Bridge-Concepts: Establishing Harmonized Networks of Ontologies

Francesco A. Zaccarini[1], Arkopaul Sarkar[2], Emanuele Ghedini[1] and Ilaria M. Paponetti[1]

[1]*University of Bologna (DIN), Bologna, Italy*
[2]*Université de Technologie Tarbes Occitanie Pyrénées, Tarbes, France*

Abstract
This study introduces a tool and methodology for the creation of harmonized networks of ontologies, a pre-condition for the full exploitation of data in federated distributed systems. Bridge-Concepts are designed to alleviate well-known challenges in ontology mapping and to address network-specific issues, such as scalability and consistency. As standalone ontology entities, they function as data pipelines in hub-and-spoke structures. Designed with FAIR (Findable, Accessible, Interoperable, and Reusable) principles in mind, their rich informal characterization makes them user-friendly interfaces and candidates for a vocabulary tailored specifically for ontology usage. Bridge-Concepts form the central element of a network-specific alignment methodology based on pragmatic criteria that can be further improved by introducing high-level ontologies and automatic tools in the loop. This approach is based on an analysis of the limits of meaning-encoding within semantic artifacts.

Keywords
Applied Ontology, Bridge-Concepts, Network Harmonization, Methodology, Ontology Alignment

1. Introduction

Formal ontologies are one of the core knowledge representation technologies and the fundamental infrastructure for the Semantic Web; however, their practical effectiveness in supporting interoperability is impeded by the existence of a plurality of frameworks –even with overlapping, or equivalent, domains of application. Not only is there a prevailing inclination among industrial stakeholders to prefer ontologies developed internally (to exert greater control over proprietary data), but different ontological frameworks can exhibit varying degrees of suitability with respect to specific pragmatical goals, making a pluralistic approach actually desirable, especially in industrial contexts [1]. Given the difficulties and drawbacks associated with the creation, and imposition, of an universal standard, ontology harmonization[1] has emerged as a valid, albeit not unproblematic, alternative: indeed, the process is complex, time-consuming and error-prone. Problematically, the benefits of interoperability increase exponentially with the number of ontologies, and elements per ontology, linked.

This study introduces a methodology and related tools (referred to as "Bridge-Concepts") to ensure the comparability of core ontology entities employed by different ontologies in order to set up a FAIR-compliant network of partially aligned, harmonized ontologies. These minimal, pinpointed data pipelines are meant to support effective integration and interoperability among a plurality of knowledge bases. The discussion will proceed as follows: Section 2 provides a short introduction to relevant issues concerning meaning-encoding in semantic artifacts (2.1), as well as a general framework for the evaluation of alignments (2.2) and an overview of issues specific to ontology networks (2.3). In Section 3, Bridge-Concepts are introduced, following Bridge-Concept templates' structure. Section 3.1 elaborates on Bridge-Concepts' formal role, explaining how they support mediated alignments among

Proceedings of the Joint Ontology Workshops (JOWO) - Episode X: The Tukker Zomer of Ontology, and satellite events co-located with the 14th International Conference on Formal Ontology in Information Systems (FOIS 2024), July 15-19, 2024, Enschede, The Netherlands.
francesco.zaccarini@unibo.it (F. A. Zaccarini); arkopaul.sarkar@uttop.fr (A. Sarkar); emanuele.ghedini@unibo.it (E. Ghedini); ilaria.paponetti2@unibo.it (I. M. Paponetti)
0009-0008-8009-5009 (F. A. Zaccarini); 0000-0002-8967-7813 (A. Sarkar); 0000-0003-3805-8761 (E. Ghedini); 0009-0002-8345-0295 (I. M. Paponetti)

[1]Usually semantic alignments (matchings), without adjustments to the involved artifacts – to avoid interference with usage [2].

ontologies, and how they address heterogeneity, while Section 3.2 presents their role as a "ontology-specific vocabulary", promoting FAIR-ness and alleviating issues related to lack of documentation. Section 3.3 explains how a contained set of Bridge-Concepts can establish a controlled, open network, touching on points related to Bridge-Concept engineering, framework consistency, and the possibility of improving the system by exploiting High-Level, Foundational Ontologies, and automatic tools.

Tool and methodology were developed in the context of two European Projects, OntoCommons (https://ontocommons.eu/) and OntoTrans (https://ontotrans.eu/), and are part of larger toolbox building on the Linked Open Terms (LOT) approach [3]. This paper focuses on theoretical foundations and general principles; the reader can refer to OntoCommons' D 2.9 (available at https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e503ae3f85&appId=PPGMS) and [4], for a more extensive discussion of practical and implementation aspects, as well as examples of usage.

