Applied Computing and Geosciences 20 (2023) 100142

Contents lists available at ScienceDirect

Applied Computing and Geosciences

journal homepage: www.sciencedirect.com/journal/applied-computing-and-geosciences

The cultural-social nucleus of an open community: A multi-level community knowledge graph and NASA application

Ryan M. McGranaghan [a,b,*], Ellie Young [c], Cameron Powers [d], Swapnali Yadav [d], Edlira Vakaj [d]

[a] NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[b] NASA Goddard Space Flight Center, Greenbelt, MD, USA
[c] Common Action, USA
[d] Birmingham City University, B47 XG, United Kingdom

ARTICLE INFO

Keywords:
Knowledge representation
Knowledge graph
Collaboration
Community
Data science
Open science
Inclusivity
Accessibility
Information organization
Heliophysics
Space physics
Collective intelligence

ABSTRACT

The challenges faced by science, engineering, and society are increasingly complex, requiring broad, cross-disciplinary teams to contribute to collective knowledge, cooperation, and sensemaking efforts. However, existing approaches to collaboration and knowledge sharing are largely manual, inadequate to meet the needs of teams that are not closely connected through personal ties or which lack the time to respond to dynamic requests for contextual information sharing. Nonetheless, in the current remote-first, complexity-driven, time-constrained workplace, such teams are both more common and more necessary. For example, the NASA Center for HelioAnalytics (CfHA) is a growing and cross-disciplinary community that is dedicated to aiding the application of emerging data science techniques and technologies, including AI/ML, to increase the speed, rigor, and depth of space physics scientific discovery. The members of that community possess innumerable skills and competencies and are involved in hundreds of projects, including proposals, committees, papers, presentations, conferences, groups, and missions. Traditional structures for information and knowledge representation do not permit the community to search and discover activities that are ongoing across the Center, nor to understand where skills and knowledge exist. The approaches that do exist are burdensome and result in inefficient use of resources, reinvention of solutions, and missed important connections. The challenge faced by the CfHA is a common one across modern groups and one that must be solved if we are to respond to the grand challenges that face our society, such as complex scientific phenomena, global pandemics and climate change. We present a solution to the problem: a community knowledge graph (KG) that aids an organization to better understand the resources (people, capabilities, affiliations, assets, content, data, models) available across its membership base, and thus supports a more cohesive community and more capable teams, enables robust and responsible application of new technologies, and provides the foundation for all members of the community to co-evolve the shared information space. We call this the Community Action and Understanding via Semantic Enrichment (CAUSE) ontology. We demonstrate the efficacy of KGs that can be instantiated from the ontology together with data from a given community (shown here for the CfHA). Finally, we discuss the implications, including the importance of the community KG for open science.

1. Introduction

The challenges faced by science, engineering, and society are increasingly complex, requiring new levels of collective knowledge, cooperation, and sensemaking (Council, 2014). Groups hoping to respond to these challenges must be broader and more cross-disciplinary and must coordinate cohesively (Council, 2015). Creating more cohesive and informed teams rests on the ability of scientists and decision makers to successfully navigate two simultaneous challenges: (1) traversing voluminous technical information to understand evolving, multifaceted domains of study and (2) creating connections across broad, distributed, high-diversity teams with shifting networks of internal collaborations and external partnerships.

A variety of information technology tools have evolved to address the first of these challenges—access to technical information. Knowledge graphs are increasingly gaining traction as a promising infrastructure

* Corresponding author at: NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA.
E-mail address: ryan.mcgranaghan@gmail.com (R.M. McGranaghan).

https://doi.org/10.1016/j.acags.2023.100142
Received 29 March 2023; Received in revised form 28 September 2023; Accepted 1 November 2023
Available online 3 November 2023

to present scientific assets in an interoperable manner, and form the backbone of major scientific information networks, such as the National Science Foundation's EarthCube[1] and Open Knowledge Network[2] programs. Knowledge graphs are also widely used in Europe for academic and government applications (Chatterjee et al., 2021), and increasingly adopted by global companies, as described by Gartner.[3]

While these initiatives show promise for organizing dynamic data, the utility of knowledge graphs remains primarily limited to the backend of applications. For users, information discovery still requires manual search through vast information corpuses and monitoring of new publications across multiple sources. With content published voluminously across disparate disciplines and distributed through centralized, large-scale publishers, or a myriad of small, highly diverse outlets, merely identifying relevant resources represents a significant time investment for scientists.

Furthermore, the volume, breadth, and complexity of scientific coordination is no longer limited to the discovery and organization of scientific materials; the complexity in collaboration between scientists has also increased, even as scientists' coordination capacity, and supportive tooling, has remained fundamentally static. Consider the informal knowledge coordination activities of an individual scientist. To create knowledge that is tailored to unique, evolving circumstances, researchers must identify fruitful technical partnerships, including with colleagues studying related topics but with whom they may not yet have personally connected. In addition to following new work of personal interest, they had like to stay abreast of the work that close colleagues are pursuing, as well as that of their new partnerships. They must follow developments in funding and strategic landscapes. They can also benefit from knowledge about their organization's existing resource base, including ongoing projects, or available expertise. And they must be open to diverse perspectives, avoid bias, and incorporate emerging scientists; by corollary, emerging scientists must also develop the skills to pursue funding, partnerships, and publication and speaking opportunities. All of the information required to achieve these various priorities flows through informal communication lines, consisting of interactions at conferences, via email, or through personal investigation (e.g., of funding opportunities), at great time cost to each individual and low visibility to the scientist's broader network.

By providing a machine-readable structure through which information of this nature can be connected to common metadata terms, knowledge graphs (KG) are capable of alleviating these information and collaboration challenges. However, to date knowledge graph projects have largely been targeted to formal knowledge management goals, with specific focus on technical data collection, graph construction, and maintenance (a natural focus given the relative ease with which these data can be encoded compared to informal communications). This manuscript provides the foundation for transferring those successes to more informal settings, detailing an ontological design methodology for a community knowledge graph, developed by the authors in a pilot conducted with NASA's Center for HelioAnalytics.

As we will illustrate below, this use case is naturally deliverable through the flexible, interlinked structure of a knowledge graph. While many definitions of a knowledge graph (KG)[4] have been offered, for the purposes of this paper we adopt the following description: a KG is a logically consistent definition of the concepts and their relationships of a given domain of knowledge that is both human and machine readable (Bonatti et al., 2018). The graph representation implies that the information is organized into entities (nodes) and their relationships (edges) (e.g., Hogan et al., 2021).

