Evidence Graphs: Supporting Transparent and FAIR Computation, with Defeasible Reasoning on Data, Methods, and Results

Sadnan Al Manir[1] [0000-0003-4647-3877],
Justin Niestroy[1] [0000-0002-1103-3882],
Maxwell Adam Levinson[1] [0000-0003-0384-8499],
and Timothy Clark[1,2] [0000-0003-4060-7360]*

[1] University of Virginia, Department of Public Health Sciences, Charlottesville VA USA
[2] University of Virginia, School of Data Science, Charlottesville VA USA

*Corresponding author: twclark@virginia.edu

Abstract.
Introduction: Transparency of computation is a requirement for assessing the validity of computed results and research claims based upon them; and it is essential for access to, assessment, and reuse of computational components. These components may be subject to methodological or other challenges over time. While reference to archived software and/or data is increasingly common in publications, a single machine-interpretable, integrative representation of how results were derived, that supports defeasible reasoning, has been absent.
Methods: We developed the Evidence Graph Ontology, EVI, in OWL 2, with a set of inference rules, to provide deep representations of supporting and challenging evidence for computations, services, software, data, and results, across arbitrarily deep networks of computations, in connected or fully distinct processes. EVI integrates FAIR practices on data and software, with important concepts from provenance models, and argumentation theory. It extends PROV for additional expressiveness, with support for defeasible reasoning. EVI treats any computational result or component of evidence as a defeasible assertion, supported by a DAG of the computations, software, data, and agents that produced it.
Results: We have successfully deployed EVI for large-scale predictive analytics on clinical time-series data. Every result may reference its evidence graph as metadata, which can be extended when subsequent computations are executed.
Discussion: Evidence graphs support transparency and defeasible reasoning on results. They are first-class computational objects and reference the datasets and software from which they are derived. They support fully transparent computation, with challenge and support propagation. The EVI approach may be extended to include instruments, animal models, and critical experimental reagents.

Keywords: Provenance, Evidence Graphs, Computation, FAIR

1 Introduction

1.1 Motivation

It is now increasingly understood that dramatically enhanced capabilities for generating and analyzing very large datasets, with increasingly sophisticated methods, require systematic referencing of archived datasets and software as persistent first-class objects, with some machine-readable record of their provenance. There is now beginning to be an additional, and necessary, focus on software citation, and provenance. Principles and methods for achieving these goals have been defined and are in various stages of transition to practice in the scientific communications ecosystem [1–11].

Incorporating these practices has been advocated to improve verifiability, replicability, reproducibility, and reusability of computational results. The goal, which was established as a requirement at the dawn of modern science [12–15], is to make the process by which results and the claims they support are arrived at, transparent, and to allow the methods involved to be inspected – at least virtually – for adequacy, and if possible, reused, and improved upon in various applications.

We would like to do this for computations, using a clean, formal, and integrated approach, that does not require "boiling the ocean"; and in which artifacts such as provenance records, which may benefit a broad community of research, are generated principally as side-effects of normal computational work. We do not want to place unwanted burdens or requirements on researchers, with which they cannot realistically be expected to comply. Any system or method that generates such artifacts ought to have other attributes of significant value to researchers. In particular, we would like an ontology providing these features to provide useful functionality in a *digital commons* environment, where asynchronous reuse of results by various researchers occurs, and not necessarily as coherent workflows.

Of course, all published scientific results (including mathematical proofs) [16–19], are provisional. Results, methods, reasoning, computations, and interpretations may be challenged by others in the community, and frequently are [20, 21]. A major reason for methodological transparency is to support such reviews and challenges. Results in science are reviewed "by a jury of peers", similarly to adversarial proceedings in law: a case is made, which may be argued for and against, based on evidence. Published "findings" or claims, and indeed computational results, are not facts. They are *defeasible assertions* [22], which rely upon a chain of evidence as warrants for belief. That makes them part of a chain of argumentation. We, therefore, treat computations and their provenance as defeasible arguments for provisional results.

1.2 Related Work

We previously undertook an analysis of the scientific communications life cycle to develop the Micropublications Ontology (MP) [23], which introduced a focus on chains of evidence and defeasible reasoning, and an emphasis on the nature of scientific claims as embedded in arguments. This approach was inspired by argumentation theory and

by Greenberg's detailed model of citation distortion [24, 25], which highlights empirical issues with citation chains in the scientific literature. It was also motivated by a perceived tendency amongst some computer scientists to take claims in the biomedical literature as "facts", without subjecting them to further scrutiny.

Argumentation frameworks [26–34] and abstract dialectical frameworks [35, 36] are important sets of tools and concepts developed in the broader AI community, with an extensive literature. Bipolar Argumentation Frameworks (BAFs) as developed by Cayrol and others [31, 37, 38], allow both supporting and challenging arguments to be asserted and reasoned over in formal models. We discuss the relationship of our work to formal argumentation frameworks further in the Methods section.

