



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. Athanasiadis³, Caterina Caracciolo⁴

¹Basque Centre for Climate Change (BC3), IKERBASQUE, Basque Foundation for Science, University of the Basque Country, Leioa, 48940, Spain

²Basque Centre for Climate Change (BC3), University of the Basque Country, Leioa, 48940, Spain

³Wageningen University, Wageningen, 6706 KN, Netherlands

⁴Food and Agriculture Organization of the United Nations, Rome, 00153, Italy

v1 First published: 17 May 2017, 6:686 (doi: [10.12688/f1000research.11638.1](https://doi.org/10.12688/f1000research.11638.1))
Latest published: 17 May 2017, 6:686 (doi: [10.12688/f1000research.11638.1](https://doi.org/10.12688/f1000research.11638.1))

Abstract

Correct and reliable linkage of independently produced information is a requirement to enable sophisticated applications and processing workflows. These can ultimately help address the challenges posed by complex systems (such as socio-ecological systems), whose many components can only be described through independently developed data and model products. We discuss the first outcomes of an investigation in the conceptual and methodological aspects of semantic annotation of data and models, aimed to enable a high standard of interoperability of information. The results, operationalized in the context of a long-term, active, large-scale project on ecosystem services assessment, include:

1. A definition of interoperability based on semantics and scale;
2. A conceptual foundation for the phenomenology underlying scientific observations, aimed to guide the practice of semantic annotation in domain communities;
3. A dedicated language and software infrastructure that operationalizes the findings and allows practitioners to reap the benefits of data and model interoperability.

The work presented is the first detailed description of almost a decade of work with communities active in socio-ecological system modeling. After defining the boundaries of possible interoperability based on the understanding of scale, we discuss examples of the practical use of the findings to obtain consistent, interoperable and machine-ready semantic specifications that can integrate semantics across diverse domains and disciplines.

Open Peer Review

Referee Status:

Invited Referees		
	1	2
version 1		
published	report	report
17 May 2017		

- 1 **Carole A. Goble** , University of Manchester, UK
- 2 **Pier Luigi Buttigieg** , Helmholtz Centre for Polar and Marine Research, Germany

Discuss this article

Comments (2)



This article is included in the **Global Open Data for Agriculture and Nutrition** gateway.

Corresponding author: Ferdinando Villa (ferdinando.villa@bc3research.org)

Competing interests: No competing interests were disclosed.

How to cite this article: Villa F, Balbi S, Athanasiadis IN and Caracciolo C. **Semantics for interoperability of distributed data and models: Foundations for better-connected information [version 1; referees: 2 approved with reservations]** *F1000Research* 2017, **6**:686 (doi: [10.12688/f1000research.11638.1](https://doi.org/10.12688/f1000research.11638.1))

Copyright: © 2017 Villa F *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: Development was originally funded by the US National Science Foundation (grant 9982938) and received partial support from the ASSETS and WISER projects funded by ESPA/NERC (grants NE-J002267-1 and NE/L001322/1). ARIES is developed with partial support from the EU-Horizon 2020 project AQUACROSS (grant agreement no. 642317).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

First published: 17 May 2017, **6**:686 (doi: [10.12688/f1000research.11638.1](https://doi.org/10.12688/f1000research.11638.1))

1. Introduction

In an increasingly connected world, the value of information depends not only on the ability to use it for the purposes for which it was collected, but also of reusing and linking it within an expanding information landscape. The term *interoperability* refers to the ability of information to be reused and linked across and beyond the institutional and disciplinary contexts where it originates. In recent years, much attention has been paid to interoperability, not only in empirical science, but also in sectors such as government, industry, the military, policy making and information management.

Disciplines emphasizing the study of information in a variety of sciences (e.g. bioinformatics, ecoinformatics, geoinformatics) have emerged to focus on reusability and integration of data artifacts and models. Reusability, versatility, reproducibility, extensibility, availability, and interpretability of information were identified as key requirements for sustainability¹. Wilkinson *et al.*² outlined the FAIR principles for data stewardship and management, calling for Findable, Accessible, Interoperable, and Reusable scholarly data publication. In practice, these goals can be enabled in different ways, with an exact interpretation that depends on the application. The most demanding interpretation of the FAIR principles can be seen as the one sought in support of the Linked Open Data paradigm³, in which information can be found, retrieved, linked and operated upon in the ultimate “machine actionable” way: *unsupervised* and automated, so that distributed computational workflows and models can be not only built and run, but also *discovered* on distributed repositories with negligible risk of misalignments. We refer to this interpretation as FAIR+, where the “+” in FAIR is rigorous enough to be trusted for automated, unsupervised linking in model and data workflows. A corollary requirement of FAIR+ is the need for information products to carry enough metadata to allow ranking of multiple candidates for linking, in order to choose the one most appropriate for the context of use.

The work presented here is part of a wider investigation on a methodology we call *semantic meta-modeling* (SMM), which enables the definition and execution of potentially complex, distributed scientific computations (*scientific workflows*⁴), based on automated semantic inference and powered by FAIR+ interoperability. In SMM, data and models can be (1) discovered on linked repositories based on semantics alone; (2) ranked for appropriateness to the intended context of use; and (3) assembled automatically into coherent, working scientific workflows. The authors, in collaboration with others, have been working on SMM for about a decade and produced a proof-of-concept software stack, named k.LAB, which operationalizes the approach. The first large-scale project building on the SMM paradigm, ARIES (ARtificial Intelligence for Ecosystem Services⁵), has provided the primary rationale and testing ground for the development of SMM. The ARIES project implements a distributed semantic web platform for ecosystem services⁶ modelling, where users are presented with the result of computing scientific workflows built automatically as a response to conceptually stated queries (e.g. “observe carbon sequestration in year 2010 in the Danube watershed”). In this article, first in a series of planned contributions illustrating the different aspects of SMM, we describe the semantic principles and methods that underlie and enable interoperability in the approach, incorporating the

feedback of ~15 researchers and ~150 ARIES modelers since 2007. Further contributions will expand on aspects of SMM not described here, in particular (1) assembling and running model workflows, (2) the details of the software implementation, and (3) the community process that has allowed us to build a distributed base of semantically annotated informational resources.

1.1. Background and rationale

The vision of a Semantic Web^{7,8} brought semantics to the foreground as an instrument to integrate diverse, independently developed information. The use of digitally stored *ontologies* (formal vocabularies paired with logical axioms describing their relationships and intended meaning⁹) has since become commonplace for *annotating* informational assets, i.e. adding concepts from ontologies to the associated metadata to enable their integration and reuse (e.g. 10,11). Research and progress in ontology-mediated interoperability have been significant, and interest in it remains high. Yet, the promise of semantic annotation has often been disappointing in practice, as describing the conceptual underpinnings of information in a way that is complete, consistent and understandable across disciplines and communities has proven difficult and elusive. Attributing stable, reliable and shared meaning to information is difficult because of a lack of accepted best practices, confusion about the phenomenological nature of observed entities and attributes, lack of accepted rules on how to choose, specialize and connect concepts, among other inevitable logical challenges. The result has been a confused landscape of mixed, incompatible attempts, which Goguen¹² described as “... the creation of a constantly shifting foreground and background, with the latter being called ‘context’”.

Ontologies for science come in many varieties. *Foundational* (also *upper* or *reference*) ontologies aim to provide philosophical foundations upon which to build lower level, domain-specific ones. They describe abstract, high-level concepts with the aim of establishing foundational logic for the definition of domain-specific concepts in derived ontologies¹³. For example, they may define the difference between abstract and concrete or establish the logical underpinnings of spatial, temporal, or part-whole relationships. Well-known foundational ontologies include DOLCE¹⁴, BFO (15, also see 16), and more comprehensive efforts like SUMO¹⁷, endorsed by the IEEE. Foundational ontologies have been successfully used and some have seen relatively broad adoption, but due to their high abstraction they cannot alone solve the issue of “what is what”, aside from providing a base for more specific *domain* ontologies (see below).

A second class of conceptualizations, *observation* ontologies (e.g. OBOE¹⁸, O&M¹⁹), also includes high-level concepts, but uses the notion of observation as the main device to introduce semantics, focusing on the “how” of observation rather than the “what” of phenomena. Emphasis is given to aspects such as the type of observation (e.g. measurements vs. rankings vs. classifications) and their use context. Observation ontologies thus try to occupy a foundational niche without committing to a specific phenomenology. Observation ontologies have been relatively successful in terms of adoption, but due in part to the lack of a phenomenological underpinning they cannot guide investigators in choosing

appropriate observables, nor, by themselves, guarantee any of the FAIR principles for interoperability.

Domain ontologies, in contrast, describe specific areas of interest. Although in principle they should be used in conjunction with foundational ontologies, they are typically produced to serve the needs of a specific community, and are seldom committed to interoperability with ontologies from other domains¹³. The interoperability enabled in such a situation is primarily syntactic (“we use the same vocabulary, even if the semantics are not well thought out”) and therefore limited to users of the same ontology. Commonly these are formed as *taxonomies*, using a hierarchy of specialization (*is-a* relationships, for example: *Precipitation is-a AtmosphericPhenomenon*) to organize and systematize the terms used within the communities endorsing their development. Examples abound in various domains, including earth and environment: for example, SWEET²⁰, ENVO²¹, SPAN/SNAP²², Gene Ontology²³, and PlantOntology for plant anatomy and morphology²⁴. Some of them, such as the Gene Ontology²³, have been highly successful; yet the terminology is easily drained of meaning when confronted with other disciplinary contexts that use the same terms differently. For example, a *crop* is, to an agricultural economist, the agricultural product that reaches the market, possibly further processed after harvest, while to an agronomist the same term refers to the producing plant species. In general, domain ontologies commonly strive to endorse the term usage that is most popular in a community; this is both the reason for their success within those communities and their primary limitation.

Large investments have been made in developing *controlled vocabularies* to describe and discover artifacts of interest for specific disciplinary sectors, meant to facilitate annotation and sharing of information objects. Controlled vocabularies are typically domain-centered, with little or no pretension of phenomenological adequacy, but have strong links to the culture, language and applications in their community of origin, usually endorsing terms and assumptions that are best recognized in a community of reference. When formally expressed, their structure is often inspired by organizational, rather than logical, reasons. Many high-adoption examples exist, including e.g. AGROVOC²⁵ and CABT²⁶ for agriculture, CUAHSI for hydrology²⁷, and MESH for medicine²⁸. Generally, vocabularies contain a large number of terms (e.g., for taxonomic or chemical species of interest), are often multilingual, and can grow rapidly with use, while, by contrast, ontologies strive for minimality and robust logics. The differences between ontologies and controlled vocabularies are often misinterpreted, and in common practice the terms are sometimes used interchangeably.

Exploring ontology repositories, such as the OBO Foundry²⁹, immediately shows that notable inroads have been made in both fundamental and domain ontologies, and that efforts to produce multi-domain conceptualizations based on common semantics are made regularly. Despite these attempts, duplications, ambiguities and inconsistencies continue to hamper the development and adoption of semantic annotation standards in practitioner communities. With the current state of the art, and with the semantic web community seemingly more interested in enabling technologies than in the conceptual aspects of interoperability,

FAIR+ interoperability remains all but impossible, perhaps aside from within very restricted communities. In our view, a major need for progress towards this goal is a solid, uncontroversial phenomenological base, i.e. the basic semantics for the types of phenomena and entities that can be understood by human observers. This phenomenology needs to be general enough to work across domains and worldviews. Formalisms and toolsets must be built to support it, to ease the specification of domains and allow for extension, while enforcing a consistent design discipline. We need clear best practices for specialization and connection of terms, and guidelines on how to integrate always growing, and potentially infinite, domain content from vocabularies without breaking the logical integrity of the resulting annotations. We start our discussion by focusing on the definition and preconditions of interoperability itself, before illustrating the details of the approach we have found useful as a possible starting point towards FAIR+ interoperability.

1.2. Interoperability of observations

In this article, we are concerned with the process of creating informational artifacts describing an **observable** concept in a chosen context, to provide evidence for a scientific deduction or computation. By **observation** we refer any artifact that is *resolved* from the perspective of a scientific process, i.e. can be used without requiring any further observation. Observations in this sense are commonly called “data”; yet, this term can be ambiguous with respect to the semantics of the observables involved, as we will explore later.

Definitions for interoperability vary by purpose, and while emphasis is sometimes given to formats and protocols for encoding and transmission (syntactic), legal compatibility (i.e., copyright and licenses) or organizational aspects (e.g., openness and purposes of the data), most definitions involve semantics - the *meaning* of what the information represents. The most rigorously structured information, such as scientific data and models, present the most stringent challenges in establishing semantic equality. Interoperability in such situations concerns links between two informational endpoints: for example, finding data to link to model inputs, or linking the outputs of one model to the inputs of another, so that a single computational chain can be established without fear that its results will be invalid. We term observations *compatible* to refer to this definition of interoperability between two informational endpoints (we will provide a more formal definition in Section 2.6).

We maintain that this kind of information alignment has three nearly independent semantic dimensions: semantics of the *observable*, the *observation* and the *context*.

Observable semantics describes *what* observations are about: physical objects, events, processes, agents, or characteristics that may be “observed” or measured. The human observer recognizes relevant observables, e.g. elevation (a *quality* pertaining to a location on Earth), households (*subjects*, part of villages or cities), or surface water flow (a *process* observable in watersheds). Much of what we call “data” consists of observations of qualities; their inherent subjects (e.g., the location on Earth whose elevation we observe) are often specified indirectly or implicitly. In order for observations to be made, and for their interoperability to be possible, it is crucial that such identities are fully specified and

unchanging. For two observations to be compatible, their observables must be described by the same concept.

If a physical object, event, process or relationship can be simply acknowledged to exist in a context of interest, *qualities*, such as the elevation of a mountain or the temperature of a body, can only be observed *indirectly*, i.e. by comparison with reference observations. Units of measurement, currencies, rankings and classification systems define the ways human observers quantitatively or qualitatively describe such observations. **Observation semantics** describe *how* the observation activity is carried out, detailing the choice of reference metrics to ensure that a state can be understood and mediated. Mediation between different observation semantics is often possible, sometimes exactly (units of measurement, typically converted with negligible loss of precision), sometimes approximately (prices in different currencies can be more roughly compared by adjusting for inflation and purchasing power) and sometimes hardly (e.g. different land cover classifications systems are often extremely difficult to mediate). To be assessed for compatibility, qualities need the full statement of how they are observed; mediation operations may be necessary to harmonize two observations before computation involving both can take place.

