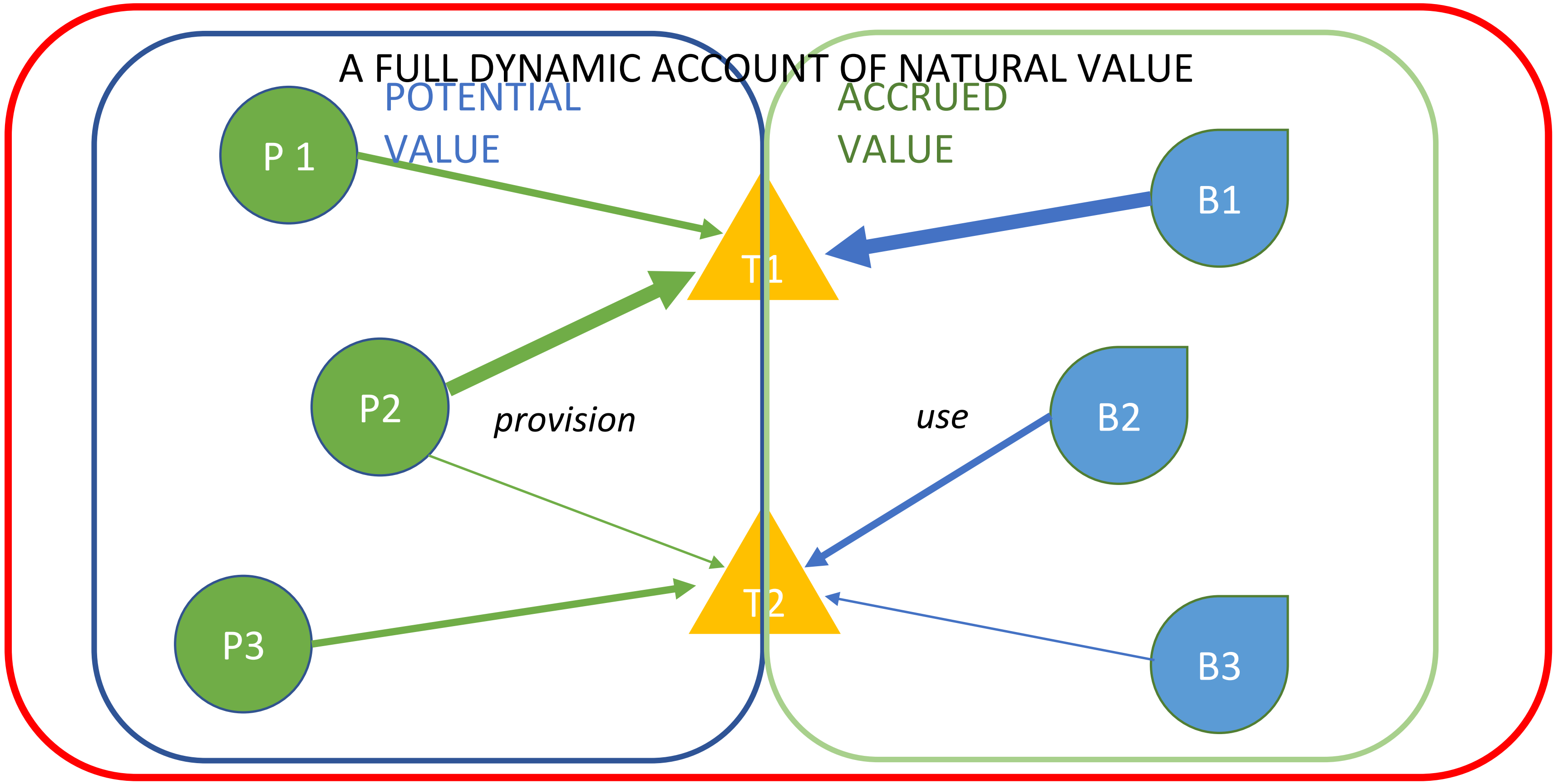
- We know the limitations of “proxy” models – and it’s not because of decision makers.
- Still, building accurate models of the *true* system is hard – and impossible in rapid assessments



@ time = t

Adaptive, assisted system characterization

Driven by semantics and by *roles*, supporting a specific view of physical phenomena without introducing ambiguities