The W3C PROV model [39–41] provides a well-thought-out and extensively tested set of core classes and properties, which may be used to document almost any computational provenance, in arbitrarily fine detail, and can serve as the basis for useful extensions. While PROV deals comprehensively and rigorously with core internal aspects of provenance, it does not engage explicitly with the role that provenance may play as evidence for computational results and interpretive claims. Nor does it conceptualize provenance as part of an argument for the validity of results, which may be countered, for example by later researchers finding bugs in code [42–45], flaws in datasets [46, 47], statistical errors [20], or fallacies in mathematical arguments [18].

Ontologies or schemas directly engaging the topic of experiments and experimental results include the Evidence and Claims Ontology (ECO) [48]; the Investigation, Study, Assay (ISA) model [49]; and the Ontology of Biomedical Investigations (OBI) [50]. All of these works are capable of representing machine-interpretable instances of scientific experiments within their proposed models and of offering limited provenance information. However, their focus is on characterizing individual experiments – principally in the wet lab – sometimes in exquisite detail, with over 4,000 classes in OBI, which make it best suited to use by highly trained specialist annotators. None of these models directly treat computational results, and none treat results as components of argumentation.

The ambitious Nanopublications model [51–54], was developed to standardize the form of research claims by recasting them as RDF triples, and aggregating holotypic claims with their paratypes, thus making the claims computable in some sense. What the nanopublications model currently lacks, is the ability to show evidential support in argumentation for the results it models.

Finally, there are related works that deal with computational results by packaging them up with their provenance records and other materials for citation or reference in metadata. These would include the Research Objects (RO) model [55–57] and its companion, RO-Crate [58]. The initial RO publication states as an explicit goal, replacing the "static PDF" form of scientific publication, and as such RO provides an integration framework across many related models involved in scientific specification, description, and communication. RO-Crate is a lightweight packaging initiative, or implementation realization, for RO and related material.

2 Methods

Cayrol and Lagasquie-Schiex's work on BAFs supplied inspiration for our approach; which has, however, somewhat different semantics from their model, regarding the support relation, and the meaning of an argument.

In BAFs, arguments are opaque, and without internal structure. This extremely abstract treatment derives from Dung's original presentation [33], enhanced to provide explicit bipolarity. In BAFs, if an argument A supports B, it agrees with B, similar to having an ally in a dispute. Therefore, an attack of C upon A also attacks B.

Our model, in contrast, treats support as it occurs in the scientific literature, as supporting evidence cited by the author of an argument. If C challenges A, a challenge by C on the supporting evidence for A cannot be inferred. However, as in Toulmin argumentation, and the work of Verheij [59–61], a challenge to A's supporting evidence undercuts A – it reduces the *warrant for belief* in A.

Evidence graphs in our approach may be represented abstractly as:

$$EG = \langle A, R+, R- \rangle, \text{ where:} \quad (1)$$
$$A \text{ is a set of representations;}$$
$$R+ \in A \times A \text{ is the set of } \textit{support} \text{ relations;}$$
$$R- \in A \times A \text{ is the set of } \textit{challenge} \text{ relations; and}$$
$$R+ \cap R- = \emptyset$$

In our treatment, Representations are generalized forms of the various types of assertions and evidence found in the scientific literature. All Representations have provenance, if only to the extent that they have authorship. A Representation may be a statement, or a set of assertions, claims, or declarations, with provenance as its support.

The simplest form of provenance is the attribution of an assertion to an *Agent*. A declarative sentence is an argument of this simple form, where a statement's attribution is its simplest supporting evidence. This corresponds to Aristotle's view of argument as having a significant measure of support from ήθος (ethos), the character of the speaker [62]; and relates to our common experience of tending to give more credence to statements from trusted, highly reputed sources. This notion can be extended to results produced by computational agents whose operations are well-validated and transparent.

A *Method* or *Material* representation in a publication, provided as provenance, constitutes a set of assertions about what was done to get a result. Software source code is a set of assertions of this form when used in a provenance description for a computation.

Source code has a dual nature, in that as a set of instructions, it is a collection of performatives [63], not evaluable as strictly true or false. We distinguish here between software as a set of instructions to a computer, which does something; and software provided as a description of what was done by the computer, i.e. "it ran this code".

We have adapted these notions to computational provenance representation. We base our model on the following key ideas:

- All data, methods descriptions, and results are sets of defeasible assertions.

- The evidence for the correctness of any result is the record of its provenance.
- Subsequent research or discussion may challenge results, datasets, or methods.

The EVI ontology is a formal representation of the evidence for any result or claim as an EvidenceGraph, which unifies existing models for both software and data citation and supports machine-based defeasible reasoning.

We first introduced several of the concepts used in EVI in our previous work on Micropublications. EVI simplifies, revises, and adapts many features of the micropublications model to a purely computational digital commons environment, where evidence graphs may be generated by a computation service. The computation service we developed, described elsewhere [64], provides affordances to the user by greatly simplifying access to very large-scale data and underlying parallel computation and workflow services. At the same time, it produces and extends evidence graphs transparently, as a side effect.

3 Results

EVI is an extension of W3C PROV, based on argumentation theory, which enables defeasible reasoning about computations, their results, data, and software. It can be extended to incorporate non-computational evidence important to the results of a computation, for example, the specific instrument (manufacturer and catalog number) used to obtain a dataset or the reagent used in a binding assay.