Observation always happens in a **context**, providing observables with a *when* and a *where*. The context is usually chosen *a priori* by the actor who created the original artifacts, and may differ, subtly or greatly, between two compatible observations. Just like the scale of a geographical map determines what entities are visible in it (urban streets disappear in a 1:2,000,000 map), certain observables only come into focus at a given geographical scale, and certain phenomena emerge only at a given temporal scale. **Context semantics** describes these aspects for an observation. While differently scaled observation can be mediated to some extent through aggregation or propagation (with loss of information), scale also reflects deeply on semantics; large scale shifts will determine incompatible semantic misalignments, with (e.g.) uncountable processes becoming countable events when their time scale is changed beyond a threshold. For example, lightning is seen as an event by a meteorologist and as a process by a high-energy physicist; subjects, e.g. the microorganisms in a lake, become visible only through qualities (the color of the lake) over a spatial extent that makes the observer lose sight of them. For this reason, **scale** is key to establishing *meaning* in more ways than usually recognized; scale depends entirely on the *chosen* observation context, therefore on the human decision of what to observe. The semantics of scale largely deal with space/time, and as such, can be formalized independently from the observables'.

The ability to accurately characterize semantics along observable, observation and context dimensions addresses the *interoperable* and *reusable* FAIR criteria. Semantic specifications can be rewritten into queries that select interoperable counterparts for an observation, addressing the *findable* and *accessible* requirements. If queries embodying all three dimensions can be executed and the resulting observations can be ranked for appropriateness, unsupervised linking becomes possible. While observation and context semantics are relatively well-understood, the characterization of what things *are* - observable semantics - remains difficult and uncertain, even

with increasing investments in ontologies and vocabularies and an engaged community behind the current state of the art.

The rest of this article details our approach to building foundations for FAIR+ interoperability through semantics, using examples. We describe a conceptual framework and examples of reasoning and specifications to support the characterization of observable semantics, resulting from field-testing in years of initial application, followed by a discussion of goals already achieved and those that remain. Here, and in all applications of these principles and methods, we argue for SMM as a driver for a *semantics-first* workflow, where the lifecycle of information begins, before data collection or model development, with the understanding of semantics, which in turn guides data collection, organization and processing up to eventual documentation, storage and curation. This contrasts with the more commonly adopted *annotation* approach¹⁰ where *data* represent the first-class artifacts, collected and stored with a logic dictated primarily by practical constraints, and semantics may complement the artifacts "after the fact" to suit the data to specific applications.

2. Materials and methods: specifying observable semantics

Semantics starts with the act of cataloguing observed reality into classes that can be referenced and communicated. Terms describing commonly acknowledged classes of physical entities (such as persons or objects) are complemented through inference, comparison, association and imagination, to encompass objects, events, processes and relationships that may not be directly perceived by the senses, but still appear in human experience, thought and communication. Such *observable* entities can be arranged along a small number of fundamental phenomenological categories (for example physical objects, processes, characteristics or events) that determine how they can be described, observed, modeled and represented. Our perception of space/time is crucial to the process of organizing reality into communicable observables, as the "resolved" units of space and time determine how we classify. This is particularly important when science's exacting descriptive needs of come into play: as described previously, spatiotemporal resolution influences an object of study's perceived structural or functional character, and shifting resolution or extent can fundamentally alter an observable's perceived category. It follows that interoperability can exist in a conceptualization, as long as the boundaries of stability of meaning for all concepts with respect to their fundamental phenomenology are stable. *Scale*, commonly defined as the choice of resolutions and extents through which we make observations of the world, binds the observables of informational artifacts to precise phenomenological categories, establishing boundaries of validity for conceptualizations.

In SMM, we call any domain conceptualization where every term has a stable and explicit phenomenological characterization a *worldview*; we recognize the worldview as the outer possible boundary for interoperability. In practice, a worldview is a set of ontologies that describes meaning under the viewpoints set by a given range of scales, where terminology is unambiguous and its relation with the chosen phenomenology is stable. Relationships

between fundamental types in the chosen upper ontology create binding constraints for the entire conceptualization and provide guidance for semantic consistency and validation. The worldview we use in the ARIES project is the primary source for the examples in the rest of this article. This worldview focuses on spatial and temporal scales broadly in tune with human life, dealing with entities, processes, events and relationships that characterize and bound socio-ecological, economic and agricultural systems. We thus anticipate that it can provide semantic building blocks for data management and modeling across a wide range of applications in socioeconomic and environmental simulation. While we found this worldview adequate to represent Earth systems data and models, we would be hard pressed to suggest its use in disciplines whose scales of interest are widely distant, such as cosmology or high-energy physics.

The formalism described in the rest of this article outlines a simple metaphysics dedicated to the practical description of observations. In this view, things exist as long as observations of them can be produced. Our later use of philosophical terms, such as *universal*, *particular*, etc.³⁰, exclusively serves this interpretation, and may slightly differ from other definitions used in philosophy and computer science. Ample discussion of these terms and their meaning can be found in the philosophical literature^{9,15,30}. Our work aims to enable FAIR+ interoperability in scientific workflows, outlining a *minimal* and *practical* phenomenological basis intended to be simple and intuitive enough to be internalized by large numbers of practitioners. Our choice of terms has evolved along a span of about eight years, reflecting design and planning, plus several years of exposure to and feedback from diverse users in academic, governmental and non-governmental sectors.

2.1. A simple metaphysics for scientific observables

We briefly articulate the phenomenological basis for our approach below, starting with the fundamental logical dichotomy of **universals vs. particulars** and further dividing particulars into **continuants** and **occurents**. In our interpretation, these terms all refer to concepts; we are not concerned here with concrete instances (e.g., the individual tree, as opposed to the *idea* of a tree) as our only aim is to produce *observations* - informational artifacts generated through the process of observing concepts in the world.

Based on this reasoning, we use the term *particular* to refer to *concepts* that describe observables for which an observation can be made, as described above, although in some literature the term is used to refer to instances. Particulars include (1) physical objects, (2) their qualities, the (3) processes and (4) events that affect them (whose observation is likely to describe the qualities affected, causal pathways and component objects) and (5) the relationships that connect them. In contrast with particulars, we use the term *universals* to refer to concepts that cannot be observed directly. In much literature (e.g. 15), the term universal simply means a concept (the abstraction of an entity, as opposed to the entity itself), so it includes concepts that we classify as particulars, such as processes or physical objects. We take a stance closer to Platonic realism³¹, which defines universals as those notions that cannot be directly incarnated unless associated with a particular. This includes classes of concepts such as attributes (e.g. ‘black’ cannot have an instance,

but it can qualify a physical object, e.g., a cat), roles and others. In SMM, we translate universals into entities that cannot be observed in their own right. Only observations of particulars can be made, and universals are attributed to them to further specify their semantics.

Within **particulars**, the *continuant vs occurrent* distinction reflects how observed entities stand with regard to space and time. This distinction is found in all foundational ontologies with slightly different definitions or terminology¹⁶; for example, DOLCE¹⁴ uses the terms *perdurant* and *endurant* instead. Continuants are entities that maintain their identity through time, including physical objects (named *subjects* in SMM) and their measurable *qualities*. For example, color or height, which can be observed only when linked to other entities through mandatory *inherency*: e.g., the height of a tree. While continuants *are*, occurents, such as events and processes, *happen*: their definition is intimately tied to time. As discussed previously, a spatial shift in the observation point can morph continuants from countable *subjects* to uncountable *qualities* as spatial resolution moves upwards and small-scale subjects lose their individual visibility in favor of larger-scale ones. Similarly, countable *events* can morph into uncountable *processes*, as temporal resolution shifts to allow appreciating change within what was formerly seen as an individual event. *Relationships* between two observations also reflect the continuant-occurrent dichotomy; accordingly, they can be seen as *structural* (unconcerned with time, such as parent-child) and *functional* (such as flows, whose expression is a time-dependent process).

Figure 1 illustrates how observational scale and the categorization of particulars are intimately linked. A temporal scale gradient (Y axis) separates occurents - for which fine temporal resolution allows an observer to appreciate change - from continuants, the meaning of which can be appreciated independent of time, due to the temporal scale being coarse enough to make change invisible. On a spatial scale gradient (X axis), “close” observation focuses *within* individual observations, impeding the appreciation of their individuality (therefore the “counting” of separate individuals), but enabling the observation of their inherent qualities and processes. As spatial scale is made coarser, the point of view moves *outside* the individual observation, allowing an observer to appreciate first the individual relationships between two of them, then an arbitrary number of them in the context of a larger-scale observation (not shown in the image). The property of *countability* tracks meaning along a spatial scale gradient in the same way that the dichotomy “happens/is” tracks meaning along time scale gradients. The diagram in Figure 1 has proven intuitive enough for ARIES users to remember and use as guidance in the first steps of semantic annotation.

2.2. A formal language to specify interoperable semantics

During the development of the ARIES project, it quickly became apparent that the explicit statement of semantics was key to achieving our goals of building and linking community-driven, interoperable repositories of independently developed data and models. At the same time, it became clear that no community of modelers, data scientists or other prospective users would consider an investment in OWL or other semantic web-endorsed formalism as the vehicle to express the semantics in data and models, and that a different

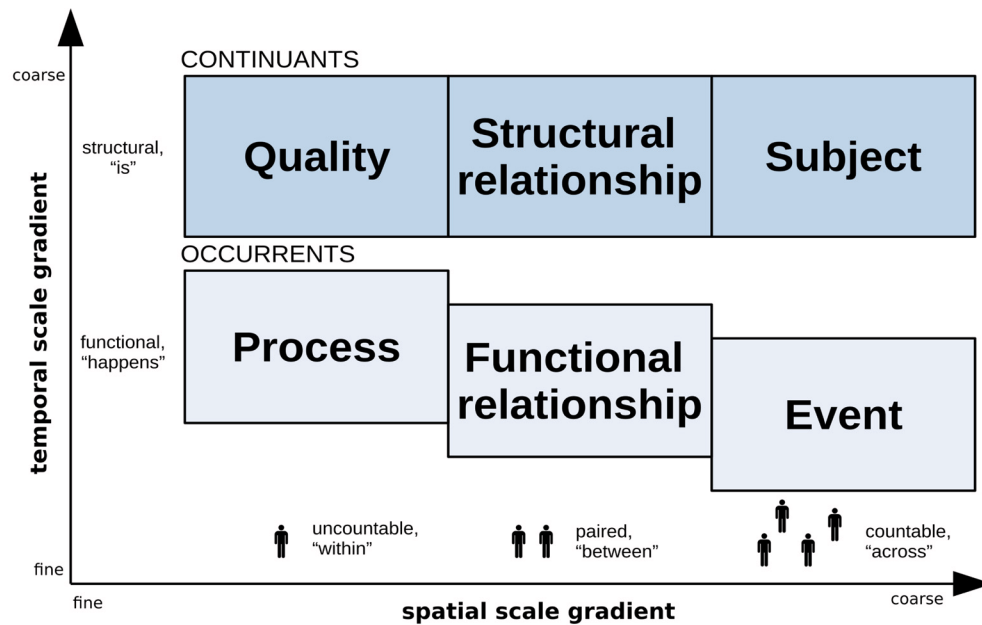


Figure 1. A synthetic representation of how the spatial and temporal dimensions of observational scale influence the identification of the fundamental categories of observables discussed in the article. Refer to the text for explanations.

approach was necessary. Our solution was the design of a custom semantic specification and annotation language, for which we laid out four main requirements.

1. **Full compatibility** with accepted semantic web standards. In the current implementation, this translates into the ability of any specification to compile to OWL2.
2. **Expressiveness:** syntax and keywords should intuitively relate to the phenomenology and experience of scientific observation, so that the terminology and complex logical constraints in the underlying ontology do not need to be learned or exposed.
3. **Readability:** the language should read as close as possible to English, using familiar terms that are as easy as possible to learn and memorize.
4. **Parsimony:** the language should support flexible composition of terms to allow the terminology to remain as small as possible, enabling the greatest possible reuse of terms.

The result of many years of design and user feedback is the **k.IM language**, which currently makes worldviews accessible to ARIES modelers. The k.IM (for “knowledge-Integrated Modeling”) language is complemented by an open source software stack named k.LAB, which provides integrated tools to develop and use conceptualizations and models using k.IM. The software, which in its current alpha stage requires **training** to be applied, will not be discussed here, but can be **freely downloaded** and **explored in source form**.

In k.IM, particulars and universals are combined to specify observables; these can later be used to annotate data and models. Keywords and syntax rules are designed to make k.IM statements readable and understandable by mimicking English syntax, while specifying much more complex, correct and consistent OWL³² axioms. All k.IM statements compile to OWL2, the most widely used and accepted representational standards for ontologies. Conceptualizations written in k.IM can thus be exported and used in OWL-based systems with no loss of information.

Experience with developing and teaching k.IM has highlighted three clearly distinguishable tiers of sophistication in semantic annotation practice, arranged here by decreasing levels of experience required and progressively shorter learning curves. Tier 1 is annotation of data and models, performed by minimally trained users utilizing the terms from domain ontologies, facilitated by context-aware search tools. Tier 2 is domain definition, where domain experience is the essential skill, but an investment in knowledge engineering remains necessary. Tier 3 is worldview definition, limited to knowledge engineers with ample time to invest in work with domain experts. All three tiers are enabled in the k.LAB community and will be discussed and exemplified below. Tier 3 usage is discussed in Section 2.3. Examples from tiers 2 and 1 will be given in Section 2.5.