2. Background

2.1. Ontologies & Meaning-Encoding

Computational ontologies serve as tools for data structuring, integration and retrieval. They play a preeminent role in knowledge representation by providing schemas for knowledge bases (e.g., knowledge graphs), thus supporting interoperability and knowledge discovery, among other things. They can be understood as a representation of certain systems or as a systematization of domains of discourse [5].

Per the orthodox definition, ontologies are *formal, explicit specifications of shared conceptualizations* [6][7]. The nuances of "explicit" and "shared" are often underappreciated. Ontologies are *explicit* inasmuch as the core ontology entities (classes, and related properties) are formally characterized *intensionally*, following, the classic Carnapian approach adding the layer of possible worlds to extensionality [8]. In practice, this means that the responsibility of characterizing ontology entities, restricting the admitted interpretations (and thus models) to align as closely as possible with the intended ones, is delegated to axiomatization. However, it is arguably operationally impossible to have rich enough formal characterizations to avoid unintended interpretations; and even assuming that it was theoretically feasible, ontologies would likely become computationally intractable well before reaching that point, at least given current technological limitations. Hence, while axiomatization establishes negative constrains in interpretation, helps in clarifying concepts, and distributes meaning across the network (multiplying the number of hinges, thus reducing ambiguity), it is ultimately *labels, and informal documentation* attached to ontology entities, which is responsible for *semantic grounding*, through the medium of an interpreter situated in a given *context*. The reliance on informal elements, *e.g.,* (human) interpreters, can be considered one of the core limits of current semantic technologies.

Hence conceptualizations have to be *shared*: for ontologies to be effectively employed, all users have to converge towards (approximately) the same intended interpretations of the ontology entities, and, most importantly, of the assumed primitives. The criticality is accentuated by the fact that the domain which should be the target of the interpretation is not universally accepted by all interpreters, *i.e.,* it is an empirical fact that different interpreters, and even the same interpreter across different contexts or timeframes, may subscribe to slightly divergent worldviews; moreover, even without getting into tangled issues concerning meaning indeterminacy, vagueness and referential failure are well-known and widespread phenomena. It is pivotal to recognize that users needn't share exactly the same concepts, or a worldview across the board: in the same way as communication through natural languages is not compromised by speakers' idiosyncrasies and borderline cases, ontologies are effective insofar as they deal successfully with most cases, and the remaining ones cause no significant practical frictions.

2.2. Heterogeneity & Harmonization

Euzenat [9][10] delineates a non-rigid classification of heterogeneity types, which can shed light on harmonization (understood in terms of resolution of heterogeneity). Leaving aside heterogeneity types that can be dealt with through the adoption of W3C's implementation recommendations, it is possible to distinguish *terminological*, *semantic/conceptual* and *semiotic/pragmatic* forms of heterogeneity.

Terminological heterogeneity occurs due to variations in identifiers and labels among ontology entities which purportedly refer to the same world entities. Differences pertaining to identifiers are the standard if ontologies are developed separately; differences pertaining to the labels can be due to the use of different natural languages ("gatto" [IT]; "cat" [EN]), cases of synonymy ("coat"; "jacket"), or preferences of specific communities. The latter can be tackled through synset analysis and more complex techniques employed by language models; however, *interpretation* remains the gold standard due to polysemy, jargons, vague or misleading labeling and the salience of contextual variants in practical scenarios.

Semantic/conceptual heterogeneity has to do with formal idiosyncrasies in modeling a given domain. This may manifest as the utilization of diverse concepts, or different choices in axiomatization, and is closely related to points discussed in Sections 1 & 2.1. Semantic heterogeneity is for instance exemplified by geometric theories adopting different primitives, leading to entirely different axiomatizations. This is not problematic when the resulting theories are logically equivalent, as direct correspondences between the entities involved could be established. However, the inherent incompleteness of ontologies, coupled with their worldview/use/goal-specific nature, makes this an exception rather than the rule. In interesting cases, reliance on interpretation is thus necessary. Therefore, informal elements are crucial.

Finally, *semiotic/pragmatic heterogeneity*, comes down to idiosyncrasies in interpretation proper, for ontology entities which appear to be (terminologically and semantically) similar, given context/user variance. The most relevant cases involve discrepancies among "false friends", which can, if undetected, have significant consequences given erroneous alignments. The issues related to the absence of a univocal interpretation discussed above are greatly magnified if ontologies are brought outside their context – a precondition for the establishment of interoperability.