Evidence Graphs are directed acyclic graphs, DAGs, produced by a service when a computation is run. They are first-class digital objects and may have their own persistent identifiers and be referenced as part of the metadata of any result. They may be arbitrarily deep. We model these using an OWL 2 vocabulary and set of rules [65, 66] which provide for propagation of support and challenge relations, with direct supports/challenges distinguished from indirect. A diagram of the classes and relations is provided in **Fig. 1**.

The classes within the dotted-line box in the figure are basic PROV classes: *Entity*, *Agent*, and *Activity*. All relations in the box are subproperties of PROV relations.

The argumentation relations *supports* and *challenges* are introduced on the new class, *Representation*, a subclass of *prov:Entity*. A *Representation* is a *prov:Entity* that *represents* another *Entity*. All digital objects in EVI are considered to be *Representations*; which may contain other *Representations*.

Supports and *challenges* are superproperties of *directlySupports* and *directlyChallenges* (not shown in the Figure). *Supports* is transitive, *directlySupports* is not. *DirectlySupports* is a superproperty of the relation *usedBy* (inverse of *prov:used*), and of the property *generates* (inverse of generatedBy) so that if a recorded *Activit*y D used a particular *Representation* C as input, and *generates* E as output, then C *directlySupports* D and D *directlySupports* E.

This is shown in Example 1 below, illustrating how the property chain rule evaluates distant support in the graph, where <s> stands for the *supports* property, and <dS> stands for *directlySupports*.

$$(D \langle used \rangle\ C \wedge D \langle generates \rangle\ E) \Rightarrow \tag{2}$$

$$(C \langle dS \rangle\ D \wedge D \langle dS \rangle\ E) \Rightarrow$$

$$(C \langle s \rangle\ D \wedge D \langle s \rangle\ E) \Rightarrow C \langle s \rangle\ E$$

Supports and *directlySupports* properties are distinguished in this way because they serve different purposes in the model. The *directlySupports* property gives us our DAG,

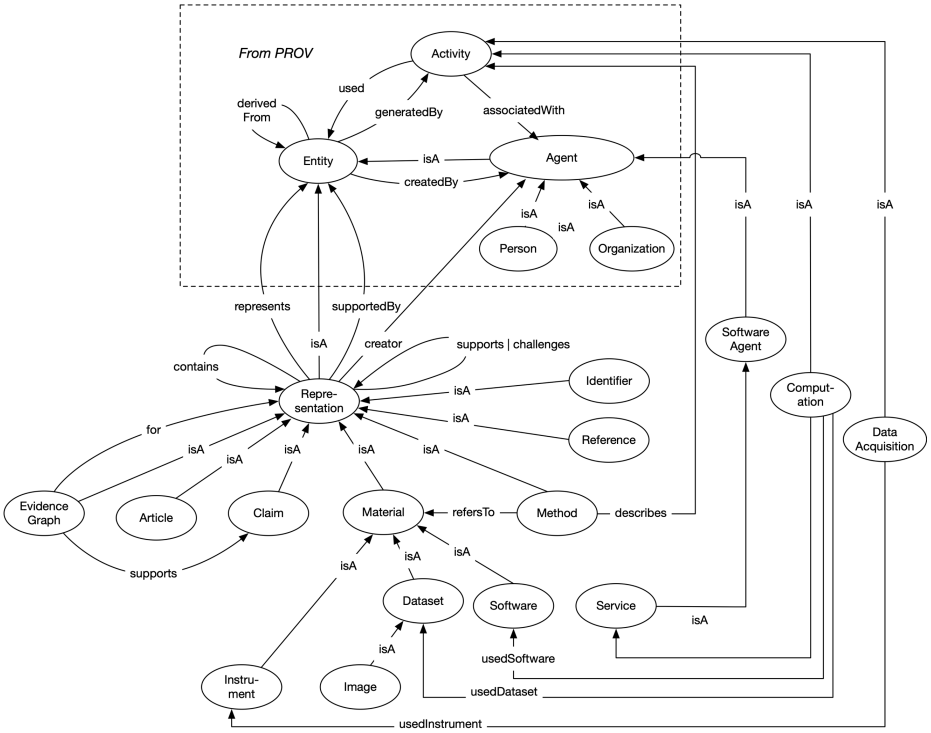

Fig. 1. Classes and relations in the Evidence Graph Ontology, EVI. https://w3id.org/EVI

connecting the various Representations to form an evidence graph. It is simply a generalization over several PROV properties. The *supports* property is transitive and allows us to infer distant support relations. By analogy with genealogical trees, *directlySupports* equates roughly to *hasParent* relations; *supports* equates roughly to *hasAncestor* relations.

Articles, *Claims*, *Material* (as used in a typical "Materials and Methods" section), *Methods* (likewise), and *EvidenceGraphs* themselves, are *Representations*. A *Method* describes an *Activity* that may *referTo* a *Material*. The class *Material* is iteratively subclassed into *Instrument, Dataset, Image, Software*, and *Downloads*. *Downloads* are particular distributions of a *Dataset* or *Software*, representing e.g. different formats of a *Dataset*, in the same manner, this approach is used in the schema.org class *DataDownload*. An *Agent* in EVI may be a *SoftwareAgent*, *Person*, or *Organization*. A *Service*

usedBy a *Computation* is a kind of *SoftwareAgent*. An *Activity* may be a *DataAcquisition* or a *Computation*.