We further specify a community knowledge graph as one representing information about the members of a community-based organization (including their partnerships and collaborations outside of the organization), allowing individuals to both share and access information about and across the communities in which they are situated. While this variation may sound simply like a conceptual specification, the schema is not the only essential feature of a community knowledge graph. Additionally, community knowledge graphs must be *designed for in-situ use*, as the information to be delivered through the graph is also generated by the user community, and is accessed in highly dynamic conditions (such as, for example, at a conference). Successfully identifying information categories (information design), representing data in them (schema and instantiation), and accessing the graph (user design), then, requires an explicit socio-technical design process. Knowledge graph projects typically commence with a process of defining competency questions; what we illustrate in this manuscript is that this typical design strategy must be augmented to identify a design concept appropriate to the unique context of a given community, capable of delivering value both for immediate needs as well as to support longer-term community support goals.

The gap that we attempt to address is that existing academic search platforms tend to privilege data that are more easily aggregated (e.g., publication data). Largely lacking from those existing systems are ways of incorporating activities that go on in a community that are not captured by publications or other more visible research artifacts, yet are instrumental to community functioning. As principles from library and information science and knowledge management have been used to inform the development of knowledge graphs delivering technical content, so are principles from design, communication, psychology, and ethnography relevant to the community knowledge graph.

The paper is structured as follows. First, we illustrate our work with and within the Center for HelioAnalytics through (Section 1.1). We then provide a detailed methodology, including the sociotechnical approach that adopted a human-centric philosophy of ontology engineering and incorporated ethnographic activities (Section 2). We demonstrate the application of the philosophy and ethnography in instantiating the ontology with survey data collected from the CfHA in the KG development sections (3 and 4). Finally, we discuss the broader significance of this work (Section 5).

1.1. The NASA center for HelioAnalytics

The NASA Center for HelioAnalytics (CfHA) is an exemplary case of a scientific community exposing the need for new forms of organization to manage increasingly complex demands of inquiry. CfHA emerged from the recognition that the study of a star, interplanetary space, and the interaction with planets (the domain of 'heliophysics' Eddy, 2009) required approaches that permitted systems-level understanding, the broad and robust use of data, and interdisciplinary collaboration. Out of this context came the term 'HelioAnalytics.' The Center for HelioAnalytics (CfHA) was established informally in 2019 and more formally under funding from the National Aeronautics and Space Administration (NASA) in 2020 to build the field and community of HelioAnalytics, by focusing on existing and important problems that had been obstinate to progress with established methods and required more modern approaches.

The objectives and inherently distributed, cross-group nature of the Center require a sophisticated approach to knowledge management and organization. CfHA is composed of a cross-disciplinary convergence of communities of physicists, statisticians, and computer scientists connected to NASA that is intended to foster fundamentally integrated research into advanced methodologies for heliophysical research, and

[1] https://www.earthcube.org/

[2] https://nsf-gov-resources.nsf.gov/2022-09/OKN%20Roadmap%20-%20Report_v03.pdf

[3] https://www.gartner.com/en/information-technology

[4] information structured into a graph form by a specific data model/schema/ontology that defines entities (objects, events, situations or abstract concepts) and their relationships. It is a collection of interlinked descriptions of entities—objects, events or concepts (Bonatti et al., 2018; Hogan et al., 2021; Boccaletti et al., 2006; Torres et al., 2021).

to promulgate such methods into the broader community. Thus, the CfHA must connect a diverse and changing community while making visible and accessible to it a wide variety of information and knowledge.

2. Methodology

Knowledge graphs typically focus on delivering technical information to users. This type of project emphasizes a user base of one: e.g., the user in direct exchange with the knowledge base. However, since a community knowledge graph is specifically focused on empowering a dynamic base of users to interact *with each other*, this typical design process is insufficient to deliver this goal. In effect, each of the components of the knowledge graph must support a system of users to engage with each other.

For this project we thus re-conceived the traditional knowledge graph design process, introducing a design phase which augments the identification of competency questions[5], the typical beginning of the technical construction of the graph itself. We note that there exists some work, mostly diffuse and of rather marginal focus, into working within a community to design a KG (e.g., 'customer involvement and feedback' principle within The eXtreme Design Methodology Blomqvist et al., 2016) upon which we build, however we adapt that previous research for the purpose of creating a community knowledge graph and that this requires novel approaches. Since community knowledge graphs are primarily focused on emerging tacit information, identifying *what* information should be supported by the resulting tool is likely to be a substantial activity on its own; user communities are often not yet aware of how technological tools may be able to support existing processes, even processes that are extremely high-friction at present (e.g., how do I know who in my group is working on what?). Further, what is important to a community may not be identified easily, even by conscientious users. This is largely because community knowledge graphs, by definition, exist outside of the scope of view of any given individual; thus, a number of "measurements", that is, samples of user stories, must be made, in order to comprehend the systemic operations of a community, and indeed, the prioritization of capabilities to build into the graph, to gain initial interest and enthusiasm.

Thus, we propose that a suitable design process is as integral to the production of the community knowledge graph as the representation itself, e.g the OWL file. The approach that we provide below was designed to the unique requirements of the CfHA use case, and serves as an instructive example; the particular design process must always be tailored to the specific context of the community it serves.

In our use case, we produce the resulting design concept through an integrated synthesis of CfHA institutional goals, member survey data, and research of best practices profiled in the literature and evident in industry. We structured our implementation to simultaneously develop a community of interest alongside necessary data capture and ontological design work, to support legibility and collaboration with members of the recipient community.

2.1. Establish community of interest

A *community of interest*[6] is a group of representatives from within the community for which the community knowledge graph is designed, who partner with the design and ontology team as in-situ collaborators, initial users, and informal champions of the project. A championing community is instrumental to promoting, adopting and maintaining the knowledge graph system into maturity, and ensure key design requirements are incorporated.

We deemed the establishment of a community of interest at project outset essential to the long-term success of this initiative, as scientists at CfHA are time – and attention – constrained, and can be skeptical of unfamiliar tools. Cohesive communities of interest are important for generating network effects for emerging technologies.

The CfHA is structured into numerous sub-groups. One of these subgroups, the 'Knowledge Team,' works on the information representation and infrastructure required by the CfHA to facilitate the interdisciplinary work of the Heliophysics data science community and was an ideal place to situate efforts to create a community of interest for this KG effort. It should be noted that the lead author is the lead of the Knowledge Team. The explicit objectives of the team are:

- Create and provide new resources to the Heliophysics community for improved information structuring (i.e., knowledge), such as: language, learning materials, and technologies to achieve information structuring;
- Develop the CfHA knowledge capture, organization, and access system;
- Connect with other knowledge teams;
- Develop a knowledge graph/system curriculum for Heliophysics community to connect sets of lectures with accompanying computational notebooks (e.g., Jupyter notebooks) and apply those curricular tools to make KGs and better information structuring possible for the Heliophysics community; and
- Build stronger relationships with data science/semantic web/ontology communities and existing information structuring efforts.

Biweekly Knowledge Team meetings are open to the broader membership at CfHA and often involve external collaborators and interested groups, and provided an ideal mechanism for socializing the project with users, as well as gaining ongoing design insights. Through the process of design and development of this knowledge graph, we also engaged in a number of informal conversations and observations which were continuously integrated into the design process.