2.3. Building a worldview: Definition of domains

The following examples, taken from the socio-ecological system worldview used in ARIES, illustrate some of the most important aspects of k.IM and their role in facilitating the conceptualization

of domains. For clarity, the examples are highlighted in the same way the k.LAB editor does: keywords (identifiers recognized as part of the language) are in purple; user-generated text (including concept identifiers) is in black; literal text (such as quoted strings) is in blue. **Table 1** provides a more systematic list of keywords with their associated meanings. We only discuss the features of the language

used to specify concepts; those concerned with data annotation will only be briefly described, while the features concerned with modeling will be discussed in a forthcoming contribution.

Geographical *elevation* is a quality inherent to regions of Earth, whose full specification involves different notions, some specific

Table 1. Base observable and universal types with the k.IM keywords that declare them. Keywords (in bold) are used in the language to state concepts. Other keywords indicated can be used to specify relationships between concepts (e.g. exposes, describes, implies).

Base category	Base concept	k.IM keyword	Explanation
		configuration	Any combination of qualities or other arrangement that can be recognized by an observer without being identifiable as a single quality, subject or other observable.
continuant	subject	thing	Any inanimate physical body, as defined by an external observer. From the point of view of observation, cavities and observable “absences” are also things.
		agent	Any physical body that is the context for autonomous processes that define its identity, including awareness of its own agency.
	quality	priority	A quality that can be ranked numerically, but no assumption is made about the scaling of the correspondent values beyond ordering.
		quantity	A quality whose states are expressed numerically.
		<i>physical properties</i>	k.IM provides keywords for the basic physical properties used in science, such as temperature, energy, mass, volume, length, area , among others. The keywords establish their physical nature (extensive or intensive) and enable validation of units of measurement.
		class	A class is a special quality that exposes one or more traits of the context. Universals are not observable but a class allows to attribute “data” to describe combinations of traits. For example, land cover type can be seen as a combination of traits describing forests, urban texture etc. Using a class allows the semantics in complex classifications to be preserved and reasoned on.
	relationship	structural relationship	A relationship between two subjects whose existence does not depend on time, e.g., parent-child.
occurrent	process	process	A phenomenon that happens within a single subject and is observed as it evolves through time, creating change in the value of the subject's qualities and potentially creating or destroying subjects, events or relationships.
	event	event	Phenomena resulting from dynamic action that are seen as an atomic and countable at the temporal scale of observation.
	relationship	functional relationship	A relationship between two subjects whose observation implies a dynamic process, such as a flow of money between two commercial parties.
universal	trait	attribute	Any attribute that is not an ordering, realm, or identity. Attributes usually describe a quality that cannot be observed as a particular in the context of observation.
		ordering	An attribute whose subclasses define an ordered list, e.g., high, medium, low.
		realm	An attribute that describes a subdivision of the context and qualifies the observable as belonging to it.
		identity	An attribute used to identify a subject according to a species or other identifying type.
	role	role	A function assumed by an observable when appearing in the context of another. A role adopted by an observable may imply other roles for observables related to it in that context.

to the geographic domain, others of more general relevance. We use **namespaces**, associated with separate URLs or files, to separate concepts from different knowledge domains; a namespace can import another (through a *using* clause as shown below), so that the concepts defined in it can be referenced. Concepts from imported namespaces are referred to using the namespace identifier as a prefix to the concept name, separated by a colon (for *earth:Region*, the concept *Region* is defined in the *earth* namespace). The specification in definition (1) comes from the *geography* namespace, declared at the beginning along with its imports.

```
namespace geography using im, earth;

length Elevation
  "Geographical elevation above sea level, as described by
  a digital elevation model." (1)
  is im:Height of earth:Terrain within earth:Terrestrial
  earth:Region;
```

In this specification, the **length** keyword establishes the fundamental character of geographical elevation, including its physical nature (an *extensive* property whose value changes with the

extent of the inherent subject) and the base unit for its measurement. This is done by tying the concept being defined to the core observation ontology, which lays out the phenomenological categories defined above, along with constraints and relationships for all common scientific observables, unseen by users. The language contains keywords for many fundamental quantities, allowing users easy specification in most situations (Table 1). It also provides *semantic operators* to easily and systematically modify existing concepts obtaining derived quality concepts:

```
event Earthquake; (2)

model table("data/earthquakes.csv") as probability of Earthquake;
```

Example (2) needs only one concept, Earthquake, as the annotation of its probability can be done through a semantic operator (**probability of**), which can only be followed by a concept describing an event and produces a concept for its probability. Similar operators allow the expression of presence, occurrence, distance, proportion, ratio and value (Table 2). The use of semantic operators greatly reduces the number of concepts needed in the worldview and enables validation of the modified observable.

Table 2. Semantic operators used in k.IM to produce quality concepts that describe the result of a specific way to observe another. Text in square brackets indicates optional specifications in k.IM syntax. These operators only create concepts, with no assumption over their values. Observer statements (Table 4), which also build the correspondent concepts, are the k.IM specifications that are concerned with the actual states resulting from their observation.

Operator prototype	Applies to	Produces
presence of <O>	Subjects, relationships, events and processes	A quality concept describing the presence or absence of O, whose states are true/false values.
count of <O>	Countables (subjects and events)	A quantity concept describing the numerosity of O in a context.
distance to <O>	Subjects and events in a spatial context	A length concept describing the distance to any observation of O in a spatial context.
probability of <O>	Events	A quality concept describing the probability of O happening.
uncertainty of <O>	Qualities	A quantity concept describing the uncertainty associated with an observation of O.
proportion of <O> [in <O>]	Identities in quantities of more general identity	A quantity describing the proportion of a particular identity in a medium, e.g. salt vs. water volume.
ratio of <O1> to <O2>	Quantities	A quantity describing the ratio between two other quantities.
value of <O> [over <O2>]	Any concept, including non-observables, over matching concepts	A quantity describing the value attributed by the observer to a particular concept, possibly in comparison with another.
occurrence of <O>	Subjects, relationships, events and processes	A shorthand for "probability of presence of O".

As concrete qualities (those of which observations can be made) can only exist inherently to a direct observable, the observable must be made explicit before the concepts can be used (e.g., *earth:Region* in the previous elevation example). In example (1), the concept statement starts with a description (highlighted in blue) that is indexed in the k.LAB software, so that users can easily locate concepts by textual searching. The **is** keyword introduces the semantic specification for the term *Elevation*. In it, *im:Height* (from the base namespace *im*, for “integrated modeling”, also the name of the containing worldview) is first established as its fundamental nature; then, inherency is established by means of the keywords **of** and **within**. Inherency enables validation of the contexts in which the qualities are used. For example, after the definitions in (1), it will be correct to annotate elevation within a watershed, as long as a previous statement defines the *Watershed* concept as a type of *earth:Region*. In many situations, specifying **within** is enough to establish inherency. The **of** keyword is used when the quality refers to a second, implicit observable in the context of inherency. For example, the “height of trees” quality in a region is inherent to that region, but implicitly describes tree subjects in it. In keeping with our readability requirement, we only allow two levels of specification and use two different keywords (**within** and optionally **of**). We found that legitimate chained specifications, such as “x within y within z within ...”, were awkward and difficult to understand in usage tests and decided against allowing such statements. Multiple chains of inherency of this kind can be defined using intermediate concepts.

In knowledge domains (as opposed to physical ones), the implicit inherent subject is often a **configuration**. This is a perceived, measured or inferred arrangement of observables that can be experienced and recognized by humans without being directly amenable to providing the observable of an informational artifact. For example:

```
namespace earth using physical;

configuration Terrain (3)
  "The three-dimensional configuration of land surfaces"
  is physical:PhysicalConfiguration within Terrestrial Region;
```

In k.IM, configurations can only follow **of** in inherency specifications. This constraint allows the construction of clear and unambiguous statements that relate well to scientific discourse while remaining logically consistent. Common logical errors stem from confusing legitimate observables (such as qualities, subjects or events) with “objects of study” in science that are part of daily discourse but are not actually amenable to being *directly* described by informational artifacts. Configurations often allow keeping such concepts (such as *Terrain* above) in a specification without compromising logical integrity. Other examples of configurations include bathymetry, aesthetics and all types of networks, e.g. a stream network or a social network, whose observables are the actual subjects and relationships that create the perceived configuration.

The inherency requirement for qualities is one of the primary means for semantic inference and validation in SMM. Machine reasoning can be applied to ensure proper usage of each concept in data annotations and models. For example, models of a quality that is inherent to a specific subject or process are validated to ensure that all other qualities used for its computation are inherent to a

compatible subject type (see below for a definition of compatibility). Any mismatch makes the model semantically inconsistent and must be solved before it can be computed. At the same time, many inferences are possible through reasoning on inherency. For example, a model’s requirement for “*presence of biology:Tree*” will automatically be satisfied if data for this specific concept are not available, by an observation of “*im:Height of biology:Tree within earth:Region*”. Because non-zero values of extensive physical properties imply the existence of their inherent subjects, a model for presence of trees can be automatically built using the height data and $\text{height} > 0$ as the criterion to establish presence. Interestingly, inherency underlies the mechanism through which shifts in a spatial scale affect identity and meaning in continuants. If a finer spatial scale resolves, e.g., the color of individual unicellular algae subjects in a volume of water, expanding the spatial extent and lowering the resolution may cause the algae subjects to go “out of focus”. Their color now becomes a quality inherent to a larger, previously invisible lake subject.

In specification (1), geographic elevation is established as a **length** by its fundamental keyword, but the definition introduced by **is** defines it as a *Height*, from the base *im* worldview namespace. The definition of *Height* shows the use of attributes to constrain the length concept to a specific orientation relative to the observer:

```
namespace im; (4)

abstract length Height inherits Vertical, Lineal;
```

The attributes *Vertical* and *Lineal* (whose definitions are not shown here) are attributed to *Height* using the keyword **inherits**, establishing characteristics of *Height* beyond its definition as a length. In this worldview, *Lineal* (as opposed to *Areal* or *Volumetric*) is used to ensure that transformations of qualities involving dimensional reasoning (e.g. dimensional collapse under scale aggregation) carry the information that is needed for the algorithms to properly mediate scales and values. While the concept of “length” belongs to the foundational observation ontology, outside the worldview, dimensionality is specific to worldviews, as it could be interpreted differently in other domains (e.g., in non-classical physics). The *Height* concept is declared **abstract** to ensure that observations cannot be made of it; any concrete concepts derived from it - omitting the **abstract** keyword, such as elevation in (1) - must specify their inherency or the k.IM parser will flag an error.

Attributes are an important feature for k.IM to enable fluent specifications while enforcing our parsimony requirement. Definition (1) exemplifies the English-like syntax used to specify attributes that restrict the context to terrestrial regions without creating a new concept. The context for *geography:Elevation* is declared to be “*earth:Terrestrial earth:Region*”, combining two concepts *at the time of usage* by simply mentioning them in sequence. Instead of merging them into another concept, *earth:TerrestrialRegion*, the sequential specification follows the grammatical conventions of the English language and yields more parsimonious ontologies. The ubiquitous use of *is-a* specialization to add attributes to observables (*BlackCat is-a Cat*) is a major cause for explosion of terminologies in domain ontologies. While this is legitimate from a phenomenological

perspective (a black cat certainly is a cat) and from a mathematical logics perspective (black cats certainly are a subset of the set of all cats), we adopt the convention that only clear semantic distinctions should reflect in *is-a* inheritance. As long as an attribute does not obviously modify identity (a black cat is *just* a cat) the specialization should be described without explicitly creating a new concept. Attribute composition through *is-a* relationships can also yield ambiguous inheritance graphs, logical errors and specification dead-ends when many attributes are used but subtypes are intended to inherit only some. As most universals apply to broad classes of observables (color is certainly not just an exclusive attribute of cats), the advantages of the quasi-natural k.IM syntax quickly become apparent in terms of parsimony and readability. This syntax enables the creation of ontologies that are small enough to be learned and used but retain high expressive power. The underlying infrastructure, such as k.LAB, is left in charge of handling, unseen, the axiomatic complexities of concept inheritance and attribute composition.

Attributes are often used to coarsely summarize the value of qualities. In k.IM, we preserve these relationships in order to allow inference of attributes:

```
namespace ecology using chemistry, earth;

abstract attribute Salinity
  describes chemistry:Salinity within earth:Aquatic earth:Region
  has children                                     (5)
    Saline,
    Brackish,
    Freshwater;
```

In definition (5), it is clear how an attribute (*ecology:Salinity*) with its “child” sub-categories is a synthetic and approximate way to describe the actual concentration of sodium chloride in a natural water body, defined in the *chemistry* namespace (see Section 2.4 for details on chemical identities, and Table 2 for the **proportion of semantic operator**):

```
namespace chemistry using im, physical;

quantity Salinity                                     (6)
  is proportion of (NaCl im:Mass) to (Water im:Mass)
  within physical:DelimitedBody;
```

By establishing a semantic relationship between the salinity categories in the *ecology* namespace and the proper salinity definition in the *chemistry* domain, we open the way for classification models (not discussed in this article) to define specific ways to observe *ecology:Salinity* (i.e., establish the concrete sub-trait that applies to a chosen context). This occurs by observing and checking ranges of *chemistry:Salinity* that determine each category in context-specific ways (e.g. distinguishing brackish from fresh water; see 5 for practical examples of how similar models may be chosen, assembled and used).

To ease specification and enable inferences and functionalities, attribution in k.IM uses four categories of universals, collectively named *traits*, which correspond to different keywords used at

declaration (Table 1). We distinguish general *attributes* from more specialized *orderings* (whose subtypes define an ordered sequence), *realms* (which identify mereologically arranged subdivisions of a context, such as atmospheric strata) and *roles*, which categorize the ways specific observables are seen when in the context of another. Each of these categories enables specific types of inference in applications; roles, in particular, are crucial for interoperability in modeling applications, and deserve a discussion that is outside the scope of this article. A final category of universals, *identities*, is instrumental for the use and reuse of external vocabularies and terminologies, and is described in detail in the next section.