Detailed formal definitions of all these terms, and equivalent schema.org [67] terms where they exist, are provided in the ontology description page at https://w3id.org, which has the latest updates. The OWL 2 vocabulary and its version history are on GitHub, here:

https://github.com/EvidenceGraph/EVI/blob/master/Ontology/versions/v0.2/evi.owl,

and the version we used for this article, is archived on Zenodo [68].

The diagram in **Fig. 1** omits detail relevant to rules and supports/challenges propagation, for pictorial clarity. The *supports* and *challenges* properties have subproperties *directlySupports*, *indirectlySupports*, *directlyChallenges*, and *indirectlyChallenges*, not shown in the figure. The *supports* property is propagated through the evidence graph via transitivity on *supports*.

$$(\text{A } \textit{directlySupports} \text{ B} \wedge \text{B } \textit{directlySupports} \text{ C}) \Rightarrow \text{A } \textit{supports} \text{ C} \tag{3}$$

Challenges are propagated through a property chain on *challenges* and *supports*.

$$(\text{A } \textit{supports} \text{ B} \wedge \text{X } \textit{challenges} \text{ A}) \Rightarrow \text{X } \textit{challenges} \text{ B} \tag{4}$$

4 Discussion

In developing and validating our approach, we performed a very large-scale time series analysis of ten years of vital signs time series data from the Neonatal Intensive Care Unit (NICU) at the University of Virginia Hospital, collected on 5,957 infants. This extended proof-of-concept was a significant research study in its own right, in which NICU time series were analyzed using 80 different algorithms in 11 mathematical families from diverse domains [69]. Computations were performed over a period of 15 months, in various stages, using the FAIRSCAPE microservices framework [64], producing an evidence graph of 17,996 nodes [70]. This evidence graph is queryable from a Stardog™ quad store via one of the microservices and is also deposited in the University of Virginia's Dataverse with a *prov:wasDerivedFrom* property associating it with the underlying dataset, which is also in Dataverse [71].

Whether using our FAIRSCAPE framework or some other approach, EVI statements are intended to be transparently generated by a *computation service*, similar to how ordinary PROV statements are generated by workflow engines. Thus, evidence graphs become side effects of doing useful computations. In our case, these computations are run within a FAIR digital commons environment.

Fig. 2 illustrates a section of this graph for one of the 5,957 subjects in the NICU analysis. In our microservices framework the EvidenceGraph service stores and retrieves the graphs in a Stardog™ RDF quad store, which also performs the inferencing. The PATH query, an extension to SPARQL, generates an *EvidenceGraph* DAG from the root object, using the *directlySupportedBy* abstraction (not shown here) to structure

the graph and the *supports* superproperty of its transitive inverse, *supportedBy*, to infer distant evidential support.

As can be seen in the example, every node in the graph has a persistent identifier, based on the ARK system [72] in our implementation. This approach provides the important unification of all available digital evidence for (and potentially, against) a result, by supporting persistent resolution to the cited objects.

One could ask, in a practical sense, if a challenge is made to some leaf node deep in the graph, should it always invalidate the root assertion? The answer to this is that challenges do not invalidate, they present an opposing view. They ultimately require human judgment as to their validity and strength. We do, however, wish to know about them. This becomes important in digital commons environments, and in, for example, meta-analyses reusing prior results. It can have a further impact if computational results are properly referenced and tied to textual claims in citation networks, such as those explored by Greenberg [24] and others.

We believe that EVI provides an important generalization of provenance as evidence for correctness, which can be further extended beyond computation to include the other


```
{'@context': {'@vocab': 'http://schema.org/', 'evi': 'http://w3id.org/EVI#'},
 '@id': 'ark:59853/4d38f7c2-4f19-409e-a864-fa3734fe7f2a',
 '@type': 'Dataset',
 'evi:generatedBy': {'@id': 'ark:59853/f581d2cc-8709-4110-8f1c-eac767065e72',
  '@type': 'evi:Computation',
  'evi:usedDataset': {'@id': 'ark:59853/7fe5f0a8-a65c-458d-a8dc-448a72a36da7',
   '@type': 'Dataset',
   'evi:generatedBy': {'@id': 'ark:59853/de0e548c-2471-47da-8732-eac617c63b54',
    '@type': 'evi:Computation',
    'evi:usedDataset': {'@id': 'ark:59853/b845a90e-b126-4676-bc17-3db5b2677fb1',
     '@type': 'Dataset',
     'author': {'@id': 'https://orcid.org/0000-0002-1103-3882',
      '@type': 'Person',
      'name': 'Justin Niestroy'},
     'name': 'Raw Time Series Feature Set'},
    'evi:usedSoftware': {'@id': 'ark:59853/4fedddb5-7674-4ece-8a64-c595c50f5685',
     '@type': 'SoftwareSourceCode',
     'author': {'@id': 'https://orcid.org/0000-0002-1103-3882',
      '@type': 'Person',
      'name': 'Justin Niestroy'},
     'name': 'Raw Data Analysis Script'},
    'name': 'Computation'},
   'name': 'part-00000-ed4b79c0-43f8-437d-9959-a98c9d63be39-c000.csv'},
  'evi:usedSoftware': {'@id': 'ark:59853/789b5f8e-2277-4e64-847e-86b201b4660a',
   '@type': 'SoftwareSourceCode',
   'author': {'@id': 'https://orcid.org/0000-0002-1103-3882',
    '@type': 'Person',
    'name': 'Justin Niestroy'},
   'name': 'HCTSA Cluster Script'},
  'name': 'Computation'},
 'name': 'HCTSA Heatmap.png'}
```