2.2. Define design scope

Identifying an initial starting point for a community knowledge graph requires a survey of the general dynamics and needs of a community. While the complete activities of CfHA were impossible to address within the scope of this initial pilot, we identified a use case that addressed the foundational requirements of a user community, in order to provide a suitable basis to build future components.

In collaboration with members of the CfHA, the Knowledge Team identified several notions that capture the Center's twin goals of tracking the development of the CfHA community as well as the science produced by the Center:

- Who are the CfHA partners and collaborators?
- What has the CfHA produced?
- A scientist is considering a new project. How can they find out who has already tried something similar, worked with the data before, or used a particular model, repository, etc? How do they determine the state of existing knowledge as well as meaningful trends?
- NASA Headquarters wants to know the impacts of recent data science funding opportunities. How can those impacts be revealed?

With the Knowledge Team, we analyzed these notions to derive the core concepts and utility around which to form the initial structure of the community knowledge graph. For example, the concepts required to address the first notion—who are the CfHA partners and collaborators

[5] Competency questions are natural language questions outlining the scope of knowledge represented by an ontology (Uschold and Gruninger, 1996) and are generally the starting point for ontology design

[6] A collaborative group of users (working at the appropriate security level or levels) who exchange information in pursuit of their shared goals, interests, missions, or business processes, and must have a shared vocabulary for the information exchanged. The group exchanges information within and between systems https://csrc.nist.gov/glossary/term/community_of_interest

Table 1
Ontology requirement specifications.

Scope:	The scope of this ontology is to semantically describe the NamedIndividuals within a professional community, their role, expertise, collaborators, and activities
Intended Users:	Heliophysics researchers NASA disciplinary scientists from related disciplines Partners of the CfHA Knowledge Engineers
Intended Use:	Use Case 1: Identify who the CfHA partners with and the nature of those partnerships that may inform new collaborations. Use Case 2: Support new research projects to understand what has been done, who to approach with questions, and guide team composition. Use Case 3: Create a more connected, enjoyable, flourishing community.
Ontology Requirements:	**CQ1**: Who do we engage with the most? **CQ2**: Who do I have the most associations with? **CQ3**: What projects do we work on with this person/group? **CQ4**: What missions do we work with this person/group? **CQ5**: What skills/exercise does this person/group have? **CQ6**: What research does already exist in this topic?

– references components that are also necessary to address the second notion – what has the CfHA produced (as a result of member activities, with partners and collaborators). In both cases, the representation of human contributors is primary.

Once this design concept is identified, it is possible to re-use this design concept to address other, more complex notions. For instance, to address notion three, we may experiment with modeling strategies that attach models, repositories, and other forms of scientific work to human contributors.

Through conversations with the Knowledge Team, a single notion was selected as the primary focus of this project: "Who are the CfHA partners and collaborators?"

Competency questions

From this notion we then identified *competency questions* (CQs), *entities*, and *relationships*. These competency questions were used to structure the rest of the design activities, as well as the conceptual mapping of the entities and relationships and the testing mechanism for the utility and success of the eventual knowledge graph. They are:

- Who are we engaging with the most? Who do I have the most associations (of type $< X >$) with?
- What projects do we (a member or members of the CfHA) work on with this person or group? What missions do we work on with this person or group?
- What skills and expertise does this person or group have?
- What research or research artifacts already exist on this topic?

These CQs are presented alongside the scope, users, and use for the KG that come from them in Table 1.

The CAUSE ontology was formalized in a machine understandable format using the Web Ontology Language (OWL) language.

2.3. Survey real-world user context

While we identified the structural goals of the KG design in conjunction with the Knowledge Team, we also surveyed the CfHA member community directly for qualitative information about their working contexts, in order to identify immediately useful day-to-day tasks that we could also address. Our choice to include a human-centered design phase in our development process originated from the awareness that knowledge graph projects are often underutilized by their intended audience—despite being built intentionally to those audiences' specific requirements. We hypothesized that incorporating insights about community members' workflows into the knowledge graph's technical requirements would produce a final, integrated product that could be successfully operated by users in their actual working environments. We were conscious of the impact of often tacit limitations (such as users' capacity to learn new technical features, or to integrate new tools into various components of their workflows) on the success or failure of a software application.

The ethnographic survey included interviews with members of the CfHA via in-person conversations, whiteboarding, and demos, and interviews conducted by members of the Knowledge Team. Additionally, workshops (virtual and in-person) were held throughout 2022. The in-person activities also included iterative conceptual modeling and competency question design processes. To augment conversational discovery, a more formal survey instrument was designed and shared to acquire information from CfHA members. The next step, that is beyond the scope of this paper, is to create a user interface that reflects the community discovery and makes the knowledge graph instantiated in this paper accessible and navigable by the CfHA members.

2.4. Define initial design concept

After identifying the primary strategic goal of the Center, and surveying members to understand day-to-day user activities and goals, we synthesized our learnings into an initial design concept.

The Center is inherently distributed across disciplines, groups, institutions and projects. Efficiently making sense of the myriad activities requires a collective view of the partnerships, collaborations, community activities, and research artifacts (e.g., presentations and publications) of CfHA members.

The members of CfHA are involved in more than 100 associated projects (proposals, committees, papers, presentations, conferences, organizations). That presents a challenge: how can the CfHA team track its connections and activities in a valuable, minimally burdensome, and fully searchable way?

We wanted to develop an ontology that could augment 'traditional' ways of coordinating within a team: e.g., organizing and attending regular team meetings, meeting new people and engaging new ideas at conferences, maintaining personal email lists, and maintaining tacit knowledge about the liaisonships and partnerships with the team. Our long-term vision was to provide an easy vector through which to promulgate news to the entire network, reducing transaction costs of manual notification and removing barriers of scope. In turn, these benefits would optimize the self-organizing capacity of the community, and afford expanded collaboration potential.

We also identified several core design constraints that our solution needed to satisfy. Although beyond the scope of the initial project, adhering to these design constraints would be essential to support eventual implementation, and thus were essential to include for consideration during the initial design phase:

- To reduce barriers of adoption, the resulting application would have to be easily accessible;

- The resulting application must allow for contributions from the entire community; and
- The resulting application must facilitate users to provide continual data contributions, and thus must provide an incentive for doing so.

2.5. Review of related work

Once we had a clear view of the nature of the community knowledge graph we intended to develop, we next assessed relevant publications and examples of community knowledge graphs; the examples are relatively few and dispersed.

The term "community knowledge graph" has typically referenced large-scale, open-source projects, such as Wikipedia and its linked data counterpart Wikidata (Vrandečić and Krötzsch, 2014). These may be more accurately described as community-developed knowledge graphs, whose databases are formed through the contributions of a large and diverse set of people—e.g., "the community". However, in our conception of the community knowledge graph, community members are represented as the subject of the graph itself. Therefore, there is little overlap between the ontologies used by Wikipedia to represent domain context and the ontological concepts required for our use case.