2.4. Bridging to accepted terminologies: Identities and authorities

In semantic annotation practice, it is common to encounter situations when an abstract observable (such as an individual animal, plant, or a material object such as a delimited volume of matter) must be identified by a “species”, such as a taxonomic or chemical one. For such situations, k.IM recognizes specific types of universals we name *identities*, which can be bound to observable concepts so that the use of a given identity type becomes mandatory to further specialize the observable:

```
namespace biology using physical;

agent Individual                                     (7)
  is physical:SelfAssertedBody
  requires identity Species;
```

In this case, the set of possible identities may be very large or even infinite. Since it is of course impractical to expect that ontologies can list all possible identities, this presents a problem when reasoning must compare concepts at two separate endpoints, as the identity used at one may not be known at the other. Having users create concepts for identities whenever a new one is needed would break interoperability, and the alternative - adding them to the shared worldview on an as-needed basis - would make the worldview prohibitively difficult to coordinate and maintain.

In such situations, we use **authorities** to link authoritative terminologies and ontologies. In k.LAB, authorities are software components that translate terms provided by authoritative terminologies, maintained by standard-defining organizations such as IUPAC for chemical nomenclature, into logical axioms that can be inserted into the namespaces provided in the worldview to create stable concepts that are available at all points of use. Authorities are identified in k.IM by names bound to a specific identity in a worldview:

```
namespace biology;

abstract identity Species                             (8)
  is Taxonomy
  defines authority GBIF.SPECIES;
```

This statement binds the GBIF.SPECIES authority to the *biology:Species* identity, requiring that any concrete *biology:Individual* is identified using it (based on definition 7, each *Individual* is in turn bound to adopting a *biology:Species* identity). For example, a

spatial coverage (e.g., a raster GIS dataset) describing the counted occurrences of honeybee individuals (*Apis mellifera*) per square kilometer could be annotated as follows:

```
model raster("data/bees.tif")
  as count biology:Individual identified as "1341976" by
GBIF.SPECIES per km²; (9)
```

Code 1341976 in the GBIF catalogue³³ is the identifier for the *Apis mellifera* species, tracking its unchanging taxonomic identity through any changes in nomenclature that may have occurred over time. For increased readability, definition (9) can also be written with a concept declaration that makes the identity explicit for a reader:

```
agent HoneybeeIndividual
  is biology:Individual identified as "1341976" by
GBIF.SPECIES; (10)
```

```
model raster("data/bees.tif")
  as count HoneybeeIndividual per km²;
```

In such situations, the user-defined concept (*HoneybeeIndividual*) functions as an *alias* for the GBIF honeybee concept, so that independent uses of the concept will not produce ambiguity, even if different specifications like (10) are given and different concept names are used in them. The two specifications (9) and (10) are functionally identical and compile to the same OWL axioms. Within the GBIF.SPECIES authority, producing logical axioms for the GBIF code 1341976 entails verifying that the code is a valid species identifier: a different outcome, such as using a non-existent or, e.g., a

family code, would result in a parsing error reported to the user. This mechanism guarantees the ability to reason across namespaces and allows full interoperability of taxonomic names when used at independent and uncoordinated endpoints. Multiple sub-authorities (such as GBIF.FAMILY, GBIF.CLASS, etc.) allow binding different classes of identifiers managed by the same organization. The GBIF web-accessible catalog service³³ provides codes that identify species and other taxonomic names in a stable and reliable way. It also provides metadata, such as labels, common names and broader terms, that are automatically linked to each concept created, allowing full specification of the identity and automated documentation of the resulting informational artifacts.

In addition to the identities managed by GBIF, representing the full taxonomic hierarchy from kingdom to variety, k.LAB provides authorities that recognize and interpret: (i) chemical identities (using the InChi naming conventions³⁴); (ii) soil taxa according to the World Reference Database nomenclature³⁵; and (iii) several classes of agricultural terms provided in AGROVOC²⁵ (Table 3). In most cases, authorities provide both validation of identifiers and search facilities, building on services provided by the managing institutions. For example, if a user refers to a chemical compound using a wrongly formatted InChi string, an informative error is reported. In contrast, a correct string can be translated by the IUPAC authority into a molecular diagram for the user to check. Availability of a specific authority within a worldview is equivalent to an *endorsement* of that authority in it. Authorities, complemented with search tools and validation, such as those provided in k.LAB, provide consistency and a sound annotation discipline in a usage landscape characterized by widespread redundancy and

Table 3. Authorities built in k.LAB and endorsed in the im worldview at the time of this writing. Each authority uses an external service or vocabulary and can provide one or more views that bridges to a specific type of identity. The concepts produced by authorities carry the URIs of the original concepts as metadata, when those are produced by the corresponding authority.

Authority	Views	Description
GBIF	GBIF.SPECIES GBIF.CLASS GBIF.PHYLUM GBIF.GENUS GBIF.ORDER GBIF.FAMILY GBIF.KINGDOM	Enables direct use in k.IM of the GBIF codes for the specific classes of identities handled by each view. These codes track taxonomic identities of different rank throughout any changes in terminology and nomenclature that may have occurred through time. Each code is validated using GBIF's web services and metadata are added to resulting concepts, including relationships to parent classifications. GBIF search services are used to provide search facilities built into the k.LAB software, to ease locating and using the GBIF identifiers in k.IM.
IUPAC	IUPAC	Enables direct use of InChi strings in k.IM, fully specifying molecular composition and structure for any chemical compound. The authority incorporates the excellent software support provided by IUPAC and related academic projects, so that the InChi strings can be located from within k.LAB, and validated. The resulting concepts' metadata may include other information and produce molecular drawings for the k.IM user to check.
SOIL	SOIL.WRB	Enables direct use in k.IM of the World Reference Base classification of soils, including bridging to the online vocabulary hosted by FAO, parsing and validation of complex soil taxonomies expressed as WRB classifiers.
AGROVOC	AGROVOC.CROP AGROVOC.PROCESS AGROVOC.SPECIES	Enables the direct use in k.IM of URIs or URI fragments from the AGROVOC vocabulary maintained by FAO. At the time of this writing, the three views listed enable access to terms related to crop types, agricultural processes and "commonsense" species identifiers used in agriculture, less specific and not interoperable with the precise taxonomies used in GBIF.

inconsistency. “Bridging” authorities, while not yet attempted, might also be designed to accept terms from one authority and turn them into the same axioms of another covering the same domain. For example, SOIL.USDA may in the future complement the existing SOIL.WRB authority as an alternative source of soil taxonomy identifiers, producing axioms compatible with the latter. This would enable transparent mediation of competing vocabularies and further expand opportunities for interoperability and reuse of existing annotated data.

2.5. Using the worldview: annotation of data and models

With a common phenomenology, a structured language and validating supporting infrastructure, knowledge engineers can create worldviews with better prospects of consistency, expressiveness and reusability. Yet, the task of building a worldview remains daunting. We can consider the building of the worldview a **Tier 3** activity, requiring significant expertise, long-term research investments and a careful vetting process involving consultation and continuing collaboration with a large number of experts. We will briefly summarize challenges and successes of worldview development for ARIES in the discussion and better discuss the topic in forthcoming contributions.

When a suitable worldview is available, it should become possible to compose domain semantic annotations by combining existing concepts from the worldview and its endorsed authorities. We can consider this **Tier 2** of difficulty in semantic annotation; it is the scope of many initiatives of which a representative example is the Agrisemantics initiative³⁶. The Agrisemantics vision statement mentions *height of corn* as an archetypical example of a common observable whose interoperability for existing data resources is desirable. A fitting ontology available in the OBO foundry²⁹ that provides concepts adequate for this task is the Plant Trait Ontology (PTO,³⁷), which draws on the work of many experts and enjoys good community acceptance. The PTO provides a hierarchy of concepts starting at *quality* (imported from the BFO ontology, also a base ontology for k.LAB) specialized to *morphology*, then further into *size height* and *plant height*. One can assume that identifying *height of corn* would require a further specialization of *plant height*, and the corn identity would simply be implied syntactically by using a *Corn...* prefix in the term assigned to the concept. Further exploration of the PTO reveals that *giant embryo* (a gene type) is a sibling of *plant height*, both specialized from *whole plant size* through *is-a* inheritance. Further *is-a* specialization of *plant height* defines, among others, concept *plant height uniformity* (a quality not physically commensurable with height) and *relative plant height* (seemingly adding an observation-related attribute, *relativity*, out of many possible). The PTO is one of the most advanced domain ontologies in use with respect to phenomenological characterization, and its terms have proved useful to large communities. Yet it is clear from this example that no ontology can *force* users to adopt cogent annotation practices, ensuring that physical and biological identities are preserved along inheritance chains and attributes retain traceable and stable meaning. These are key requirements to help prevent inconsistencies and better assist annotation in service of the FAIR goals. If the same exercise were replicated in k.IM, for example to annotate a raster map file describing corn height in cm in a given

region, the language itself would have driven the specification of the semantics:

```
model raster("data/cornheight.tif")
  as measure im:Height
    of agriculture:PlantIndividual identified as "c_12332"
by AGROVOC.CROPS
  within agriculture:CropField
in cm; (11)
```

Or for increased readability:

```
agent CornPlant is agriculture:PlantIndividual identified
as "c_12332" by AGROVOC.CROPS;

model raster("data/cornheight.tif") (12)
  as measure im:Height of CornPlant within agriculture:CropField
in cm;
```

The *measure* (observable) *in* (unit) syntax (see below and Table 4), one of k.IM’s *observer* statements (Table 4), embodies the semantics for the *how* of observation discussed above, and requires that the primary observable, in this case *im:Height*, be a physical property, simultaneously enforcing the use of units of measurement appropriate for its physical nature. Definitions (11) and (12) intentionally use *agriculture:PlantIndividual* instead of *biology:Individual*, as the latter requires a precise species identity (definition 7), while the former references the commonsense taxonomy used in AGROVOC for crop types, reflecting the intended semantics for the data. Most importantly, the adoption of rigorous phenomenological inheritance and specification syntax requires the realization (and the explicit statement) that *height* is first of all a quality *of* a plant subject, and that the data refer to plants *within* a cropfield subject. These logical axioms are a necessary base for any reasoning that can assess their compatibility within applications. While such details still need to be learned by a user, the syntax itself serves as a guide for the annotation workflow: the use of inconsistent observables or the lack of proper inherency would yield ungrammatical statements that are reported as errors. For example, leaving out the **of** specification would cause height to become abstract, therefore not usable for data annotation; leaving out **within** would leave the context of inherency for the quality blank, reported as an error for any non-abstract quality. The result is readable by a non-expert and compiles to axioms specifying a single OWL concept, which can be transferred to a remote endpoint in axiomatic form and reconstructed for reasoning or database querying; the shared worldview is the only requirement for its interpretation. The concept as constructed carries information about physical nature, dimensionality, domain of application, agricultural identity, biological identity and context of inherency (plants within a cropfield). These are assembled through consistent logical restrictions and are robust to validation and machine reasoning. On this basis, inferences can be performed that use the annotated dataset to satisfy queries beyond the asserted quality, e.g. for *presence of CornPlant* as discussed previously. Simply through reasoning on the concept, a query for an observation of the height of generic *agriculture:PlantIndividual* in any *earth:Region* (of which *agriculture:CropField* is a subtype) could be satisfied, in absence of a more specific match, by the same corn height data.

Table 4. Observer types from the annotation of qualities in k.IM, with a description of the observation semantics implied.

These statements are used in data and model annotation (as opposed to ontology definition) to express either data semantic or model dependencies; when necessary, they automatically apply the semantic operators of Table 2 to build the correspondent concepts. In addition, they specify observation semantics (such as units, currencies or categories) so that the concepts can be associated to specific data values and mediated when necessary.

Prototype	Description
measure <O> in <unit>	Specifies the unit for a concrete physical property and ensures that it is compatible with the physical nature and the spatiotemporal context of use. Ensures that units are converted when dependencies are matched to data.
rank <O> [min to max]	Used with priorities, can specify a scale for bounded ranks and ensure that scales are properly converted when dependencies are matched to data.
classify <O> [into <O1> [if <condition>], O2, ...>] [according to <metadata field>] [as identified by <authority>]	Used with class concepts, enables many useful ways of specifying the semantic content of categorical classifications; in addition to the direct specification of the concepts that each possible value or range of values should map to, it allows specifying metadata for conversion (e.g. the standard encodings of categories in common land cover datasets) and to match values to concepts by converting identifiers through a specified authority.
value <O> [over <O2>] [in <currency>] min to max]	Values can be direct or relative (an example of the latter is the pairwise comparisons used in multiple criteria analysis) and refer to a currency (monetary or conceptual) or have a scale like in the case of rankings. When the currency is monetary, a year must also be specified; k.LAB contains functionalities that bridge to conversion services so that values can be adjusted for inflation and converted to different currencies in many cases.
distance to <O> in <length unit>	A distance observer will observe all the objects of the type mentioned in the context of observation and compute the distance to them. In k.IM, this observer can also be used with reference to the URI of a specific observation, which can be located anywhere.
count <O> [per <extent unit>]	Count observers observe all the objects of the type mentioned and produce their numerosity, if necessary distributed over space and/or time. A count concept is produced unless O is already a count.
ratio [of] <O1> [to <O2>]	Ratio observers describe ratios between qualities. A ratio concept is produced unless O is already a ratio.
proportion [of] <O1> [in <O2>] percentage [of] <O1> [in <O2>]	Proportion and percentage are differently scaled ways to observe a proportion concept, which is created according to rules in Table 2 unless O1 is already a proportion.
uncertainty [of] <O>	The numeric scaling and computation of uncertainty is not mandated in k.IM. In k.LAB, currently, numeric uncertainties are computed as standard deviations of probability distributions, and the Shannon index of diversity is used for categorical information.
probability [of] <O>	Probability observers validate their data in the [0-1] interval. A probability concept is produced unless O is already a probability.
occurrence [of] <O>	This observer is a "fluent" shorthand to specify the probability of a presence.
presence [of] <O>	Validates data as boolean (true/false). A presence concept is produced unless O is already a presence.