Fig. 2. JSON-LD of a portion of the evidence graph for one of the 5,997 infant subjects from clustering step in comparative time-series analysis, adapted from [64].

types of evidence presented in scientific publications, for example by including identifiers such as RRIDs [73] of important experimental reagents and animal models in the

protocols from which the data was derived. Our end goal is to be able to provide, with the metadata of any datasets we store in an archive, a link to the corresponding evidence graph. With this practice, any result presented in a publication, with appropriate data citation, ought to be resolvable to its entire evidence graph, and transitively closed to its components. Future research directions include extended support for packaging evidence graphs; support for extended descriptive metadata; service integration with data and software archives; and continued alignment with other initiatives in this space. As a component of the FAIRSCAPE digital commons framework, we plan for EVI to continue its evolution with direct input from computational users.

Information sharing statement

- The EVI ontology OWL2 vocabulary is available at https://w3id.org/EVI# under MIT license.

Acknowledgements

We thank Chris Baker (University of New Brunswick), Carole Goble (University of Manchester), and John Kunze (California Digital Library) for helpful discussions. This work was supported in part by the U.S. National Institutes of Health, grant NIH 1U01HG009452; and by a grant from the Coulter Foundation.

References

1. Cousijn, H., Kenall, A., Ganley, E., Harrison, M., Kernohan, D., Lemberger, T., Murphy, F., Polischuk, P., Taylor, S., Martone, M., Clark, T.: A data citation roadmap for scientific publishers. Sci Data. 5, 180259 (2018). https://doi.org/10.1038/sdata.2018.259.
2. Data Citation Synthesis Group: Joint Declaration of Data Citation Principles. Future of Research Communication and e-Scholarship (FORCE11), San Diego CA (2014).
3. Fenner, M., Crosas, M., Grethe, J.S., Kennedy, D., Hermjakob, H., Rocca-Serra, P., Durand, G., Berjon, R., Karcher, S., Martone, M., Clark, T.: A data citation roadmap for scholarly data repositories. Scientific Data. 6, 28 (2019). https://doi.org/10.1038/s41597-019-0031-8.
4. Groth, P., Cousijn, H., Clark, T., Goble, C.: FAIR Data Reuse – the Path through Data Citation. Data Intelligence. 2, 78–86 (2020). https://doi.org/10.1162/dint_a_00030.
5. Juty, N., Wimalaratne, S.M., Soiland-Reyes, S., Kunze, J., Goble, C.A., Clark, T.: Unique, Persistent, Resolvable: Identifiers as the foundation of FAIR. Data Intelligence. 2, 30–39 (2020). https://doi.org/10.5281/zenodo.3267434.
6. Katz, D., Chue Hong, N., Clark, T., Muench, A., Stall, S., Bouquin, D., Cannon, M., Edmunds, S., Faez, T., Feeney, P., Fenner, M., Friedman, M., Grenier, G.,

Harrison, M., Heber, J., Leary, A., MacCallum, C., Murray, H., Pastrana, E., Perry, K., Schuster, D., Stockhause, M., Yeston, J.: Recognizing the value of software: a software citation guide [version 2; peer review: 2 approved]. F1000Research. 9, (2021). https://doi.org/10.12688/f1000research.26932.2.