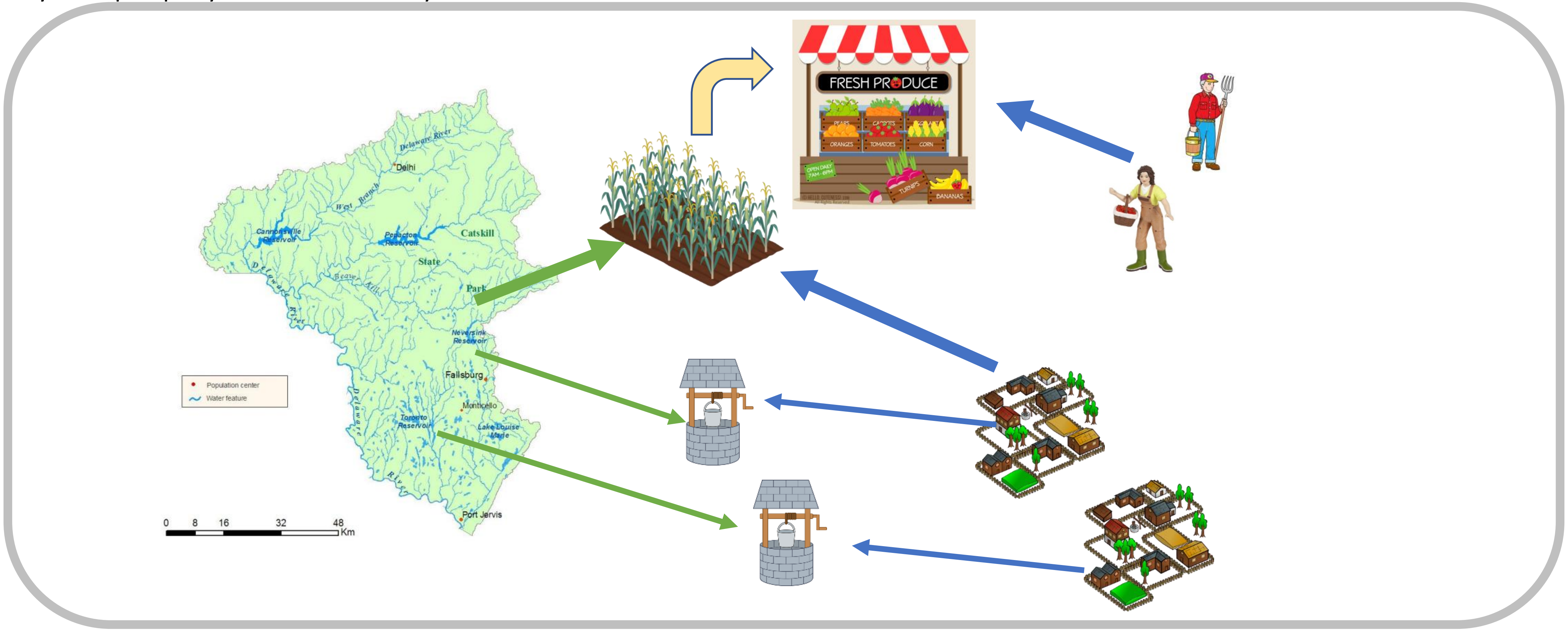
Providers (e.g. forests, watersheds): where valuable ecosystem function happens

Transactors (e.g. wells, crops, atmosphere): where natural value is generated

Beneficiaries (e.g. farmers, coastal dwellers): demand agents for natural value

Example: building an eco-social flow network

Triggered by a simple query: “observe social dynamics of water in watershed X”



The model for the *system* then creates and classifies all relationships, starting with *provision* (provider->transactor)…

…and following with use (beneficiary <- transactor), building a (potentially) differently scaled model for each flow.

Intermediate transactors (e.g. markets) are brought in according to the ontologies. They can be local or remote.