The simplest examples of usage (**Tier 1**) exploit pre-defined concepts from the worldview to annotate resources. In such cases, knowledge of the syntax and simple search tools allow a user to produce annotations that can accompany informational assets for automated discovery and indexing:

```
worldview im;
model raster("elevation.tif") as measure geography:Elevation
in m; (13)
```

Specifications such as (13) are simple enough to be added to metadata or "sidecar files" – files with a ".kim" extension that accompany data files with the same name - which may be automatically detected and indexed by specifically designed web crawlers, so that indexes

of web-accessible, annotated datasets can be built and maintained. Specification (13) is complete and correct, as *geography:Elevation* is fully characterized in terms of inherency within the worldview, as seen in statement (1). The statement of the worldview name is enough to load the web-accessible worldview and use it to interpret the specification that follows. To annotate qualities, which encompass a majority of data artifacts, the syntax for *observer* statements (such as **measure** in the example above) is enough to represent all *observation semantics* known to k.IM. The set of available observer statements (Table 4) is small and has proven easy to learn and use in ARIES coursework and test user communities.

While this article does not fully describe modeling and annotation features of k.IM and k.LAB, we note that data annotation is not

restricted to qualities. Subjects can also be annotated with ease using a slightly different syntax:

```
model each vector("data/roads.shp") as
infrastructure:Road; (14)
```

As subjects are observed directly (without needing units or other known observations for comparison), the simple acknowledgement of the semantics is enough to annotate a source of objects, such as roads in a vector file. The keyword **each**, only applicable to countable observables, reflects the fact that such sources can produce one or more subjects with the specified semantics; the **model** statement also allows annotation of semantics for any attributes of the observed subjects (not shown). The k.IM language and k.LAB infrastructure build on these semantic foundations to enable a distributed modeling infrastructure, in which the resulting observations can be complemented, through further extension to the **model** syntax, with procedural information to create “live” observations interacting on a networked infrastructure, in compliance with their semantics. Such features, briefly described for the ARIES application in 5, will be more thoroughly illustrated in forthcoming contributions and documentation.

2.6. Assessing compatibility

FAIR+ interoperability requires the unsupervised assessment of *compatibility* between semantically annotated resources. We use the term *compatibility* to refer to concepts and *interoperability* to refer to observations of compatible observables. In k.LAB, compatibility enables interoperability in two fundamental ways:

1. Validation of connections, for example in ensuring that a model’s dependent observables are compatible with the computed output in terms of inherency;
2. Discovery and retrieval of compatible observables for queries stated only through their semantics, so that the best source of information (data or model) for a required observable can be located on the network when requested by users or models being computed.

Using the notion of interoperability illustrated in Section 1.2 and the semantic foundations illustrated so far, the assessment of compatibility for interoperability can be defined as follows.

Two **observables** (O1, O2) are compatible if and only if:

- The main observable concept in O1 (without considering traits and inherency) equals, or is a more specialized version of, the main observable in O2;
- O1 adopts all the same traits and roles as O2 (which may have additional traits); e.g. the main observable in O1 and O2 may be a generic length, but if O1 is vertical, O2 must also be;
- If O1 has an inherent type (**of**), O2 must have a compatible one;
- If O1 has a context type (**within**), O2 must have a compatible one.

If observables are compatible, their **observations** are interoperable. They are FAIR+ interoperable if and only if:

- Their observables are compatible;
- Their observation semantics can be mediated (e.g. both are measurements in compatible, but possibly different units);
- Their context can be mediated: the intersection of the extents (e.g. space, time) of the scale for both observables is non-empty and the resolution of each extent is the same as, or can be resampled to fit, the other’s.

This definition is amenable to being incorporated in an unsupervised algorithm. Mediation may engender information loss (e.g. aggregation error) and other uncertainties (e.g. when bridging different classification systems), which should be recorded as provenance³⁸ in separate records kept with the dataflow. In queries, when more than one interoperable observation may be returned, any potential information loss can become part of the criteria used to rank the appropriateness of each candidate observation that matches the observable. On a semantic level, the match may also be incomplete. For example, some traits of the matching observation may not be stated in the query, e.g. a vertical length could match an unspecified one. This offers a base to develop ranking strategies considering, among other criteria, metrics of semantic accuracy or distance; the latter is an important criterion in k.LAB and will be discussed in detail in further contributions.

3. Discussion and perspectives

Distributed databases with their contents annotated according to a common worldview can allow the kind of large-scale, yet precise, semantically-driven interoperability that has so far remained a high-ranking wish in the semantic web community. SMM, a modeling approach where FAIR+ interoperability is an integral requirement, sees data and models as *definitions* for possible observations: while datasets can produce, possibly through mediation of observation or context semantics, the requested observations in a self-contained way, models do so through computation that may involve the observation of other concepts they depend on, to be *resolved* through other data or models. Distributed databases of k.IM-annotated data and models can be built using k.LAB and accessed through modern web services³⁹ with distributed, certificate-based authentication. These services form an operational semantic web whose nodes contain FAIR-compliant scientific observations and models. In forthcoming contributions, we will describe the ways that k.LAB enables the assemblage, validation and computation of scientific workflows that observe an arbitrary user-requested concept in a user-defined context. These functionalities have been informally described in the context of the ARIES project⁵.

While the coverage and scale of our applications so far remains too small to warrant claims of large-scale success, our experience with ARIES indicates that building such distributed knowledge bases is possible and practical. Large-scale initiatives, such as NEON⁴⁰, CSDMS⁴¹, directives such as INSPIRE⁴² and many others, are seeking interoperability of data, increasing requirements and initiatives for data openness and publication of data⁴³, and implementing new data release standards that emphasize accessibility of

information. Approaches that can facilitate the development of consistent semantics beyond textual metadata and controlled vocabularies become essential. FAIR criteria outline a way to gather all those observations in a way that will greatly advance science synthesis⁴⁴. Faced with a state of the art in which semantic interoperability is still often understood as “matching of terms”⁴¹, we argue that the semantic research and infrastructure available to date are still not ready for a FAIR+ interpretation of interoperability, and propose the work presented here as a contribution towards it.

At the time of this writing, the k.LAB infrastructure and the *im* worldview are used to annotate datasets and models numbering just below one thousand, and have been exposed to about 150 users, of which only about 20 use it for their daily work. These are very small numbers compared to the ambition of open data and the importance of interoperability in scientific discourse. Our experience in ARIES has highlighted both strong and weak points in the attempt of creating a systematic and accessible path to rigorous semantic annotation for practitioners. Advantages recognized by the user community are:

- Clarifying the **components of interoperability**, so that conceptualization efforts are focused and a suitable workflow can be identified. The most important aspect in this sense is the clear focus on observable semantics: no time is wasted seeking semantics to express model-related concepts (“model”, “variable”), observation-related ones (“measurement”) or context-related ones (“spatial resolution”), all of which figure prominently in commonly used ontologies.
- Formalizing a **simple phenomenology** for observables and universals. The base observables in [Figure 1](#) have proven intuitive enough to be understood and remembered by diverse users, helping them “home in” quickly on observable semantics as described in the previous point. Also, the use of independently defined and flexibly attributed universals to express attributes, identities and roles has effectively and intuitively solved, in our applications, the plaguing issue of excessive and improper specialization.
- The **k.IM language** and **k.LAB platform** make ontologies and annotations immediately actionable, enforcing the logical consistency of each definition both by enforcing syntactical correctness through intelligent editing tools and by employing a machine reasoner⁴⁵ to identify and report logical errors to the user. The language guides, simplifies and validates the definition of knowledge; the support software provides feedback and allows users to immediately perform user queries and compute workflows whose results enable at-a-glance validation of the semantic correctness of the concepts employed.

At the same time, clear difficulties remain in instrumenting a path to large-numbers adoption of an approach like the one we propose. For example, the use of a custom language to specify ontologies has disadvantages: the choice was inevitable for us due to the need to

reach *large bases of* users other than knowledge engineers, but connecting to semantic web research and communities with a custom approach is of course much more difficult despite our commitment to OWL2. Another important difficulty is the need for complex, custom software to make the approach actionable, with obvious costs and difficulties related to its development, distribution and maintenance.

Finally, building and sharing worldviews that reflect large and complex domains remains a daunting task, despite the guidance of a systematic conceptual framework and methodology. In particular, developing a collaborative process to ensure that the worldview reflect the uncontroversial thinking of large communities requires both large collaboration investments and sophisticated tooling for harmonization and refactoring. Despite the success of mid-size initiatives like ARIES, we are at the very beginning of an ambitious effort whose challenges may well prove too large for large-scale adoption.

In our mind, these difficulties are offset by the potential for the collaborative, wider use of scientific products that would be enabled by such a rigorous, semantically-driven interoperability. The ability to automatically discover and compute dataflows based only on conceptual queries opens pathways that may lead to much larger use of scientific products, with a potentially much larger involvement of decision-makers and citizen scientists. Our efforts are sustained and motivated by the realization of the potential of effective, actionable interoperability to promote and enable a more efficient economy of knowledge, creating clear incentives to the sharing of data and models, so that they may become part of large and yet undiscovered computational chains.

Software availability

The k.LAB software is available in **source form from Bitbucket** and in binary form from the **Integrated Modelling collaboration site**.

Author contributions

FV is responsible for the original SMM vision, led the development of the approach, designed the k.IM language and the supporting k.LAB infrastructure, and wrote the majority of the text. SB and IA participated to the development of the approach, led crucial case studies, contributed to the design of the semantic principles, the software implementation and to the writing. CC contributed to the work on authorities, provided experience with philosophical underpinning, provided context for agricultural applications, and contributed to the writing.

Competing interests

No competing interests were disclosed.

Grant information

Development was originally funded by the US National Science Foundation (grant 9982938) and received partial support from the ASSETS and WISER projects funded by ESPA/NERC (grants

NE-J002267-1 and NE/L001322/1). ARIES is developed with partial support from the EU-Horizon 2020 project AQUACROSS (grant agreement no. 642317).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

Ken Bagstad, Simon Willcock and Mikel Egaña Aranguren provided advice and feedback on early drafts. Many participants to

the ARIES project and the International Spring University on Ecosystem Services Modeling provided valuable feedback and testing on the approach, the k.IM language and the software infrastructure over almost a decade. Discussion with the Global Agricultural Concept Scheme (GACS) working group (in particular Johannes Keizer, Thomas Baker, Devika Medalli, Elizabeth Arnaud, Medha Devare, Sophie Aubin) also greatly helped focus, shape and improve the details of the approach. Giovanni L'Abate drove the development of the SOIL.WRB authority.

References

- Kumazawa T, Saito O, Kozaki K, *et al.*: **Toward knowledge structuring of sustainability science based on ontology engineering.** *Sustain Sci.* 2009; **4**: 99. [Publisher Full Text](#)
- Wilkinson MD, Dumontier M, Aalbersberg IJ, *et al.*: **The FAIR Guiding Principles for scientific data management and stewardship.** *Sci Data.* 2016; **3**: 160018. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Jain P, Hitzler P, Sheth AP, *et al.*: **Ontology alignment for linked open data.** *International Semantic Web Conference.* Springer; 2010; 402–417. [Publisher Full Text](#)
- Ludäscher B, Lin K, Bowers S, *et al.*: **Managing scientific data: From data integration to scientific workflows.** *Geol Soc Am Spec Pap.* 2006; **397**: 109–129. [Publisher Full Text](#)
- Villa F, Bagstad KJ, Voigt B, *et al.*: **A methodology for adaptable and robust ecosystem services assessment.** *PLoS One.* 2014; **9**(3): e91001. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Millennium Ecosystem Assessment: **living beyond our means - Natural assets and human well-being.** [Internet]. 2005. [Reference Source](#)
- Berners-Lee T, Hendler J, Lassila O.: **The semantic web.** *Sci Am.* 2001; **284**(5): 34–43. [Publisher Full Text](#)
- Antoniou G, Van Harmelen F: **A semantic web primer** [Internet]. MIT press; 2004. [Reference Source](#)
- Guarino N: **Formal ontology and information systems.** *Proceedings of FOIS.* 1998; 81–97. [Reference Source](#)
- Villa F, Athanasiadis IN, Rizzoli AE: **Modelling with knowledge: A review of emerging semantic approaches to environmental modelling.** *Environ Model Softw.* 2009; **24**(5): 577–587. [Publisher Full Text](#)
- Porter CH, Villalobos C, Holzworth D, *et al.*: **Harmonization and translation of crop modeling data to ensure interoperability.** *Environ Model Softw.* 2014; **62**: 495–508. [Publisher Full Text](#)
- Goguen JA: **Data, schema, ontology and logic integration.** *Log J IGTL.* 2005; **13**(6): 685–715. [Publisher Full Text](#)
- Keet CM: **The use of foundational ontologies in ontology development: an empirical assessment.** *Extended Semantic Web Conference.* Springer; 2011; 321–335. [Publisher Full Text](#)
- Gangemi A, Guarino N, Masolo C, *et al.*: **Sweetening ontologies with DOLCE.** *International Conference on Knowledge Engineering and Knowledge Management.* Springer; 2002; 166–181. [Publisher Full Text](#)
- Arp R, Smith B, Spear AD: **Building ontologies with basic formal ontology** [Internet]. Mit Press; 2015. [Reference Source](#)
- Mascardi V, Cordi V, Rosso P: **A Comparison of Upper Ontologies.** *WOA.* 2007; 55–64. [Reference Source](#)
- Pease A, Niles I, Li J: **The suggested upper merged ontology: A large ontology for the semantic web and its applications.** Working notes of the AAAI-2002 workshop on ontologies and the semantic web. 2002. [Reference Source](#)
- Madin J, Bowers S, Schildhauer M, *et al.*: **An ontology for describing and synthesizing ecological observation data.** *Ecol Inform.* 2007; **2**(3): 279–296. [Publisher Full Text](#)
- Cox SJ: **An Explicit OWL Representation of ISO/OGC Observations and Measurements.** *Proceedings of the 6th International Conference on Semantic Sensor Networks.* Aachen, Germany; CEUR-WS.org; 2013; **1063**: 1–18. [Reference Source](#)
- Raskin RG, Pan MJ: **Knowledge representation in the semantic web for Earth and environmental terminology (SWEET).** *Comput Geosci.* 2005; **31**(9): 1119–1125. [Publisher Full Text](#)
- Buttigieg PL, Morrison N, Smith B, *et al.*: **The environment ontology: contextualising biological and biomedical entities.** *J Biomed Semant.* 2013; **4**: 43. [Publisher Full Text](#)
- Grenon P, Smith B: **SNAP and SPAN: Towards dynamic spatial ontology.** *Spat Cogn Comput.* 2004; **4**: 69–104. [Publisher Full Text](#)
- Ashburner M, Ball CA, Blake JA, *et al.*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet.* 2000; **25**(1): 25–29. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ilic K, Kellogg EA, Jaiswal P, *et al.*: **The plant structure ontology, a unified vocabulary of anatomy and morphology of a flowering plant.** *Plant Physiol.* 2007; **143**(2): 587–599. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Caracciolo C, Stellato A, Morshed A, *et al.*: **The AGROVOC linked dataset.** *Semant Web.* 2013; **4**(3): 341–348. [Publisher Full Text](#)
- Hood MW, Ebermann C: **Reconciling the CAB Thesaurus and AGROVOC.** *Q Bull IAALD IAALD.* 1990. [Reference Source](#)
- Tarboton DG, Horsburgh JS, Maidment DR: **CUAHSI community Observations Data Model (ODM) version 1.1 design specifications.** *Des Doc.* 2008. [Reference Source](#)
- Nelson SJ, Johnston WD, Humphreys BL: **Relationships in medical subject headings (MeSH).** *Relationships in the Organization of Knowledge.* Springer; 2001; **2**: 171–184. [Publisher Full Text](#)
- Smith B, Ashburner M, Rosse C, *et al.*: **The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration.** *Nat Biotechnol.* 2007; **25**(11): 1251–1255. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Armstrong DM: **A theory of universals: Universals and Scientific Realism** [Internet]. CUP Archive; 1978; **2**. [Reference Source](#)
- Jowett B, others: **The republic of Plato.** Clarendon press; 1888. [Reference Source](#)
- Motik B, Patel-Schneider PF, Parsia B, *et al.*: **OWL 2 web ontology language: Structural specification and functional-style syntax.** *W3C Recomm.* 2009; **27**: 159. [Reference Source](#)
- Free and Open Access to Biodiversity Data | GBIF.org** [Internet]. [cited 3 Mar 2017]. [Reference Source](#)
- Heller S, McNaught A, Stein S, *et al.*: **InChI - the worldwide chemical structure identifier standard.** *J Cheminform.* 2013; **5**(1): 7. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Food and agriculture organization of the united nations: **World reference base for soil resources.** *World Soil Resour Rep.* 1998; **84**: 21–22. [Reference Source](#)
- Baker T, Caracciolo C, Doroszenko A, *et al.*: **GACS Core: Creation of a Global Agricultural Concept Scheme.** *Metadata and Semantics Research: 10th*