7. Katz, D.S., Gruenpeter, M., Honeyman, T.: Taking a fresh look at FAIR for research software. Patterns. 2, (2021). https://doi.org/10.1016/j.patter.2021.100222.
8. Smith, A.M., Katz, D.S., Niemeyer, K.E., FORCE11 Software Citation Working Group: Software citation principles. PeerJ Computer Science. 2, e86 (2016). https://doi.org/10.7717/peerj-cs.86.
9. Starr, J., Castro, E., Crosas, M., Dumontier, M., Downs, R.R., Duerr, R., Haak, L.L., Haendel, M., Herman, I., Hodson, S., Hourclé, J., Kratz, J.E., Lin, J., Nielsen, L.H., Nurnberger, A., Proell, S., Rauber, A., Sacchi, S., Smith, A., Taylor, M., Clark, T.: Achieving human and machine accessibility of cited data in scholarly publications. PeerJ Computer Science. 1, e1 (2015). https://doi.org/10.7717/peerj-cs.1.
10. Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data. 3, 160018 (2016).
11. Wimalaratne, S.M., Juty, N., Kunze, J., Janée, G., McMurry, J.A., Beard, N., Jimenez, R., Grethe, J.S., Hermjakob, H., Martone, M.E., Clark, T.: Uniform resolution of compact identifiers for biomedical data. Scientific Data. 5, 180029 (2018). https://doi.org/10.1038/sdata.2018.29.
12. Dear, P.: Revolutionizing the Sciences: European Knowledge and Its Ambitions, 1500-1700. Princeton University Press, Princeton and Oxford (2009).
13. Holmes, F.L.: Argument and Narrative in Scientific Writing. In: Dear, P. (ed.) The Literary Structure of Scientific Argument: Historical Studies. p. 224. University of Pennsylvania Press, Philadelphia (1991).
14. Rossi, P.: Philosophy, Technology, and the Arts in the Early Modern Era. Harper & Row, New York (1970).
15. Shapin, S.: Pump and Circumstance: Robert Boyle's Literary Technology. In: Hellyer, M. (ed.) The Scientific Revolution. Blackwell, Oxford (2003).
16. Committee on Science, Engineering, and Public Policy of the National Academies: On Being a Scientist: Responsible Conduct in Research. National Academies Press, Washington DC (1995).
17. Lakatos, I.: Proofs and Refutations. Cambridge University Press (1976).
18. Maxwell, E.A.: Fallacies in Mathematics. Cambridge University Press, Cambridge, UK (1959).

19. Krabbe, E.C.W.: Strategic Maneuvering in Mathematical Proofs. Argumentation. 22, 453–468 (2008). https://doi.org/10.1007/s10503-008-9098-7.
20. Ioannidis, J.P.A.: Why Most Published Research Findings Are False. PLoS Med. 2, e124 (2005). https://doi.org/10.1371/journal.pmed.0020124.
21. Ioannidis, J.A.: Contradicted and initially stronger effects in highly cited clinical research. JAMA. 294, 218–228 (2005). https://doi.org/10.1001/jama.294.2.218.
22. Koons, R.: Defeasible Reasoning, http://plato.stanford.edu/archives/spr2014/entries/reasoning-defeasible/, (2013).
23. Clark, T., Ciccarese, P.N., Goble, C.A.: Micropublications: a semantic model for claims, evidence, arguments and annotations in biomedical communications. J Biomed Semantics. 5, 28 (2014). https://doi.org/10.1186/2041-1480-5-28.
24. Greenberg, S.A.: Understanding belief using citation networks. Journal of Evaluation in Clinical Practice. 17, 389–393 (2011). https://doi.org/10.1111/j.1365-2753.2011.01646.x.
25. Greenberg, S.A.: How citation distortions create unfounded authority: analysis of a citation network. Bmj. 339, b2680 (2009). https://doi.org/10.1136/bmj.b2680.
26. Bench-Capon, T.J.M., Dunne, P.E.: Argumentation in artificial intelligence. Artificial Intelligence. 171, 619–641 (2007). https://doi.org/10.1016/j.artint.2007.05.001.
27. Besnard, P., Hunter, A.: Elements of Argumentation. MIT Press, Cambridge MA, USA (2008).
28. Boella, G., D.M.Gabbay, Van Der Torre, L., Villata, S.: Support in Abstract Argumentation. In: Baroni, P., Cerutti, F., Giacomin, M., Simari, G.R., Guarino, N., Kok, J.N., Liu, J., Lopez de Mántaras, R., Mizoguchi, R., Musen, M., S.K.Pal, and Zhong, N. (eds.) Computational Models of Argument. IOS Press, Amsterdam (2010).
29. Brewka, G., Polberg, S., Woltran, S.: Generalizations of Dung Frameworks and Their Role in Formal Argumentation. IEEE Intell. Syst. 29, 30–38 (2014). https://doi.org/10.1109/MIS.2013.122.
30. Carrera, Á., Iglesias, C.A.: A systematic review of argumentation techniques for multi-agent systems research. Artif Intell Rev. 44, 509–535 (2015). https://doi.org/10.1007/s10462-015-9435-9.
31. Cayrol, C., Lagasquie-Schiex, M.-C.: Bipolar Abstract Argumentation Systems. In: Rahwan, I. and Simari, G.R. (eds.) Argumentation in Artificial Intelligence. Springer, Dordrecht (2009).
32. Cohen, A., Gottifredi, S., García, A.J., Simari, G.R.: An approach to abstract argumentation with recursive attack and support. Journal of Applied Logic. 13, 509–533 (2015). https://doi.org/10.1016/j.jal.2014.12.001.
33. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence. 77, 321–357 (1995). https://doi.org/10.1016/0004-3702(94)00041-x.
34. Oren, N., Norman, T.J.: Semantics for Evidence-Based Argumentation. 9 (2003).
35. Brewka, G., Woltran, S.: Abstract Dialectical Frameworks. 10 (2010).