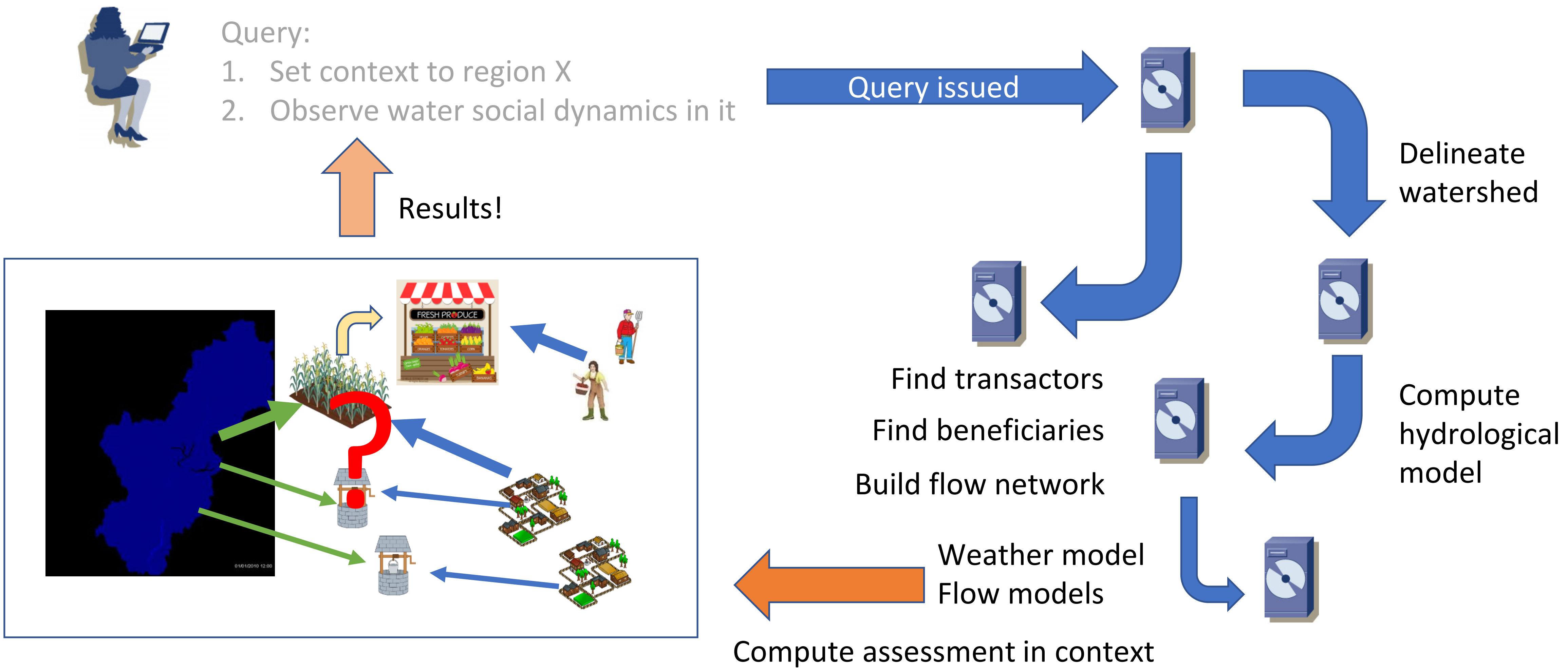
Providers (e.g. forests, watersheds) are first identified and built by the AI engine.

The ontologies define types of Transactors (e.g. wells, crops, atmosphere), identified last.

Beneficiaries (e.g. farmers, coastal dwellers) are identified next.

Models and data live on a semantic web

An extensible network hosts data, models and model services available to users



A semantic approach to modelling

Address all the "W's of information – what, where, when, why, and how – without becoming too large or complex to learn and use.




SUBJECTS:	A mountain	A population of humans	A forest	A river
QUALITIES:	Elevation (measurement)	Per capita income (value)	Percent tree canopy cover (%)	Stream order (ranking)
PROCESSES:	Erosion	Migration	Tree growth	Streamflow
EVENTS:	Snowfall	A birth	Death of a tree	A flood event

RELATIONSHIPS: ↖ Skiers using a mountain for recreation ↗ ↖ A city using a river for water supply ↗

Semantic ROLES allow to account for "alternative views" of these observables without giving up consistency

References

- https://f1000research.com/articles/6-686/v1

F1000Research

F1000Research 2017, 6:686 Last updated: 07 FEB 2018

Check for updates

METHOD ARTICLE

Semantics for interoperability of distributed data and models: Foundations for better-connected information [version 1; referees: 2 approved with reservations]