We included the evaluation of traditional social networking software – which also use knowledge graphs – in our design assessment. For example, community-oriented knowledge graph software such as Facebook and LinkedIn establish broad databases of information about people. Although these ontologies are not public, we anticipate that the concepts represent the linkages of the social graph—e.g., such as those described in the Friend of a Friend (FOAF) ontology (Brickley and Miller, 2014).

There exist pockets of progress on community-based knowledge infrastructure, notably from the ontological engineering community (Gómez-Pérez et al., 2004). Like traditional social networking software, prominent examples of ontological engineering applications that include community components predominantly come from industrial or commercial applications (e.g., Blomqvist and Öhgren, 2008). A key distinction between these examples and our initial design vision, and resulting CAUSE ontology, is our focus on supporting a more cohesive community and supporting discovery and innovation, rather than a business outcome or knowledge management goal. However, previous work from the ontological engineering community exists to understand social connections at quite broad levels (Masolo et al., 2004).

In the scientific community, community knowledge representation is perhaps even less mature. NASA has investigated data representation to understand community: at Johnson Space Flight Center to support business considerations (David Meza, personal correspondence) and at the NASA Jet Propulsion Laboratory in the creation of an Institutional Knowledge Graph (Lewis McGibbney, personal correspondence). The Earth Science Informatics Partners (ESIP)[7] community investigated an Earth Science Knowledge Graph (ESKG) Testbed Project as an automatic approach for building interdisciplinary Earth Science knowledge graphs to improve data discovery (McGibbney et al., 2017). Most of these efforts have not published results nor do they maintain publicly accessible graphs and could be considered exploratory rather than functional.

Finally, arguably the biomedical community has by necessity developed more mature knowledge infrastructure than most physical sciences (e.g., the Monarch Initiative Mungall et al., 2016[8]) and there has even been discussion of how to collaboratively build ontologies for biomedical domains (Noy et al., 2008). These efforts have overwhelmingly focused on physical units (e.g., genes Ashburner et al., 2000) rather than the community element.

[7] https://www.esipfed.org/

[8] https://monarchinitiative.org/

2.6. Translate insights from related work into conceptual model of the ontology

Informed by our review of other examples and research related to community knowledge graphs, we next undertook a process of synthesis to make final design decisions, solidifying the design concept and competency questions in light of implementation considerations and strategies.

Defining How to Build: Incorporating Insights from Reviewed Materials We ultimately settled on a hybrid of the approaches: defining the interlinkages between community members as a social graph, expanded to include relationships to their own work products, topics of expertise, and other technical information, and enabling ongoing user contributions to populate the knowledge graph.

We were inspired by FOAF's representation of people and their social relationships in the virtual world. Our application inherits the need to represent virtual relationships and information, although it also includes a broader set of relationships, and required dynamic intersections such as community members' conference and meeting interaction, and participation in projects and missions.

Future extensions of the CAUSE ontology will benefit from the further consideration of ontology design patterns, to inform the development of ground and conceptual ontologies, providing at the ground layer a foundation of essential terms such as 'Person' in the CAUSE ontology presented below, which can be layered with additional temporal, contextual, and further conceptual meanings to produce specific representations of dynamic multi-agent community roles. Indeed, the "building blocks" that the modular approach to ontological engineering affords forms the foundation of the new solution that CAUSE aims to provide (Gangemi and Presutti, 2009; Blomqvist et al., 2016; Hammar and Presutti, 2016; Krieg-Brückner and Codescu, 2021; Shimizu et al., 2022).

Defining the Approach for Designing the Ontology We converged on a foundational design decision to treat human contributors as the primary unit of focus, as well as topic, in the graph.

We also adopted a multi-level modeling strategy, constructing both a top-down and bottom-up model separately, and working iteratively to ultimately converge these into the final ontology.

The goal of this blended approach was to develop a harmonious integration between the types of data and manners of description that a community member of CfHA would produce in the context of a working experience, and the strategic priorities of the community itself, as well as the organizational activities of the Knowledge Team and broader agencies (e.g., NASA).

Our multi-level design process is mirrored in both the resulting ontology and instantiated graph, which are composed of integrated, faceted information provided by diverse personas, which together compose the community ecosystem.

Below we describe these two simultaneous modeling approaches in more depth.

2.7. Top-down: Translating design into ontological concepts

We began by decomposing the competency questions into data elements necessary to answer those questions, as shown in Fig. 1.

The competency questions roughly described core concepts, such as "person", "group", and "project;" active qualities of those concepts, such as "action", "relationship", and "output;" and descriptive, annotative qualities useful for second-order analysis of relationships, such as notions of frequency and quality. The latter two categories were not directly modeled into the ontology but rather indicated a range of concepts requiring differentiation, to address dynamic questions such as: What types of actions are salient for the community to identify and report into the knowledge graph? What types of measurements must be included to define "who we engage with the most?"

Who	do we	engage with	the most?
Individual	Member of CfHA - individual	Know (types of relationships: have met, shared a panel with, published a paper with, connected with representative, etc.)	Measurements of frequency, duration, number of times; ranking of results (most likely by variety of criteria, to avoid implicit bias)
Group	Group Member of CfHA - individual	Meet (conferences, workshops, meetings, etc.)	
	Department of CfHA (group of groups)	Work with (on proposals, committees, papers, presentations, conference, organizations, collaboration (formal, informal)	
		Share disciplinary expertise with	
*Can also be applied to materials (e.g., which papers do we cite the most? Etc.)			

Fig. 1. Decomposing competency questions into data elements to determine the top-down component of the CAUSE ontology. This is a demonstration of the decomposition for the first CQ: "Who are we engaging with the most?".

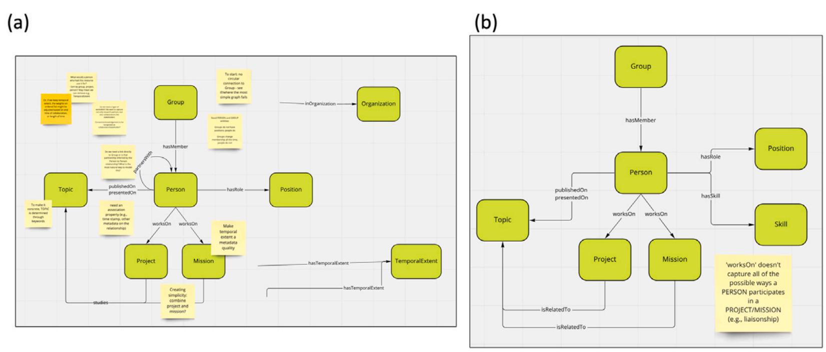

Fig. 2. Illustration of the evolution of design of the CAUSE ontology that extracted *entities* and *relationships* from a process of describing the CfHA team members' knowledge of the collaborations and partnerships use case. (a) Initial conceptual map, revealing a process of review and feedback with the CfHA community (post-it notes on the diagram); and (b) Resolving open comments and iterating the conceptual model. This evolution largely used the mind-mapping and diagramming software, Miro miro.com/.