36. Dung, P.M., Thang, P.M.: Representing the semantics of abstract dialectical frameworks based on arguments and attacks. Argument & Computation. 9, 249–267 (2018). https://doi.org/10.3233/AAC-180427.
37. Cayrol, C., Lagasquie-Schiex, M.-C.: Coalitions of arguments: A tool for handling bipolar argumentation frameworks. International Journal of Intelligent Systems. 25, 83–109 (2010). https://doi.org/10.1002/int.20389.
38. Cayrol, C., Lagasquie-Schiex, M.-C.: Bipolarity in argumentation graphs: Towards a better understanding. International Journal of Approximate Reasoning. 54, 876–899 (2013). https://doi.org/10.1016/j.ijar.2013.03.001.
39. Gil, Y., Miles, S., Belhajjame, K., Deus, H., Garijo, D., Klyne, G., Missier, P., Soiland-Reyes, S., Zednik, S.: PROV Model Primer: W3C Working Group Note 30 April 2013, https://www.w3.org/TR/prov-primer/, (2013).
40. Lebo, T., Sahoo, S., McGuinness, D., Belhajjame, K., Cheney, J., Corsar, D., Garijo, D., Soiland-Reyes, S., Zednik, S., Zhao, J.: PROV-O: The PROV Ontology W3C Recommendation 30 April 2013. (2013).
41. Moreau, L., Missier, P., Belhajjame, K., B'Far, R., Cheney, J., Coppens, S., Cresswell, S., Gil, Y., Groth, P., Klyne, G., Lebo, T., McCusker, J., Miles, S., Myers, J., Sahoo, S., Curt Tilmes: PROV-DM: The PROV Data Model: W3C Recommendation 30 April 2013. World Wide Web Consortium (2013).
42. Soergel, D.A.W.: Rampant software errors may undermine scientific results. F1000 Research. 3, 303 (2015). https://doi.org/10.12688/f1000research.5930.2.
43. Bhandari Neupane, J., Neupane, R.P., Luo, Y., Yoshida, W.Y., Sun, R., Williams, P.G.: Characterization of Leptazolines A–D, Polar Oxazolines from the Cyanobacterium *Leptolyngbya* sp., Reveals a Glitch with the "Willoughby–Hoye" Scripts for Calculating NMR Chemical Shifts. Org. Lett. 21, 8449–8453 (2019). https://doi.org/10.1021/acs.orglett.9b03216.
44. Miller, G.: A Scientist's Nightmare: Software Problem Leads to Five Retractions. Science. 314, 1856–1857 (2006). https://doi.org/10.1126/science.314.5807.1856.
45. Axelrod, V.: Minimizing bugs in cognitive neuroscience programming. Front. Psychol. 5, (2014). https://doi.org/10.3389/fpsyg.2014.01435.
46. Brown, A.W., Kaiser, K.A., Allison, D.B.: Issues with data and analyses: Errors, underlying themes, and potential solutions. Proc Natl Acad Sci USA. 115, 2563–2570 (2018). https://doi.org/10.1073/pnas.1708279115.
47. Goldberg, S.I., Niemierko, A., Turchin, A.: Analysis of Data Errors in Clinical Research Databases. 5.
48. Giglio, M., Tauber, R., Nadendla, S., Munro, J., Olley, D., Ball, S., Mitraka, E., Schriml, L.M., Gaudet, P., Hobbs, E.T., Erill, I., Siegele, D.A., Hu, J.C., Mungall, C., Chibucos, M.C.: ECO, the Evidence & Conclusion Ontology: community standard for evidence information. Nucleic Acids Research. 47, D1186–D1194 (2019). https://doi.org/10.1093/nar/gky1036.
49. Rocca-Serra, P., Brandizi, M., Maguire, E., Sklyar, N., Taylor, C., Begley, K., Field, D., Harris, S., Hide, W., Hofmann, O., Neumann, S., Sterk, P., Tong, W., Sansone, S.-A.: ISA software suite: supporting standards-compliant experimental annotation and enabling curation at the community level. Bioinformatics. 26, 2354–2356 (2010). https://doi.org/10.1093/bioinformatics/btq415.