Ferdinando Villa[1], Stefano Balbi[2], Ioannis N. Athanasiadis[3], Caterina Caracciolo[4]

[1]Basque Centre for Climate Change (BC3), IKERBASQUE, Basque Foundation for Science, University of the Basque Country, Leioa, 48940, Spain
[2]Basque Centre for Climate Change (BC3), University of the Basque Country, Leioa, 48940, Spain
[3]Wageningen University, Wageningen, 6706 KN, Netherlands
[4]Food and Agriculture Organization of the United Nations, Rome, 00153, Italy

Every observation has a *subject*

- Countable, physical, recognizable object
- Examples:
  1. A mountain
  2. A population of humans
  3. A population of trees (i.e., a forest)
  4. A river




Data describe a specific *quality* of a subject

- They require a reference quantity to describe (e.g., measurement unit, category system)
- Examples:
    1. The elevation of a mountain (measurement)
    2. Per capita income of a group of humans (value)
    3. Percent tree canopy cover (proportion)
    4. A river’s stream order (ranking)




Over time, subjects experience *processes*

- Examples:
    1. Erosion of a mountainside
    2. Migration of a human population
    3. Tree growth in a forest
    4. Streamflow in a river




A process is an *event* when we see it happen as a unit

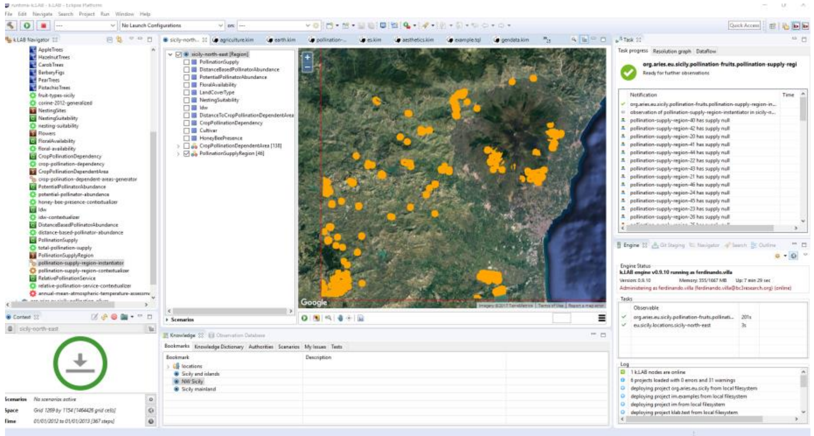
- Examples:
  1. A snowfall event on a mountain
  2. The birth of a new human in the population
  3. The death of a tree in the forest
  4. A flood event on a river