International Conference, MTSR 2016, Göttingen, Germany, November 22–25, 2016, *Proceedings*. Springer; 2016; 311–316.

[Publisher Full Text](#)

37. Yamazaki Y, Jaiswal P: **Biological ontologies in rice databases. An introduction to the activities in Gramene and Oryzabase.** *Plant Cell Physiol.* 2005; **46**(1): 63–68.
[PubMed Abstract](#) | [Publisher Full Text](#)
38. Simmhan YL, Plale B, Gannon D: **A survey of data provenance in e-science.** *ACM Sigmod Rec.* 2005; **34**(3): 31–36.
[Publisher Full Text](#)
39. Pautasso C, Zimmermann O, Leymann F: **Restful web services vs. big web services: making the right architectural decision.** *Proceedings of the 17th international conference on World Wide Web.* ACM; 2008; 805–814.
[Publisher Full Text](#)
40. Hampton SE, Strasser CA, Tewksbury JJ, *et al.*: **Big data and the future of ecology.** *Front Ecol Environ.* 2013; **11**(3): 156–162.
[Publisher Full Text](#)
41. Peckham SD, Hutton EW, Norris B: **A component-based approach to integrated modeling in the geosciences: The design of CSDMS.** *Comput Geosci.* 2013; **53**: 3–12.
[Publisher Full Text](#)
42. Directive I: **Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).** *Publ Off J.* 2007.
[Reference Source](#)
43. Molloy JC: **The open knowledge foundation: open data means better science.** *PLoS Biol.* 2011; **9**(12): e1001195.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
44. Peters DP, Havstad KM, Cushing J, *et al.*: **Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology.** *Ecosphere.* 2014; **5**(6): 1–15.
[Publisher Full Text](#)
45. Poole DL, Mackworth AK, Goebel R: **Computational intelligence: a logical approach [Internet].** Oxford University Press New York; 1998.
[Reference Source](#)

Open Peer Review

Current Referee Status:



Version 1

Referee Report 11 December 2017

doi:10.5256/f1000research.12570.r22819



Pier Luigi Buttigieg 

Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

General comments:

In this contribution, the authors present part of a system to infuse models and the data linked with them with semantic content through a high-level programmatic interface, bringing digital content more in line with the FAIR guiding principles. The manuscript presents some encouraging progress and clearly represents a formidable effort, however, does reveal some concerning (but certainly addressable) issues with the system and the ontological framework it depends upon.

In my opinion, the strongest component of this contribution is the consolidated and practical approach to adding semantics to the toolbox of modellers. To me, material like code snippet (14) represents a hugely exciting and powerful advancement; I would be much more supportive of this submission if it focused primarily on these capabilities. The transparent and systematic handling of naming authorities and connection to internationally endorsed vocabularies is also promising, as is the care taken to prevent these conventions from dictating more rigorous semantics. As stated in the MS, the potential to use semantics to bridge these resources is immense and the authors' work in this direction is of high value.

I believe the weakest part of this work is the semantic framework described within it. I couldn't find the ontologies themselves, so evaluation is based only on the descriptions in the manuscript. Frankly, I find the approach to handling "observables" haphazard and something of an uncomfortable shortcut relative to semantics developed in, e.g., the Biocollections Ontology (BCO). Initially, such shortcuts may appear to relieve a burden of using semantic technologies on end users, however, this quickly becomes uncontrolled. Further, in my experience, the commitment to realism in ontology (i.e. avoiding dealing with concepts and focusing on empirical phenomena) development is essential in the scientific realm to prevent rapid semantic drift between initiatives. The treatment of scale here seems to lend itself to this issue, unless the semantics reflect that scales overlap (I couldn't find any evidence of this). Expecting a coherent "common sense" approach to prevail rarely bears fruit across multiple groups without concentrated effort at parsing out their knowledge into some mature upper level framework.

Further, I'm left wondering why there doesn't seem to be a developed attempt at reuse of ontologies such as BFO, DOLCE, or other upper level resources. The authors mention that they use BFO, but the framework they propose seems to counter many of its basic semantics without clarification or logical justifications. I'm afraid I can't really understand how the reuse of these foundational and well-adopted semantics is actually implemented or if it's defensible. That could be a paper in itself and it probably should be if the authors want to defend their proposals more completely: in this format, it's very hard to endorse the proposals. I highlight a few specific issues in my detailed comments below.

Much the same can be said for the use of mid-level ontologies like the TO and ENVO in which the authors point out some important flaws/limitations. Generally speaking, the editors of such resources are happy to receive feedback and coordinate with external systems rather than forcing the duplication of work. Were there attempts at engagement via issue trackers or similar? The lack of engagement would be odd given how much the authors emphasise FAIRness. Perhaps a clarification of how, exactly, these resources were reused and what their future role in the ARIES system will be (assuming they are willing to cooperate with its development) would help.

Additionally, there are some descriptions and claims (e.g. that k.IM makes things "immediately actionable", more below) which need some form of support or demonstration to be substantiated. This need not be exhaustive: a few illustrative examples as a supplement should suffice. If no such demonstration is provided, I recommend omitting these claims.

Overall, much of this contribution is more a specification rather than a report on results and outcomes. I suggest large parts of this (most tables and statements) be moved into a supplement and a few used to illustrate the key points made.

Despite these critiques, I am very supportive of this work. This is one of the most streamlined semantics-to-modelling systems I've encountered and I'm hopeful that it can iron out the issues outlined above. Further, I echo the authors' call for a more consolidated semantic landscape to allow such technologies the stability they need to grow efficiently. Parallel efforts should begin converging and complementing one another (e.g. OBO and other such resources providing worldviews for ARIES while ARIES reshapes their content when inconsistencies are found). I'd be very happy to support this moving forward

Detailed comments:

Below, please find a non-exhaustive set of specific comments to clarify and expand on my positions above. I don't require a response to each one, but the major points they collectively describe should be addressed in a revision.

Worldview: To me, a worldview sounds more or less like a "domain ontology". Is this the case? If so, how are overlaps between domains handled?

I feel uncomfortable with some of the syntax used in the various code snippets: e.g. how is a model a measure? This is not really intuitive and requires commitment to an unusual abstraction. This reflects some of the semantic ambiguity at the foundation of the ARIES/kIM system discussed above.

While I certainly can appreciate the effort involved, many years of design and user feedback do not necessarily translate, 1:1, into quality: I would tone down suggestions to this effect. This is especially true as the user/testing group is rather small (which doesn't mean the feedback was poor, but reduces the weight nonetheless). Simply stating the begin date of the work and the number of active testers at the beginning and at the end would be enough.

Regarding the statements regarding OWL/OWL2: It seems that authors are proposing that those working with OWL or using other upper level ontologies should either 1) switch to the ARIES system or 2) maintain translators between OWL2 and ARIES-compliant languages. This is quite a bold position for a relatively new and untested system. I would suggest that this should be the other way around for the early phases: ARIES should develop translation layers to the establish technologies.

Perhaps I misunderstand, but I find it quite odd that the authors criticise other ontologies for requiring commitment to a given system of operation while simultaneously recommending that users subscribe to the IM system in the first few lines of discussion. As with any technology, users must understand the tools they work with and, indeed, commit to some sort of convention.

What is promised in the forthcoming contributions is quite ambitious and exciting, but has little relevance to this submission and, naturally, no evidence to support the claims made. I thus suggest this be toned down considerably. An "operational semantic web" that is able to pull data together into a coherent data set for immediate analysis would indeed be very useful. There are application-specific examples of this in the biomedical domain (e.g. the Monarch Initiative). However, simply being linked in such a network does not make data and models FAIR-compliant. How do you know they are really interoperable and reusable unless the contributors of those datasets have made them so? k.LAB cannot magically make this happen, or can it? I think a careful look at what FAIR specifies is needed before claims of compliance are made.

"Faced with a state of the art in which semantic interoperability is still often understood as "matching of terms"" By whom? It is true that many in the community tragically think of semantics as terminological mapping, but there are quite a few others that really do target the meaning behind terms. I think that statements like this don't do the community justice and give the false impression that the SMM/k.LAB solution is the only one addressing true semantics.

"In our view, a major need for progress towards this goal is a solid, uncontroversial phenomenological base, i.e. the basic semantics for the types of phenomena and entities that can be understood by human observers."

I'm afraid that controversy is not likely to be dispelled: there will always be disagreement as human understanding is continually changing and growing, with researchers and other observers frequently disagreeing. Statements like this presuppose some form of authority which is antithetical to open progress. I would suggest that systems should be able to handle controversy while offering systemic stability.

"Formalisms and toolsets must be built to support it, to ease the specification of domains and allow for extension, while enforcing a consistent design discipline. We need clear best practices for specialization and connection of terms, and guidelines on how to integrate always growing, and potentially infinite, domain content from vocabularies without breaking the logical integrity of the resulting annotations." This sounds more or less identical to the objective and actual modus operandi of the OBO Foundry and Library. I would suggest that OBO Principles be discussed here.

I take strong exception to how concepts are nearly equated to the objects of empirical observation – in my experience, this quickly becomes toxic to handling the informatics of natural science. This also doesn't seem to resonate with how the authors treat particulars like real-world entities across the manuscript. This and superficial treatments of ideas like "Platonic realism" concern me. If particulars are concepts, how can they maintain their identity through time unless they are being conceptualized? What is exactly mean when the authors discuss the observation of a concept? I think such claims and arguments should be the put forth in a separate paper reviewed by logicians and developers of upper level ontologies. The value of this paper is more the implementation of the semantics-to-model link.

"If a physical object, event, process or relationship can be simply acknowledged to exist in a context of interest, qualities, such as the elevation of a mountain or the temperature of a body, can only be observed indirectly, i.e. by comparison with reference observations."

I don't think this works - the same can be said for the physical objects themselves. Every distinction is

based on comparison to one or more points of reference. The idea of indirect observation of a quality but direct acknowledgement of a physical object, event, process, or relationship is also occult to me. Further, I see no meaningful distinction between a process and event save an arbitrary fiat boundary.

I recognise the importance of handling entities across scales, but the arguments offered don't really convince me. Formalising scale semantics independently of the "observables" seems off: the observables don't disappear or cease to influence a system simply because the scale changes. Indeed, the scales being handled are anchored to the observables identified in a given application. This is true even if they are not "directly" observable. I'm not sure how the authors' handling of scale is useful as it sounds like something that would actually amplify bias. Some sort of demonstration of why this approach is more useful is needed.

"The ability to accurately characterize semantics along observable, observation and context dimensions addresses the interoperable and reusable FAIR criteria. Semantic specifications can be rewritten into queries that select interoperable counterparts for an observation, addressing the findable and accessible requirements."

Looking at the description of the FAIR criteria, I wouldn't say that they've been addressed so easily. Semantic technologies are of course needed to make data FAIR (especially the "I" criterion), but FAIRness is something that can only be evaluated in action, not in theory. I would weaken these claims until a demonstration is provided.

"While observation and context semantics are relatively well-understood, the characterization of what things are - observable semantics - remains difficult and uncertain, even with increasing investments in ontologies and vocabularies and an engaged community behind the current state of the art."

Such a statement would need more support - how are these levels of understanding assessed by the authors? Some more clear arguments would go a long way here.

"Terms describing commonly acknowledged classes of physical entities (such as persons or objects) are complemented through inference, comparison, association and imagination...Such observable entities..." These entities cannot be observable if they are imagined. The argumentation gets somewhat murky here. Or is imagination meant in terms of the creativity of the knowledge engineer? This is slightly less worrying.