After identifying these conceptual sets, we compared the collection of core concepts with the model draft that was produced from community contributions via Knowledge Team meetings. We sought to define a fundamental process map representing the activities of community members within work environments, which could function as the basis for the rest of the ontology. Fig. 2 shows a portion of the evolution of the process map, which proceeded from hand-written (not shown) to diagramming software (a couple of iterations shown).

Using the concept "people" as the core building block upon which to base our arrangements, we defined a structure that sufficiently represented all core concepts.

The resulting ontology extends across five axes:

- Topics (such as 'astrophysics,' 'machine learning,' 'the Sun')
- Objects (such as publications, papers, conferences)
- Activities (processes and projects)
- NamedIndividuals (community members, as well as teams, departments, organizations, etc.)
- Affiliations (relationships between NamedIndividuals and groups)

We specifically chose concepts that represent fundamental components of communities, and avoided academic jargon, in order to produce a pattern ontology that can be adopted by communities in other domains. This may fill an existing absence in knowledge representations: a core and broadly available ontology for the representation of professional communities.

We next defined relationships between these entities, of two types: (1) basic relationships to express the semantic context of the use case, such as (isContributor, isMemberOf), as well as (2) relationships for connecting topics, objects, activities, and role descriptors to human contributors. Affiliations were also designated as relationships. As an organizing principle, we attempted to select broad relationship tags that could be reused, such as "resultsIn" (applicable to many processes) versus "published" (narrower scope). Some of these relationships are structural, such as "isContributor", and thus, broadly applicable across domains and use cases; however, others, particularly Affiliation labels, are more specific to the CfHA, and would require tailoring for other use cases.

To represent notions of quality, we created additional "meta-metadata" terms, such as measurements of time (e.g., frequency, duration, number of times), as well as concepts for representing the outcome of various investments. Such concepts are of relevance to the Knowledge Team as well as NASA quite broadly (e.g., NASA Headquarters), and other decision makers who allocate resources across opportunities within the organization. To support the administrator's view on each of these levels, additional attributes were modeled, such as starting dates, resources, and quality.

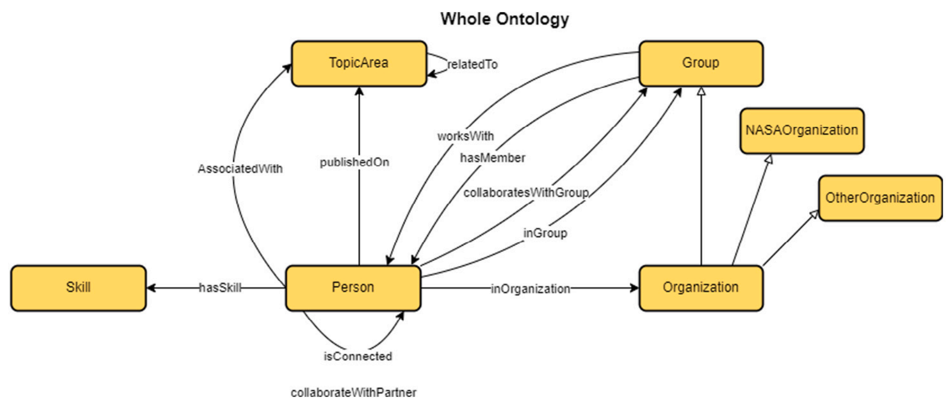

Fig. 3. Visualization of the CAUSE ontology.

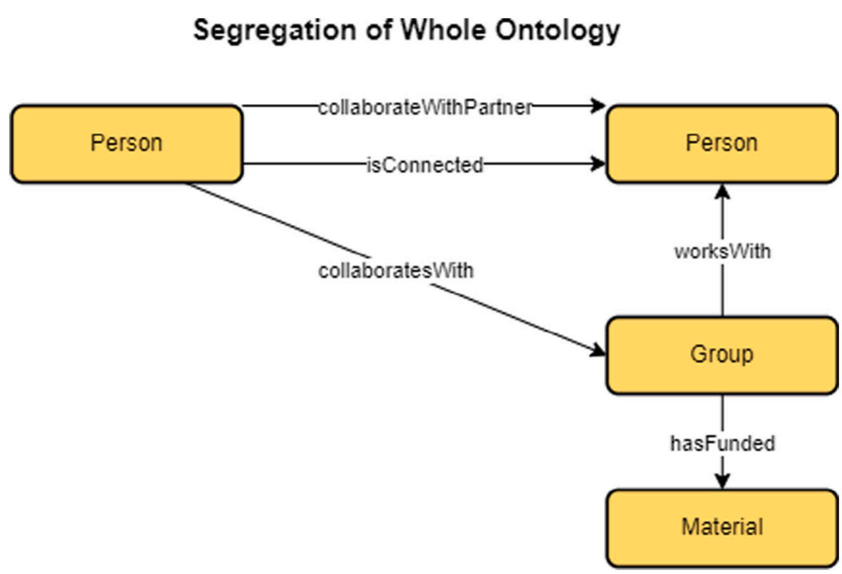

Fig. 4. Visualization of the segregated CAUSE ontology: the 'Person' component.

2.8. Bottom-up: Extracting concepts from collected data

The simultaneous bottom-up approach began with data collected from CfHA members in-situ. These data were analyzed for recurrent patterns, which could be described through the high-level categories defined in the previous section. This work followed the finalization of top-level ontological concepts, to verify that they were able to represent actual user data accurately and completely.

Biweekly calls with the Knowledge Team throughout the design process created regular feedback on the conceptual map. Additionally, we observed community members using the opportunity of the call to manually share data that could be alternatively organized through the completed knowledge graph, such as:

- Identification of which journals they relied on for various types of information discovery;
- Match graduate students with CfHA mentors who could provide advice on specific, technical project;
- Discover upcoming grant opportunities; and
- Compose teams to respond to grant opportunities.

Thus working directly with a community to create a KG leads to observing and discovering more granular and diverse data that are instrumental to community collaboration and functioning.

3. KG development and description

Coupled with the high-level design development (described in the previous section), we implemented the ontology in the Protégé software (Musen, 2015). Fig. 3 shows the Protégé visualization of the CAUSE ontology.

CAUSE further defines 'Person' and 'Project' components as shown in Figs. 4 and 5, respectively.

The human contributor-focused conception of the knowledge graph guided its development from the start and is reflected in the centrality of the 'Person' entity in the general classes of CAUSE (represented as a 'NamedIndividual,' more on this choice below).

After establishing general classes, information from the CfHA community was imported into the ontology. Data instances are responses to the CfHA survey we conducted. Each survey response was submitted by a member of the CfHA community, such that each is associated to a 'NamedIndividual.' We use the owl:NamedIndividual class for declaring named (in contrast to anonymous) individuals per the OWL 2 specification.[9] In the CAUSE ontology, this covers both 'Person' and 'Organization.'