50. Bandrowski, A., Brinkman, R., Brochhausen, M., Brush, M.H., Bug, B., Chibucos, M.C., Clancy, K., Courtot, M., Derom, D., Dumontier, M., Fan, L., Fostel, J., Fragoso, G., Gibson, F., Gonzalez-Beltran, A., Haendel, M.A., He, Y., Heiskanen, M., Hernandez-Boussard, T., Jensen, M., Lin, Y., Lister, A.L., Lord, P., Malone, J., Manduchi, E., McGee, M., Morrison, N., Overton, J.A., Parkinson, H., Peters, B., Rocca-Serra, P., Ruttenberg, A., Sansone, S.-A., Scheuermann, R.H., Schober, D., Smith, B., Soldatova, L.N., Stoeckert, C.J., Taylor, C.F., Torniai, C., Turner, J.A., Vita, R., Whetzel, P.L., Zheng, J.: The Ontology for Biomedical Investigations. PLoS ONE. 11, e0154556 (2016). https://doi.org/10.1371/journal.pone.0154556.
51. Velterop, J.: Nanopublications: the future of coping with information overload. LOGOS: The Journal of the World Book Community. 21, 119–122 (2010). https://doi.org/10.1163/095796511X560006.
52. Gibson, A., van Dam, J., Schultes, E., Roos, M., Mons, B.: Towards Computational Evaluation of Evidence for Scientific Assertions with Nanopublications and Cardinal Assertions. In Proceedings of the 5th International Workshop on Semantic Web Applications and Tools for Life Sciences (SWAT4LS) Paris, France. 28–30 (2012).
53. Groth, P., Gibson, A., Velterop, J.: The Anatomy of a Nano-publication. Information Services and Use. 30, 51–56 (2010). https://doi.org/10.3233/ISU-2010-0613.
54. Schultes, E., Chichester, C., Burger, K., Kotoulas, S., Loizou, A., Tkachenko, V., Waagmeester, A., Askjaer, S., Pettifer, S., Harland, L., Haupt, C., Batchelor, C., Vazquez, M., Fernández, J.M., Saito, J., Gibson, A., Wich, L.: The Open PHACTS Nanopublication Guidelines V1.8. EU Innovative Medicines Initiative - Open PHACTS Project RDF/Nanopublication Working Group (2012).
55. DeRoure, D., Goble, C.: Lessons from myExperiment: Research Objects for Data Intensive Research. Presented at the eScience Workshop 2009 (2009).
56. Bechhofer, S., Roure, D.D., Gamble, M., Goble, C., Buchan, I.: Research Objects: Towards Exchange and Reuse of Digital Knowledge. Presented at the The Future of the Web for Collaborative Science (FWCS), 19th International World Wide Web Conference (WWW'10) April 26 (2010).
57. Belhajjame, K., Zhao, J., Garijo, D., Gamble, M., Hettne, K., Palma, R., Mina, E., Corcho, O., Gómez-Pérez, J.M., Bechhofer, S., Klyne, G., Goble, C.: Using a suite of ontologies for preserving workflow-centric research objects. Journal of Web Semantics. 32, 16–42 (2015). https://doi.org/10.1016/j.websem.2015.01.003.
58. Carragáin, E.Ó., Goble, C., Sefton, P., Soiland-Reyes, S.: A lightweight approach to research object data packaging. (2019). https://doi.org/10.5281/ZENODO.3250687.
59. Toulmin, S.E.: The Uses of Argument. Cambridge University Press, Cambridge UK (2003).
60. Verheij, B.: Evaluating Arguments Based on Toulmin's Scheme. Argumentation. 19, 347–371 (2005). https://doi.org/10.1007/s10503-005-4421-z.
61. Verheij, B.: The Toulmin Argument Model in Artificial Intelligence. Or: how semi-formal, defeasible argumentation schemes creep into logic. In: Rahwan, I.

and Simari, G. (eds.) Argumentation in Artificial Intellgence. Springer, Dordrecht (2009).
62. Aristotle: Rhetoric. Dover Publications, Mineola NY (2004).
63. Austin, J.L.: How to Do Things With Words. Harvard University Press, Cambridge MA (1962).
64. Levinson, M.A., Niestroy, J., Al Manir, S., Fairchild, K.D., Lake, D.E., Moorman, J.R., Clark, T.: FAIRSCAPE: A Framework for FAIR and Reproducible Biomedical Analytics. 2020.08.10.244947 (2020). https://doi.org/10.1101/2020.08.10.244947.
65. OWL 2 Working Group: OWL 2 Web Ontology Language: W3C Recommendation 27 October 2009. World Wide Web Consortium, Cambridge MA, US (2009).
66. Al Manir, S., Niestroy, J., Levinson, M., Clark, T.: EVI: The Evidence Graph Ontology, OWL 2 Vocabulary. Zenodo (2021).
67. Guha, R.V., Brickley, D., Macbeth, S.: Schema.org: evolution of structured data on the web. Communications of the ACM. 59, 44–51 (2016). https://doi.org/10.1145/2844544.
68. Troupin, C., Muñoz, C., Fernández, J.G.: Scientific results traceability: software citation using GitHub & Zenodo. 4 (2018).
69. Niestroy, J., Levinson, M.A., Manir, S.A., Clark, T., Moorman, J.R., Fairchild, K.D., Lake, D.E.: Discovery of signatures of fatal neonatal illness in vital signs using highly comparative time-series analysis. BioRXiv. 2021.03.26.437138, (2021). https://doi.org/10.1101/2021.03.26.437138.
70. Niestroy, J., Levinson, M.A., Al Manir, S., Clark, Timothy.: Evidence Graph for: Discovery of signatures of fatal neonatal illness in vital signs using highly comparative time-series analysis, https://doi.org/10.18130/V3/HHTAYI, (2021).
71. Niestroy, J., Levinson, M.A., Al Manir, S., Clark, T., Moorman, J.R., Fairchild, K.D., Lake, D.E.: Replication Data for: Discovery of signatures of fatal neonatal illness in vital signs using highly comparative time-series analysis, V2, https://doi.org/10.18130/V3/VJXODP, (2021).
72. Kunze, J., Rodgers, R.: The ARK Identifier Scheme, https://escholarship.org/uc/item/9p9863nc, (2008).
73. Bandrowski, A.E., Martone, M.E.: RRIDs: A Simple Step toward Improving Reproducibility through Rigor and Transparency of Experimental Methods. Neuron. 90, 434–436 (2016). https://doi.org/10.1016/j.neuron.2016.04.030.