"It follows that interoperability can exist in a conceptualization, as long as the boundaries of stability of meaning for all concepts with respect to their fundamental phenomenology are stable. Scale, commonly defined as the choice of resolutions and extents through which we make observations of the world, binds the observables of informational artifacts to precise phenomenological categories, establishing boundaries of validity for conceptualizations."

Ideas like "fundamental phenomenology" are dubious: many ontology developers have been in search of this Holy Grail, but we can only deal with what empirical studies have predictably and repeatedly confirmed. Knowledge and meaning are maintained around these anchors. Further, scale is typically imposed by observers unless one considers the scale associated with an entity itself. I don't see how this connects to the validity of a conceptualisation. Is it not valid to assert that roads are present in the European continent even though they can't be observed due to coarse scale? Indeed, the idea that continuants themselves can be morphed is problematic: shifting perspective doesn't make a continuant cease to be. This perceiver-based semantic basis seems counterproductive and no results are put forth to support arguments that it is superior to alternate modes of handling scale.

I really don't see how subjects can morph into qualities. Alluding to the cyanobacteria in water, the cyanobacteria haven't disappeared or changed, one is simply talking about the green quality of the chlorophyll in these cells integrated over the expanse of a different entity (a lake, pond, etc). A rigorous

semantic solution would track this over multiple scales, rather than 'fudge' it and directly claim the water body is green.

There needs to be simplification of how "structural" and "functional" relationships are discussed. This is nothing more than differentiating instantaneous (in the SNAP logic) relationships from those that require a temporal window to be realised (e.g. realisations of functions and dispositions in the BFO world). Indeed, why do we need yet another way of doing something that is done by other upper level ontologies? If the authors must re-invent this, they should, at the very least, make the relationships to existing upper level resources clear and argue why their approach is superior (again, I think that should be a different paper).

"At the same time, it became clear that no community of modelers, data scientists or other prospective users would consider an investment in OWL or other semantic web-endorsed formalism as the vehicle to express the semantics in data and models"

Really? Why not? Why do they even need to know that OWL is being generated in the background? Why abandon a functional system because some unspecified community of users won't interact with it? Should we abandon HTML5 because average web users don't understand how to write it? Obviously not, it's just hidden and tools that generate it created (e.g. CMS tool).

Claiming that some unspecified experience the authors have demonstrate the necessity of a completely new semantic standard is not scientific reporting or even convincing.

I'm afraid that Figure 1 makes little sense to me. Not all occurrents are finer in scale than all continuants and, as the authors themselves argue in their examples, not all qualities are finer in spatial scale than all subjects. Full logical elucidations and robustness under counter examples must be provided (again, another paper).

Regarding the criteria of compatibility, expressiveness, readability, and parsimony:

Compatibility: What sort of maintenance do these compilers require? Who will maintain them? How is accuracy evaluated? Are there use cases? I feel as though this needs more clear specification and I would dedicate more room to this than the attempts at upper level semantics in this manuscript. The authors may be considering publishing these in a follow up paper, but they really should be here as these are the elements that are of the highest value.

Expressiveness: Intuition is quite a subjective thing and actually the source of many semantic errors in the first place. I would consider it common knowledge that usage of similar terminology even by interacting expert groups often does not correspond to 'uncontroversial' usage unless they're using a standardised guide. So I can't support the idea that 'keywords' are sufficient to satisfy expressiveness, especially from a machine perspective. Further, most expressive (i.e. well-axiomatised) ontologies I'm aware of don't require users to know anything about the underlying axiomatisation in order to use their content, so this isn't really novel. Sections like this switch the nature of the manuscript into something more like a review, rather than focus on the outcomes of the project at hand, which is confusing. Further, are synonyms of classes handled/matched or do keywords (i.e. primary labels) have to be memorised by the user?

Readability: Naturally, very important, however do the examples provided really improve on expressions that are entered into ontology editors such as TopBraid or Protégé? Further, the readability outlined here requires some effort by the user to subscribe to and understand the ARIES upper level, which must be learned to be useful, as with any other system.

Parsimony: I assume the authors mean that the system should allow post-composition. This is already done with "dead simple design patterns" or TermGenies. These should be acknowledged so as not to give the impression that this well-recognised issue hasn't been at least partially addressed thusfar.

"In k.IM, particulars and universals are combined to specify observables;"

This makes very little sense to me - perhaps this just needs rephrasing, but I don't see how or why there would be a need to combine universals and particulars. One always observes the particular, which is linked through instantiation with the universal.

"k.IM statements readable and understandable by mimicking English syntax, while specifying much more complex, correct and consistent OWL2 axioms"

In principle, I see this as a useful step forward - but as the more friendly syntax is less specific and logically fuzzy (Table 1), what happens when new content is added and the syntax expands? Will this not just result in another layer of ambiguity through the creation of a strange k.IM dialect that silos users into its translation system? This connects to the issues with "expressiveness" outlined above.

I would like to see more elucidation of the definitions in Table 1 and also some developed justification as to why existing upper level treatments are not sufficient for the ARIES objectives. There may not be 1:1 matching in all cases, but I don't see great difficulty in pre-composing the classes noted in this table. Once again, reuse should be preferred to duplication unless there's a very clear reason not to go this route.

Subject: Is this not just the subject of observation? This then presupposes that an observation action has occurred, which suggests that this is not a high-level class, but something lower-level and operational.

configuration: This seems superfluous: isn't this true of any precomposed and well-axiomatised class?

Thing: how is this different from entity? Why should this be inanimate? As this is a common term (due to Protégé and other resources), would it not be more sensible to call this "inanimate object"? Again, if this has to be the output of an observation process, is this not too low-level to warrant the use of the word "thing"?

agent: This suggests all agents must be self-aware, correct?

Priority: This is certainly not the normal, intuitive, or common-sensical usage of this word, which runs counter to the authors' claims and reasoning above. Further, ranking is a numerical operation which can be arbitrarily applied across qualities, making this definition suspect. The same is true for quantity. There is a layer missing in these semantics.

Class: Again, a very dubious use of a term that is more or less reserved in the semantics community and not naturally or intuitively associated with qualities.

Process: I'm not sure why a process would only include a single subject. This doesn't seem very realistic. Do qualities need to change during participation in a process?

I could go on with each row of this table. The bottom line is that this manuscript does not provide sufficient arguments to show that this treatment is stable: this should be done in a separate publication reviewed by developers of upper level semantic resources.

Concerning the geographical elevation: it's very peculiar that the work done in other ontologies handling geology/geography and qualities is not used or their developers engaged to coordinate with this approach: again, it seems that this system is creating a silo, rather than linked, interoperable, and portable products. Given that this work is focused on FAIR thinking, I see this as a significant issue.

Statement (1) is quite readable, but I don't really feel that this is so much easier than composing statements in Protege, using a reasonable set of ontologies. What is much easier is that all of the components can be called by just identifying the namespace. I can imagine this as a very approachable working environment, calling in external semantics that are more rigorously developed. I would encourage the authors to make this more prominent relative to the claim that k.IM/k.Lab offers an intuitive experience: it seems that it's just as idiosyncratic as other solutions.

A graph/network figure illustrating the semantic model that corresponds to statement (1) - including a few

classes not identified in the statement, but linked to those that are identified - would help show the interconnectivity and architecture of the system. This would make statements like "This is done by tying the concept being defined to the core observation ontology" much more meaningful.

"The language contains keywords for many fundamental quantities, allowing users easy specification in most situations (Table 1)." Claims like this need some sort of support or explanation of why the authors think the coverage is so high. Which domain scientists have found this to be sufficient in most situations? If that data is not available, a description of what the authors have covered (as a supplement perhaps, as Table 1 cannot go into such depth) would be needed.

"along with constraints and relationships for all common scientific observables"

Similar issue to the above - I find it hard to believe or parse this statement. What is common here? How can the complete coverage be so confidently claimed?

Operators in general:

Is this not similar to/the same as what the Relations Ontology and other stores of Object Properties does? Are the properties defined for reasoning (more formally than the list in Table 2)? Is their reflexivity and inversion properties coded for reasoners and query systems to make use of? The language here suggests more novelty than there is and, once again, this doesn't seem reuse well-adopted object properties that already exist, making claims of FAIRness weaker.

Table 2: This reveals some confusing patterns - how does an operator produce a quality? I can understand presence being represented as a BFO:'dependent continuant' or PATO:quality, but I don't see how a relation (operator) can "produce" one.

- What is the distinction, metaphysically, of a quality and a quantity?
- Are "Countables" defined in a class? or are range restrictions placed on the operator?
- What is meant by "in a (spatial) context"? Is not everything that concerns this system in a spatial context?
- It seems strange to have occurrence as a shorthand for 'probability of'.

There are more points of imprecision/ambiguity that I could list here, but perhaps I'm not appreciating how these are treated internally. This should be more clear in the paper: how are these checked for logical consistency?

Statement (2): This seems like lax modelling: the CSV file is not the model, it's a CSV file which, perhaps, pertains to a model. I appreciate that the condensed syntax may expand internally to express things more precisely, but - if this is not the case - claims that this is semantically rigorous need to be softened.

"As concrete qualities (those of which observations can be made) can only exist inherently to a direct observable, the observable must be made explicit before the concepts can be used (e.g., earth:Region in the previous elevation example)"

This seems like semantics for no clear purpose. What is a non-concrete quality? Who defines the limits of observation? Naturally, a thing must exist (be instantiated) to be observed. Perhaps this is just poorly phrased?

"...so that users can easily locate concepts by textual searching."

This assumes that users all use the same syntax and/or terminology - which is hardly supportable without training. A useful feature, nonetheless. However, does this mean that users can specify any "concept description" they wish? How is this semantically controlled? If these definitions are done haphazardly, this could defeat the purpose of a semantic resource.

"is first established as its fundamental nature"

Claims like this seem un-falsifiable, and thus suspect. Does this mean the authors take these as primitives? Is there a list of all these primitives with their logical elucidations present? If not, it's hard to support this as a semantic resource.

"it will be correct to annotate elevation within a watershed, as long as a previous statement defines the Watershed concept as a type of earth:Region."

I'm not sure I follow. Does elevation inhere in the watershed or does it inhere in some entity that is, itself, somewhere within the watershed? This is not linguistically clear, and thus cannot be logically clear.

"The of keyword is used when the quality refers to a second, implicit observable in the context of inherency. For example, the "height of trees" quality in a region is inherent to that region, but implicitly describes tree subjects in it."

This is hard to support - this quality does not actually inhere in a region. Like the chlorophyll in the lake example, this can be understood as shorthand. While I appreciate the convenience, unless this is expanded in the background, this feels like a very risky route in semantic modelling with too many shortcuts.

"In keeping with our readability requirement, we only allow two levels of specification and use two different keywords (within and optionally of)."

This makes the language in k.IM very idiosyncratic. It's debatable whether this is useful for readability that results in reasonable semantics. This feels like setting one's own goalposts without objective criteria.

"We found that legitimate chained specifications, such as "x within y within z within ...", were awkward and difficult to understand in usage tests and decided against allowing such statements."

True, these are awkward, but why prevent them if they're logically valid? What's the alternative?

"Multiple chains of inherency of this kind can be defined using intermediate concepts"

This must be explained further.

"In knowledge domains (as opposed to physical ones), the implicit inherent subject is often a configuration."

This raises a few red flags for me - it's very easy to create semantic resources that have little bearing on physical reality and are thus of questionable use in the natural sciences. I don't see why topology or terrain (3) is treated as not "being directly amenable to providing the observable of an informational artifact". It is the bearer of a quality, which is claimed to be linkable to an information artifact. Following this logic, wouldn't events also be indescribable? It is useful to be able to pre-compose semantics for a group of phenomena (~ a configuration) that are often referenced together, but these are still physical.

"One can assume that identifying height of corn would require a further specialization of plant height, and the corn identity would simply be implied syntactically by using a Corn... prefix in the term assigned to the concept."

Are we talking about TO here? There is no PTO in OBO.

Regarding TO and other OBO ontologies, this wouldn't/shouldn't be limited to a syntactic specialisation, but include an axiom linking the class to a taxonomy. Many classes in TO follow this recommendation. Also, "giant embryo (a gene type)" is not correct. It's not a gene type, it's a quality. Also, these are not "concepts" as OBO resources take a realist stance on knowledge representation.

"plant height uniformity (a quality not physically commensurable with height)"

One would have to be more clear about this. I see the issue here as one of inherence (this quality inheres in a collection of whole plants, thus conflicting with the assertion in the superclass). Did the authors reach out to the TO developers to correct this?

"relative plant height (seemingly adding an observation-related attribute, relativity, out of many possible)." Again, this has to be more clear. I agree that this does not sit well in the TO as this is more an information artifact than a new type of quality and requires cross-axiomatisation with an ontology dealing with measurement processes. In OBO, this is not an uncommon occurrence and numerous "stubs" exist. These ontologies have issue trackers where ambiguities like these can be raised for later revision.

"Yet it is clear from this example that no ontology can force users to adopt cogent annotation practices, ensuring that physical and biological identities are preserved along inheritance chains and attributes retain traceable and stable meaning."

I'm not sure that the examples provided illustrate this point. They do show that even well-developed ontologies aren't perfect and need some classes to be refactored. However, annotation practices are an order removed from the development of a semantic backbone. It's of course possible to annotate in many ways, but whether these are going to work with the reference ontology and suit a user's objectives is a different question. Traceable and stable meaning is handled by the URI scheme and obsolescence best practices.

Further, I don't quite see how the k.IM example that follows (11) fixes these issues. The "language itself" is often rife with ambiguity and jargon and not a stable shortcut to develop rigorous semantic models (but can be good for a first pass). In the example, a "measure" is roughly the same as a BFO:quality, so they both require a material entity to inhere in. A quality and a unit do not embody the "how", one would need a process ontology like the "protocol" branch of OBI to do that. A particular quality/measure can only enforce a set of units of such constraints are hard-coded in the ontologies used for k.IM - is this the case? Would it not be sensible and more sustainable to create a working relationship with TO and other ontologies to fix the errors spotted and provide ARIES with better worldviews?