Connections between nodes were either modeled as Object Properties or Data Properties. A note that while we used Protégé to build the

[9] https://www.w3.org/2007/OWL/wiki/FullSemanticsNamedIndividuals

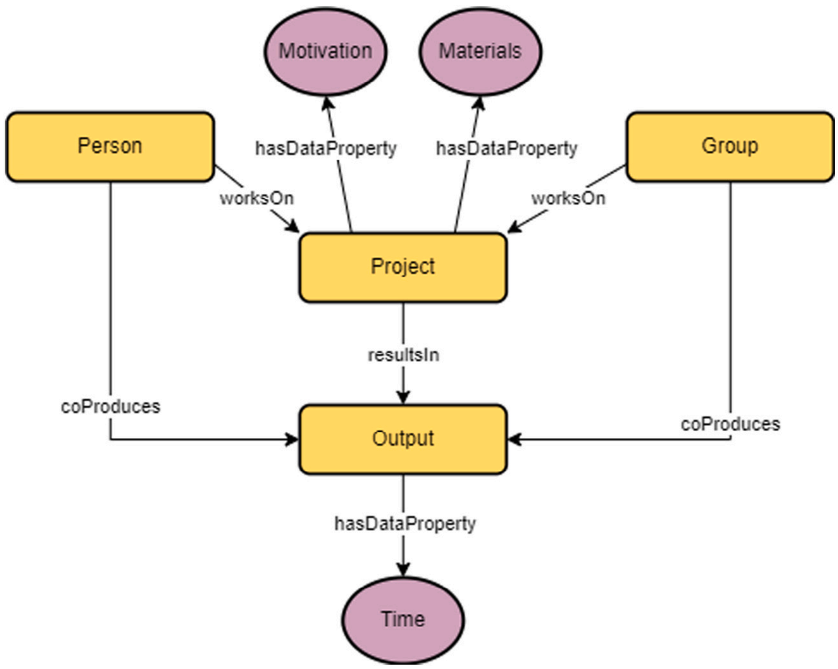

Fig. 5. Visualization of the segregated CAUSE ontology: the 'Project' component.

KG, it is not a graph database application. It was sufficient in this proof-of-concept paper because the data were small (~70 survey responses). To scale this work, a graph database would be required.

Semantic Web Rule Language (SWRL) rules (Lan, 2004) were written to support automatic instantiation of the knowledge graph; as an example, we wrote a rule based on a transitive property that classified a Topic to a Group if a Topic was associated with a Person, and a Person collaborated with a Group.

Finally, because we imagine there is an audience outside of knowledge engineers interested in this research, we provide two substantive components of the knowledge graph design that will be somewhat basic for more experienced knowledge engineers. Object Properties represent a connection between NamedIndividuals using a format that is similar to the grammatical structure of a sentence in English. The subject of the sentence is akin to the domain intersecting with the Object Property, or the acting Class. The verb of the sentence is akin to the Object Property itself, or the action being performed. Finally, the object of the predicate is similar to the range intersecting with the Object Property, or the Class being acted on. For example, the 'Earth and Space Science Department", which is a Group, has the Object Property assertion partOfGroup, which has the range 'NASA Jet Propulsion Laboratory.'

Data Properties represent a relationship between a 'Class' and a concrete piece of data. For example, a Person has a givenName, familyName, and potentially an additional middleName. This category of connection is utilized with tangible data that have not been modeled as an Individual.

3.1. Data sources

The 'gold standard' data set for the proof-of-concept knowledge graph from the CAUSE ontology were those collected directly from the CfHA community. The KG is designed such that data from other sources can be integrated (e.g., NASA's Astrophysics Data System (ADS; https://ui.adsabs.harvard.edu/) in a straightforward way: we only need to define the relationship between categories in the new database and those defined by our ontology. We have not done so in this proof-of-concept manuscript.

4. KG evaluation

CAUSE contains 36 classes, 51 object properties, and 17 data properties. It was tested using the Pellet reasoner and the Oops!(ontology pitfall scanner!) tool (Poveda-Villalón et al., 2014) to demonstrate that the schema is consistent. All CQs were converted to SPARQL queries and successfully executed towards the KG as demonstrated in the listing below. The outcome of the queries was validated for accuracy by the experts of the CfHA community.

•CQ1: Who do we engage with the most?

Listing 1: Who do we engage with the most?

```
SELECT (SAMPLE(?person) as ?p) (COUNT (DISTINCT ?y) AS
    ↪ ?groups)
WHERE {
?person :collaboratesWithGroup ?y }
ORDER BY DESC (?groups) GROUP BY (?person)
```

•CQ2: Who do I have the most associations with?

Listing 2: Who do I have the most associations with?

```
SELECT ?person ?partner
WHERE { ?person :collaborateWithPartner ?partner }
ORDER BY DESC (?partner)
```

•CQ3: What projects do we work on with this person/group?

Listing 3: What projects do we work on with this person/group?

```
SELECT ?project ?person ?group WHERE {
       ?person :worksOn ?project.
:collaboratesWithGroup ?group.
       :associatedWithTopic :AI
       }
ORDER BY (?project)
```

•CQ4: What missions do we work with this person/group?

Listing 4: What missions do we work with this person/group?

```
SELECT ?mission ?person ?group WHERE {
       ?person :worksOn ?mission.
:collaboratesWithGroup ?group.
       :associatedWithTopic :AI
       }
ORDER BY (?mission)
```

•CQ5: What skills/expertise does this person/group have?

Listing 5: What skills/expertise does this person/group have?

```
SELECT ?person ?group ?skill WHERE {
?person :hasSkill ?skill .
:collaboratesWithGroup ?group
}
ORDER BY (?person)
```

•CQ6: What research does already exist on this topic?

Listing 6: What research does already exist on this topic?

```
SELECT ?research ?topic ?person WHERE {
?person :associatedWithTopic ?topic .
?person :producesOutput ?research
}ORDER BY (?research)
```

5. Discussion and significance

The central achievement of this manuscript is the development of an ontology capable of answering six CQs central to NASA teams and groups; particularly those of a cross-disciplinary community, the CfHA. We believe that this ontology and the factors that motivate it, enabling community self-discovery and equality and inclusivity within, are perhaps an excellent foundation on which other researchers can build. The purpose of this section is to briefly describe our perspective on those opportunities to acknowledge that this ontology is not the end point of this research and to invite the reader to consider where they might build on these foundations. Because these thoughts are extrapolations from the research presented in this manuscript, we are intentionally brief.

First, the CfHA is an example of the kinds of broad and expansive groups that are becoming more common at NASA and all agencies. These groups require wider expertise and intelligence to address key societal problems (e.g., climate change) and they will have greater knowledge needs.