Tooling (1): languages and modelling software

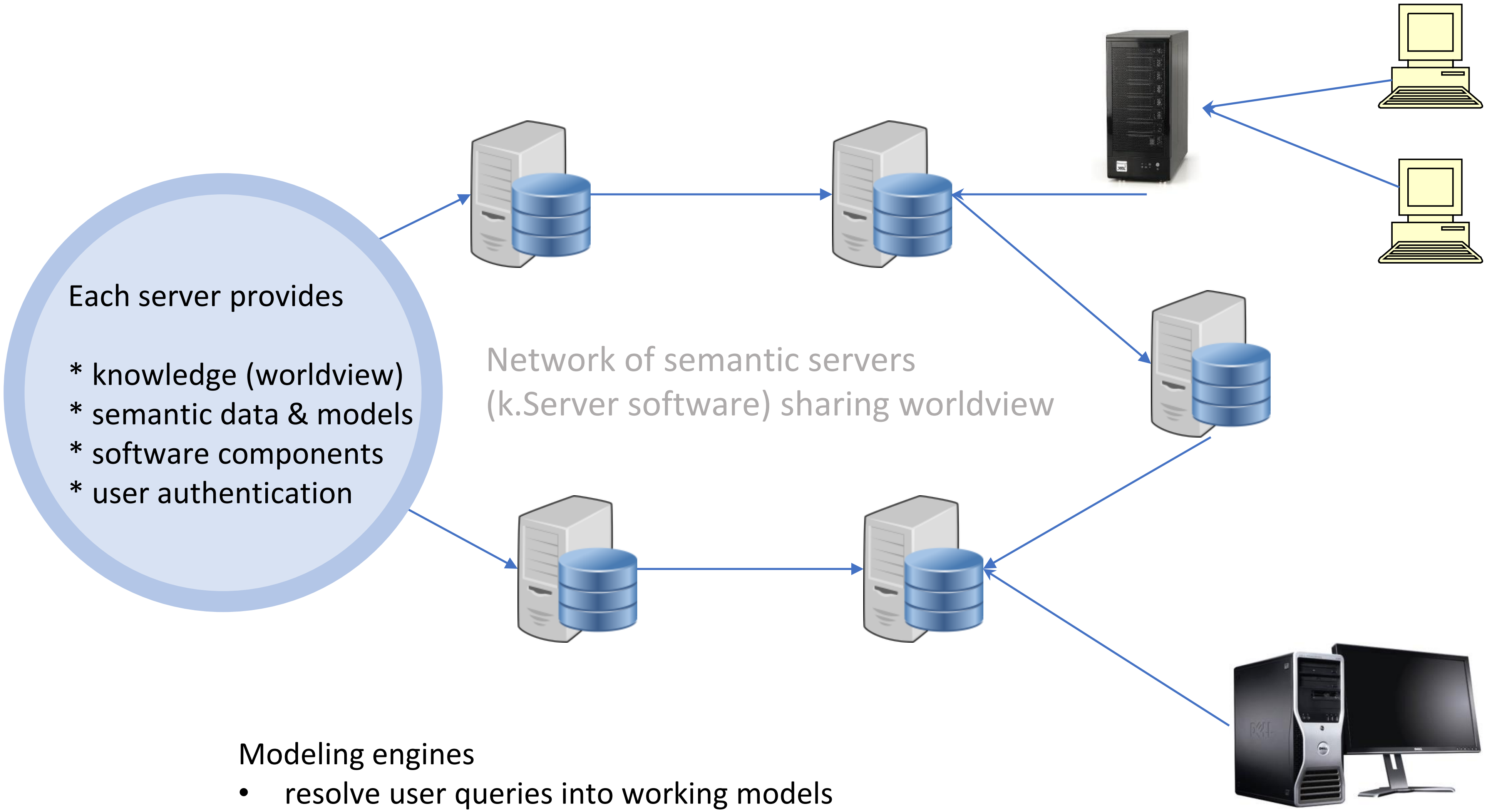
```
role PollinatorSupplier
    is ses:Provider
    applies to earth:Region
    implies PollinatorAbundance as ses:Supply;

role AgriculturalProductionDependent
    is ses:Beneficiary
    implies PollinatedYield as ses:Demand
    applies to observation:Subject;

/**
 * Roles that define the P->T and B->T relationships.
 */
role PollinationSupplyConnection
    is ses:ProvisionFlow
    applies to im:MatterTransferConnection between PollinatorSupplier and PollinationDependent;

role AgriculturalUseConnection
    is ses:UseFlow
    applies to im:MatterTransferConnection between AgriculturalProductionDependent and PollinationDependent;

/**
 * Role for the ES, tying everything together.
 */
role PollinationEcosystemBenefit
    "The benefit obtained by any user of the yield made possible by pollination. This is
     easier to monetize than most ES when defined this way."
    is ses:ProvisioningEcosystemBenefit
    implies at least 1 PollinationSupplyConnection, at least 1 AgriculturalUseConnection
;

/*
```

The k.IM language is used to express both the worldview and the data/models that use it



- Tools and interfaces enable end users, modelers, and network administrators
- Simplify the tasks of semantically describing, coding, and publishing data and models.
- Provide and maintain documentation, community resources for discussion, user support and bug reporting
- Create tools for participatory, graphical model building that can be directly translated into templates for working models.

Tooling (2): distributed semantic web infrastructure



Web users connect to modeling engine and use web interface for modeling

UI is drag-and-drop only, optimized for non-technical user

Modelers have a private engine and connect to the network using the modeling IDE

UI is optimized for developers and modelers

Modeling engines

- resolve user queries into working models
- find content on the semantic network

User Interface(s): a preview

1. ARIES Explorer: the user perspective (Monday to Wednesday)

1. k.LAB Modeler: the modeler perspective (Thursday and Friday)