"Definitions (11) and (12) intentionally use agriculture:PlantIndividual instead of biology:Individual, as the latter requires a precise species identity (definition 7), while the former references the commonsense taxonomy used in AGROVOC for crop types, reflecting the intended semantics for the data."

How do the PlantIndividual and biology:Individual interoperate? This sounds like a point where semantic drift can occur quickly. In the OBO world, one would create union classes of biological species for less specific common terminologies.

"Most importantly, the adoption of rigorous phenomenological inheritance and specification syntax requires the realization (and the explicit statement) that height is first of all a quality of a plant subject, and that the data refer to plants within a cropfield subject."

I'm not sure where this requirement is coming from - the user? The task? The worldview? Further, expressing this (along with logical constraints on qualities etc) is quite possible using an application ontology derived from PATO, PO, PCO and ENVO rather than developing an ad hoc semantic.

"While such details still need to be learned by a user, the syntax itself serves as a guide for the annotation workflow: the use of inconsistent observables or the lack of proper inherency would yield ungrammatical statements that are reported as errors."

I don't really see how this is that better than the reasoning checks used by other ontologies. Further, I'm not sure if grammatical evaluation will always be a good indicator - that largely depends on the labels used, whether the syntax is sensible (I've already noted issues with some usage above), and if the grammar is appropriate for all cases in the scope of this work (not evaluated). This is actually quite a

major issue – whatever mechanisms are used to check logic should be absolutely true to that logic, not a grammar. Are there no reasoners used to check axiomatic consistency?

"non-abstract quality" The distinction between abstract and non-abstract qualities etc remains very confusing to me. A box or a supplement is needed to help explain the need with clear examples.

"The result is readable by a non-expert and compiles to axioms specifying a single OWL concept, which can be transferred to a remote endpoint in axiomatic form and reconstructed for reasoning or database querying; the shared worldview is the only requirement for its interpretation."

There are many "ifs" along this path: it may be readable by a non-expert, but so are asserted axioms in Protege - as the authors note above, the various conventions of both the k.IM system and the worldview chosen need to be familiar to the user for this to be not just read but understood. Yes, this can be transferred to an endpoint in an axiomatic form, but that assumes understanding of and stable interaction with the resources on the other end of that endpoint. I don't see evidence that this system does that as a rule: the authors should include examples of this or make this statement more clearly aspirational. The sentences following this one are claims of a similar kind that would need to be substantiated somehow.

"Specifications such as (13) are simple enough to be added to metadata data files with the same name - which may be automatically detected and indexed by specifically designed web crawlers, so that indexes of web-accessible, annotated datasets can be built and maintained."

This is very interesting - similar approaches (tiny semantic annotation files) have shown promise (see the PhenoPackets project). I think this MS would be greatly improved by focusing on such practical approaches rather than semantics which are present in other resources. Spending more time talking about how the environment can link elements of information artifacts to ontologies (e.g. as in (14)) would target an urgent gap in research and implementation. Unfortunately, most of this is not shown.

"Specification (13) is complete and correct, as geography:Elevation is fully characterized in terms of inherency within the worldview, as seen in statement (1)."

What is meant by "fully" here? I don't really see this as complete or correct. The definition of (1) is circular and calls for a model for description, which seems extraneous. Also, "terrain" is defined as pertaining to land surfaces only (3) implying that surfaces permanently covered by water don't have elevation (which I'm assuming means elevation above sea level), which is incorrect (e.g. many lakes such as Lake Tahoe and Khövsgöl Nuur).

"SMM, a modeling approach where FAIR+ interoperability is an integral requirement, sees data and models as definitions for possible observations: while datasets can produce, possibly through mediation of observation or context semantics, the requested observations in a self-contained way, models do so through computation that may involve the observation of other concepts they depend on, to be resolved through other data or models."

Again, defining "FAIR+ interoperability" (which seems internally redundant) as accomplished without any real demonstration of it in action can't be supported. Further, data and models are not definitions and treating them as such will likely lead to more ambiguity. The rest of this paragraph seems only half-developed and I'm not sure what the authors are trying to convey.

Reference 39 isn't really about k.LAB, while the prose suggests it is.

"The most important aspect in this sense is the clear focus on observable semantics: no time is wasted seeking semantics to express model-related concepts ("model", "variable"), observation-related ones ("measurement") or context-related ones ("spatial resolution"), all of which figure prominently in commonly used ontologies."

I take strong exception to this - time is by no means wasted by handling the semantics of informational and procedural entities. The fact that these feature prominently in other ontologies should be an indicator of this. In fact, one could argue that procedural metadata is needed to handling anything FAIRly, otherwise there can be no true reproducibility. The fact that the current approach neglects these is not really a virtue to be celebrated, although I do recognise that an observational focus is a good point of initialisation.

"Formalizing a simple phenomenology for observables and universals. The base observables in Figure 1 have proven intuitive enough to be understood and remembered by diverse users, helping them "home in" quickly on observable semantics as described in the previous point. Also, the use of independently defined and flexibly attributed universals to express attributes, identities and roles has effectively and intuitively solved, in our applications, the plaguing issue of excessive and improper specialization." This is purely speculative without some sort of comparative study. As stated above, I see major issues with the phenomenology or base semantics used here and no evidence that they are more accurate or useful than other systems (assuming users have equal training in both).

"home in" --> "hone in"

"The k.IM language and k.LAB platform make ontologies and annotations immediately actionable, enforcing the logical consistency of each definition both by enforcing syntactical correctness through intelligent editing tools and by employing a machine reasoner to identify and report logical errors to the user. The language guides, simplifies and validates the definition of knowledge; the support software provides feedback and allows users to immediately perform user queries and compute workflows whose results enable at-a-glance validation of the semantic correctness of the concepts employed." This is the most exciting part of this contribution, but its capacities are not shown or discussed at a meaningful depth. Which reasoners are used? Are they valid for any external resources brought in? What is meant by immediately actionable? How, exactly, does enforcement occur? Without more context, I can't see if the system actually simplifies and/or validates semantic work, as claimed.

Is the rationale for developing the new method (or application) clearly explained?

Partly

Is the description of the method technically sound?

Partly

Are sufficient details provided to allow replication of the method development and its use by others?

No

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?

No

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

Referee Expertise: ontology, semantics, bioinformatics, omics, microbial ecology

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Referee Report 14 August 2017

doi:10.5256/f1000research.12570.r22974



Carole A. Goble 

School of Computer Science, University of Manchester, Manchester, UK

The work described represents the first in a number of expected publications from a long running and important effort to build an integrated “e-Laboratory” for ecosystems modelling. The ambition of this k-LAB is to support the automatic assembly of scientific workflows over data and models, which in turn requires that inputs and outputs are compatible. The ARIES project web site gives some indication of the ambition and the driving application of work reported – more so than the paper.

This paper chiefly sets out to describe the semantic principles and methods to support the required interoperability (dubbed “FAIR+”) through systematic semantic descriptions of observable phenomena, the observations made on the observables and the context of these observations. The majority of the paper is given over to the philosophy underpinnings of these semantics and the presentation of a bespoke language and its support software developed by the authors to support the expression of the semantics.

Thus the main thrust of the paper is

- The fundamental arguments behind the development of an ontology for the systematic representation of observable phenomena (“specifying observable semantics”) suitable for ecoservices modelling across different domains and different contexts (notably scales)
- using this ontology to make observations about data and models (or “informational artifacts”) through annotations (“specifying observation semantics”)
- The presentation of the k.IM language of the authors’ invention to express the ontology and annotations with bespoke support software.
- that these annotations serve to support model input-output compatibility to support the automatic assembly of scientific workflows to answer questions posed using this (and other?) ontologies

Each of these points is interesting and worthwhile, most notably the development of a semantics observables that can transcend the difficulties of scales. The final point – demonstrating interoperability through compatible observations – is the least well demonstrated in the paper.

The semantics of observations is an important topic and has long challenged ontologists and those working in the fields of data annotation. The paper presents valuable insights into the challenges and thought processes in the development of the authors’ phenomenon-based semantic infrastructure and makes a case for why interoperability of observations is paramount for ecosystem modelling. The notion of defining a “worldview” is a useful one. The tiers of users chimes with experiences and expectations reported in the biosemantics literature. The authors are also well aware of the limitations of their approach as discussed in section 3.

The k.IM language is interesting, and it is more than plausible that one would want a language to disguise

OWL2: though claims to its ease of readability and its compatibility with OWL2 are not really demonstrated.

It's a very dense paper that packs a lot in and requires several readings. Part of the challenge of the paper is to describe the semantics coherently and completely enough to be convincing whilst leaving examples of how the semantics are used to other papers. It partially succeeds. The paper's presentation is also often frustrating.

Nonetheless this is stimulating and valuable work and a useful contribution to the semantics of observables and observations.

Suggestions for improvements:

Ontologies, semantics, whatever you like to call them, are for a purpose. The authors note the success of domain ontologies is because of their purpose. However, very little information is given on the purpose of the semantics of observations: the application, the nature of the questions, data and models, whether the data to be interoperated is public data or privately annotated or the nature of compatibility that is being sought. Section 2.6 and some hints in section 1.2 are the only hints we have towards the driver of the work – that is input-output compatibility. Many readers will not know what socio-ecological modelling is. A much better description, all in one place with an example from the beginning would be beneficial.

- It is not clear what is even meant by a model – the annotation example in section 2.5 suggests it is a tiff file.
- An example that will exercise compatibility across scales would be ideal. A clear description – with example - of the application of the enterprise would make it much easier to judge the value of the method and to judge the claims to interoperability made in the abstract and introduction. The salinity example in section 2.3 is the most compelling, as an example of how all the work of the semantics could be put to use and it's a pity this wasn't developed further in favour of a shorter treatment of the metaphysics of the semantics of observables.

There is a tendency to introduce key points and terminology almost as asides as one goes along which makes it much harder to digest than it should be. For example:

- Inherency for qualities is an important concept and is frequently referred to but is not specifically defined.
- Configuration is briefly introduced as an aside example – does the configuration Terrain render the need to state that Elevation is within Earth:Terrestrial earth: Region redundant as Elevation is im:Height of earth:Terrain?
- The whole of section 2.2 is a drip-feed of new keywords and terms that makes it a slog to work out from the examples what the semantics and syntax of the k.IM language actually are. The software is not amenable to the uninitiated – there are few readmes and the documentation seems to be entirely in a password-guarded wiki. Table 1 gives some words, but definitions of “is”, “as”, “described”, “requires”, “has children” etc are not given.
 - Tables 1 and 2 should have an example for each entry.
 - A complete definition of the full syntax of the language is needed at least in supplementary materials

The resultant ontology of observables is not available or referenced as far as I am aware. Granted that the semantics as defined by the language are expressed through the language: nevertheless it is not clear what the “common phenomenology” referenced in section 2.5 ended up as.

- What other ontologies are imported? We discover that quality is imported from the BFO ontology as an aside on page 13. In section 1.1 the authors do a good job of discussing various classes of

ontology and their roles, but then drop any further references. Tables 1 and 2 need to do a better job of the provenance of terms.

- If the paper is aimed at ontologists (as one might expect) then more detail is needed in the description of the final ontology. Several times the authors' mention that the descriptions compile to OWL 2.0. It would have been instructive to see one such compilation or to have an example as a supplementary material. Claims to compatibility with OWL2 are not really demonstrated.
- Is the ontology of observables available?

Figure 1 gives the intuition that the authors hope for but perhaps doesn't stand up to close scrutiny. Coarse grained spatial scale and temporal scale as continuant subject? Are all subjects coarse grained?

The section on identities is well written and clear – “bridging authorities” have been attempted for example <http://www.bridgedb.org/> and to some extent identifiers.org. *Scientific lenses to support multiple views over linked chemistry data*. Batchelor C, et al In The Semantic Web – ISWC 2014, Lecture Notes in Computer Science Volume 8796, 2014, 98-113 sets out the notion of linksets used in the Open PHACTS linked data platform for pharmacological data

Other related work has a few gaps.

- The notion of ontological patterns and languages, some of which sit on top of OWL, has been addressed in the literature. OPPL (the Ontology Pre-Processor Language) dates from 2010 (<http://oppl2.sourceforge.net/documentation.html>) and more recently tawny-owl (<https://github.com/phillord/tawny-owl>); both promote the notion of Ontology Patterns, which is inherent in the k.IM approach. Webulous takes a spreadsheet approach to ontology patterns rather than a language based one (<http://www.ebi.ac.uk/spot/webulous/>). Given that the k.IM language is effectively constraining Tier3 developers to certain patterns, this literature seems relevant. There is also the ontology annotation literature: examples include Zooma (<http://www.ebi.ac.uk/spot/zooma/>) and DOME0 (<https://doi.org/10.1186/2041-1480-3-S1-S1>) the latter of which uses the W3C Open Annotation Data Model.
- The notion of automated workflows using ontological annotations on inputs and outputs has also had some past attention. A well known example is the WINGS system: A Semantic Framework for Automatic Generation of Computational Workflows Using Distributed Data and Component Catalogs. Gil, et al *Journal of Experimental and Theoretical Artificial Intelligence*, 2011. <http://dx.doi.org/10.1080/0952813X.2010.490962>

Is the rationale for developing the new method (or application) clearly explained?

Yes

Is the description of the method technically sound?

Yes

Are sufficient details provided to allow replication of the method development and its use by others?

Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?

Partly

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Discuss this Article

Version 1

Author Response 15 Aug 2017

Ferdinando Villa, Basque Centre for Climate Change (BC3); IKERBASQUE, Basque Foundation for Science, Spain

Thank you for noticing and letting us know. We will certainly correct this in the first revision after all reviewers' comments are in.

Competing Interests: No competing interests were disclosed.

Reader Comment 10 Aug 2017

Alan Allwardt, USGS, USA

In section 1.1 (p. 4), the term "worldview" is introduced before it is defined in section 2 (p. 5). Given that "worldview" has different connotations in philosophy, linguistics, sociology, religion, and (in this instance) simulation modeling, I urge you define your use of the term the first time it is mentioned.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative

- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research