With any ontology and knowledge graph effort, a vital element is how those artifacts are made accessible to those without semantic technology expertise. Therefore, user interfaces (UI) that make the graph accessible, navigable, extensible will be needed. Elements of a UI need to enable members of a community to access existing and input new data about the community and their activities. Ultimately, the impact of an ontology and knowledge graph depends on the adoption by the community, so the UI must become a part of the community member's workflow. This is a significant challenge for all knowledge graph projects.

CAUSE is of high potential utility in large part because of the flexibility with which it could be used, its impact depending on the ingenuity with which a community uses it. We perceive several ways that the ontology and knowledge graph could be used to begin to realize some of this potential:

1. Enable rapid discovery and analysis of publications and other scientific materials;
2. Support triage for communication channels that proliferate across platforms (e.g., email, social media, chat apps, etc.); and
3. Structure coordination within and between organizations.

The ultimate hope is that CAUSE could enable more fluid cross-individual and cross-organizational efforts by reducing barriers to both discovery of intra- and inter-community members, an important component of the move towards *Open Science* (Vicente-Saez and Martinez-Fuentes, 2018; Gentemann et al., 2021, 2022). Central to the move towards open science is elevating implicit information held currently in personal networks to an explicit, open knowledge network (e.g., Janowicz et al., 2022; Zhu et al., 2022; Nelson et al., 2019; Shefchek et al., 2020).

The broadest goal for extending the CAUSE ontology foundation is to work towards a *knowledge commons*. A knowledge commons is a combination of intelligent information representation and the openness, governance, and trust required to create a participatory ecosystem whereby the whole community maintains and evolves this shared information space. McGranaghan et al. (2021), McGranaghan (2022). Indeed, the knowledge graph lends itself to the development of an ecosystem of knowledge communities:

- **Between colleagues**, facilitated by data entry about individuals community members, and for their personal and extended network;
- **Between researchers**, facilitated by data entered into a technical knowledge base, which can then be accessed and curated by specific individuals; and
- **Between levels of community**, engendered by the ability to become familiar with a range of groups in the general environment and dynamics between them.

In maturation, this knowledge graph may enable an increase in creative inputs between individuals and groups as well as the development of new relational ties. This is a particularly important goal in science, where intersecting perspectives shape what is known, and greater diversity leads to better results and more creative investigations (Hochberg et al., 2017). Generative collaborations precede scientific collaborations, and are originated through the crossing of new intersections. What would be the effect, then, of removing barriers to cross-disciplinary discovery? As this knowledge graph is oriented towards the needs of scientists at CfHA, the primary activities that this ontology supports are collaboration, peer discovery, and perusal of knowledge products produced by community members. However, with some modifications the CAUSE ontology is capable of addressing communications and coordination challenges within organizations at large. Those extensions involve support for information transmission across defined roles, processes, and hierarchies (Sequeda and Lassila, 2021).

6. Conclusions

We presented a knowledge engineering solution to the fact that the challenges faced by science, engineering, and society are increasingly complex, requiring broad, cross-disciplinary teams to contribute to collective knowledge, cooperation, and sensemaking efforts. The result of our solution is a new ontology to support cross-community knowledge sharing and discovery: the Community Action and Understanding via Semantic Enrichment (CAUSE) ontology. Both the CAUSE ontology and the design approach developed to create it are significant contributions to the semantic web community. CAUSE was designed with a specific use case, the NASA Center for HelioAnalytics (CfHA), whose cross-disciplinary and cross-institutional membership exemplifies the challenges of the modern scientific team (Council, 2015). Using data collected from the CfHA and in partnership with them, we instantiated the CAUSE ontology to provide a proof-of-concept knowledge graph.

Finally, we discussed the implications of our community-based approach to knowledge graph creation, which is to enable community stewarding of the graph and open science practices.

CRediT authorship contribution statement

Ryan M. McGranaghan: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Ellie Young:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Cameron Powers:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft. **Swapnali Yadav:** Data curation, Formal analysis, Investigation, Resources, Software, Validation, Visualization, Writing – original draft. **Edlira Vakaj:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

The authors are deeply appreciative of the NASA Center for HelioAnalytics, funded by NASA ISFM, USA, for supporting this work and creating a community in which conversations like these occur. Software and publicly available artifacts supporting this work are available from a Github repository (https://github.com/rmcgranaghan/Helio-KNOW). The final ontology is available at https://github.com/rmcgranaghan/Helio-KNOW/blob/main/User_Interface_with_Interactive_Knowledge_Graph/latest_updated_ontology_file/CfHA_Ontology_V_4.2.owl. R. McGranaghan gratefully acknowledges support of NASA Grant 80NSSC21K0622 for this work.

References

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H.L., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S.E., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: Tool for the unification of biology. Nature Genet. 25, 25–29.

Blomqvist, E., Hammar, K., Presutti, V., 2016. Engineering ontologies with patterns - The extreme design methodology. In: Ontology Engineering with Ontology Design Patterns.

Blomqvist, E., Öhgren, A., 2008. Constructing an enterprise ontology for an automotive supplier. Eng. Appl. Artif. Intell. 21, 386–397.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U., 2006. Complex networks: Structure and dynamics. Phys. Rep. 424, 175–308. http://dx.doi.org/10.1016/j.physrep.2005.10.009, URL https://www.sciencedirect.com/science/article/pii/S037015730500462X.

Bonatti, P.A., Decker, S., Polleres, A., Presutti, V., 2018. Knowledge graphs: New directions for knowledge representation on the semantic web (dagstuhl seminar 18371). Dagstuhl Rep. 8, 29–111.

Brickley, D., Miller, L., 2014. FOAF vocabulary specification 0.99. xmlns.com. FOAF project. Archived from the original on 2022-03-03. Retrieved 2022-03-31.

Chatterjee, A., Nardi, C., Oberije, C., Lambin, P., 2021. Knowledge graphs for covid-19: An exploratory review of the current landscape. J. Pers. Med. 11.

Council, N.R., 2014. Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and beyond. The National Academies Press, Washington, DC, http://dx.doi.org/10.17226/18722, URL https://nap.nationalacademies.org/catalog/18722/convergence-facilitating-transdisciplinary-integration-of-life-sciences-physical-sciences-engineering.

Council, N.R., 2015. Enhancing the Effectiveness of Team Science. The National Academies Press, Washington, DC, http://dx.doi.org/10.17226/19007, URL https://nap.nationalacademies.org/catalog/19007/enhancing-the-effectiveness-of-team-science.

Eddy, J.A., 2009. The sun, the earth, and near-earth space: A guide to the sun-earth system.

Gangemi, A., Presutti, V., 2009. Ontology design patterns. In: Handbook on Ontologies.

Gentemann, C., Erdmann, C., Kroeger, C., 2022. Opening up to open science. Issues Sci. Technol. 37, 57–59.

Gentemann, C.L., Holdgraf, C., Abernathey, R.P., Crichton, D.J., Colliander, J.E., Kearns, E.J., Panda, Y., Signell, R.P., 2021. Science storms the cloud. AGU Adv. 2.

Gómez-Pérez, A., Fernández-López, M., Corcho, Ó., 2004. Ontological engineering: With examples from the areas of knowledge management, e-commerce and the semantic web. In: Advanced Information and Knowledge Processing.

Hammar, K., Presutti, V., 2016. Template-based content odp instantiation. In: WOP@ISWC.

Hochberg, M.E., Marquet, P.A., Boyd, R.T., Wagner, A., 2017. Innovation: An emerging focus from cells to societies. Philos. Trans. R. Soc. B 372.

Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., d. Melo, G., Gutierrez, C., Gayo, J.E.L., Kirrane, S., Neumaier, S., Polleres, A., Navigli, R., Ngomo, A.C.N., Rashid, S.M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., Zimmermann, A., 2021. Knowledge graphs. Commun. ACM 64, 96–104.

Janowicz, K., Hitzler, P., Li, W., Rehberger, D., Schildhauer, M., Zhu, R., Shimizu, C., Fisher, C.K., Cai, L., Mai, G., Zalewski, J., Zhou, L., Stephen, S., Gonzalez, S., Mecum, B.D., Carr, A., Schroeder, A., Smith, D., Wright, D.J., Wang, S., Tian, Y., Liu, Z., Shi, M., D'Onofrio, A.M., Gu, Z., Currier, K., 2022. Know, know where, knowwheregraph: A densely connected, cross-domain knowledge graph and geo-enrichment service stack for applications in environmental intelligence. AI Mag. 43, 30–39.

Krieg-Brückner, T., Codescu, M., 2021. Generic ontology design patterns: Roles and change over time. ArXiv abs/2011.09353.

Lan, H., 2004. Swrl : A semantic web rule language combining owl and ruleml.

Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., Guarino, N., 2004. Social roles and their descriptions. In: KR.

McGibbney, L.J., Jiang, Y., Burgess, A.B., 2017. ESIP's Earth science knowledge graph (ESKG) testbed project: An Automatic approach to building interdisciplinary earth science knowledge graphs to improve data discovery. In: AGU Fall Meeting Abstracts. IN33C–0131.

McGranaghan, R.M., 2022. The evolution of heliophysics: Complexity, community, and open science. Front. Astron. Space Sci. URL https://api.semanticscholar.org/CorpusID:251956688.

McGranaghan, R., Klein, S.J., Cameron, A., Young, E., Schonfeld, S., Higginson, A., Ringuette, R., Halford, A., Bard, C., Narock, A., Thompson, B., 2021. The need for a space data knowledge commons. Struct. Collective Knowl. URL https://knowledgestructure.pubpub.org/pub/space-knowledge-commons. https://knowledgestructure.pubpub.org/pub/space-knowledge-commons.

Mungall, C.J., McMurry, J.A., Köhler, J.P., Borromeo, C.D., Brush, M.H., Carbon, S., Conlin, T., Dunn, N.A., Engelstad, M., Foster, E., Gourdine, J., Jacobsen, J.O.B., Keith, D., Laraway, B., Lewis, S.E., Xuan, J.N., Shefchek, K.A., Vasilevsky, N.A., Yuan, Z., Washington, N.L., Hochheiser, H., Groza, T., Smedley, D., Robinson, P.N., Haendel, M.A., 2016. The monarch initiative: An integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res. 45, D712–D722.

Musen, M.A., 2015. The protégé project: A look back and a look forward. AI matters 1 4, 4–12.

Nelson, C.A., Butte, A.J., Baranzini, S.E., 2019. Integrating biomedical research and electronic health records to create knowledge-based biologically meaningful machine-readable embeddings. Nature Commun. 10.

Noy, N., Tudorache, T., d. Coronado, S., Musen, M.A., 2008. Developing biomedical ontologies collaboratively. In: AMIA ... Annual Symposium proceedings. AMIA Symposium. pp. 520–524.

Poveda-Villalón, M., Gómez-Pérez, A., Suárez-Figueroa, M.C., 2014. Oops!(ontology pitfall scanner!): An on-line tool for ontology evaluation. Int. J. Semantic Web Inf. Syst. (IJSWIS) 10, 7–34.

Sequeda, J., Lassila, O., 2021. Designing and building enterprise knowledge graphs.

Shefchek, K.A., Harris, N.L., Gargano, M.A., Matentzoglu, N., Unni, D.R., Brush, M.H., Keith, D., Conlin, T., Vasilevsky, N.A., Zhang, X.A., Balhoff, J.P., Babb, L., Bello, S.M., Blau, H., Bradford, Y.M., Carbon, S., Carmody, L., Chan, L.E., Cipriani, V., Cuzick, A., Rocca, M.G.D., Dunn, N.A., Essaid, S., Fey, P., Grove, C.A., Gourdine, J.P.F., Hamosh, A., Harris, M.A., Helbig, I., Hoatlin, M.E., joachimiak, marcin pawel, Jupp, S., Lett, K.B., Lewis, S.E., McNamara, C., Pendlington, Z.M., Pilgrim, C., Putman, T., Ravanmehr, V., Reese, J.T., Riggs, E.R., Robb, S.M.C., Roncaglia, P., Seager, J., Segerdell, E., Similuk, M.N., Storm, A.L., Thaxon, C., Thessen, A.E., Jacobsen, J.O.B., McMurry, J.A., Groza, T., Köhler, S., Smedley, D., Robinson, P.N., Mungall, C.J., Haendel, M.A., Munoz-Torres, M.C., Osumi-Sutherland, D., 2020. The monarch initiative in 2019: An integrative data and analytic platform connecting phenotypes to genotypes across species. Nucleic Acids Res. 48, D704–D715.

Shimizu, C., Hammar, K., Hitzler, P., 2022. Modular ontology modeling. Semantic Web.

Torres, L., Blevins, A.S., Bassett, D.S., Eliassi-Rad, T., 2021. The why, how, and when of representations for complex systems. SIAM Rev. 63, 435–485.

Uschold, M., Gruninger, M., 1996. Ontologies: Principles, methods and applications. Knowl. Eng. Rev. 11, 93–136. http://dx.doi.org/10.1017/S0269888900007797.

Vicente-Saez, R., Martinez-Fuentes, C., 2018. Open science now: A systematic literature review for an integrated definition. J. Bus. Res. 88, 428–436. http://dx.doi.org/10.1016/j.jbusres.2017.12.043, URL https://www.sciencedirect.com/science/article/pii/S0148296317305441.

Vrandečić, D., Krötzsch, M., 2014. Wikidata: A free collaborative knowledgebase. Commun. ACM 57, 78–85. http://dx.doi.org/10.1145/2629489.

Zhu, R., Janowicz, K., Mai, G., Cai, L., Shi, M., 2022. Covid-forecast-graph: An open knowledge graph for consolidating covid-19 forecasts and economic indicators via place and time. In: AGILE: GIScience Series.