All these types of heterogeneity should be addressed in ontology harmonization. However, to respect pluralism and to allow for non-disruptive integration, this can only be done indirectly. Terminological heterogeneity can be resolved by (not) establishing identities among individual constants and equivalences among classes or properties. Semantic heterogeneity can be addressed by setting up semantic connections among ontology entities. Semiotic heterogeneity can only be tackled by producing links that take into account actual usage in practice and other informal documentation. This arguably requires collaboration with the stakeholders employing the ontology, *i.e.,* domain experts, and close scrutiny of related knowledge bases. An ideal alignment between two ontologies would be such that information shared through mappings is indistinguishable from re-conceptualizations of the relevant systems/domains, and any form of redundancy is avoided. Needless to say, ideal alignments are purely a theoretical limit. Discarding the possibility of working directly on knowledge bases, which would undermine the benefits of employing semantic technologies while still incurring the aforementioned interpretative issues, such alignments could only be established among (locally) equally expressive ontologies, *i.e.,* already (locally) semantically and semiotically harmonized ontologies. Nevertheless, this limit can be used as a reference point to establish a metric of alignments' informativeness and error, as well as qualitative, and potentially also quantitative, criteria to guide alignment choices.

For the sake of simplicity, and in line with the examples that will be discussed below to present Bridge-Concepts, let us focus on *classes* (C) and *individual constants* (I). A set of mappings from a certain ontology $\mathcal{O}_1$ to an ontology $\mathcal{O}_2$ is *maximally information-preserving* if, as a result of the mapping, all the individual constants $i_1, i_2, ..., i_n \in I_{\mathcal{O}_1}$ would be classified under all the classes $X_1, X_2, ..., X_n \in C_{\mathcal{O}_2}$ the referred-to entities (given a hypothetical intended interpretation) would be independently conceptualized under. *Informativeness* can thus be understood in terms of a simple proportion. Similarly, *error* can be defined as the ratio of individual constants that are categorized under classes that are either disjoint from, or subclasses of, the most specific class under which the referred-to entities would be conceptualized in the target ontology. Extrapolating from common practice in ontology harmonization, the ideal goal is maximizing informativeness and minimizing error. However, in practical applications, other factors come into play. Balancing informativeness and error involves trade-offs, to be evaluated in light of contextual factors. For example, it may be acceptable to tolerate a higher error to increase informativeness, provided that this does not cause *practical* disruptions.

Additionally, scalability is a crucial factor to consider, especially when harmonizing a large number of ontologies. Thus, the aim shifts towards achieving an operationally optimal balance.

2.3. Harmonizing Ontology Networks

Harmonization is further complicated by contingent factors. Domain/Application-level ontologies are often axiomatized using languages of limited expressiveness. This can be due to the ontologies' specialized intended use, but also to the need to reduce development costs. Ontologies often either lack documentation, or it can be not accessible/not clear. Specialized jargon is often employed, with examples and terminologies understandable only to those deeply involved in the ontology's use. FAIRness sensibility is also a recent theme. Moreover, ontologies often contain outright mistakes (in conceptualization) or inconsistencies among documentation and axiomatization. This is partly due to the issues mentioned above, and partly because ontologies are "living artifacts": they are shaped by use and undergo subtle changes throughout their lifecycle. In addition, most applications, especially in industrial contexts, require informative alignments with no errors concerning *contextually salient links*, *i.e.,* they require high specificity in individuals' discrimination. And this list is far from exhaustive.

More issues can be identified for the harmonization of large ontology *networks*. First, and foremost, harmonization usually involves pairwise alignments through semantic links. Problematically, the number of required alignments increases exponentially ($\frac{n(n-1)}{2}$) with the number n of ontologies involved. Concatenations of alignments, *i.e.,* alignments mediated by sets of ontologies/alignments, are often less informative. This is especially true given heterogeneity and the diverse domains of application/coverage granularity. Establishing well-documented and richly axiomatized *Reference ontologies* [11][12] might seem an optimal solution. However, the problems outlined in the introduction concerning stakeholders' divergences can reappear, it is difficult to deal with inconsistencies in conceptualization, and the task can be outright monumental depending on the variance of the set of ontologies involved. Nevertheless, this approach can be effective if core stakeholders are interested in establishing a standard for a specific domain, though, notably, the resulting ontology may render the ones to be aligned largely obsolete, so the solution may not be fully conservative.

Another challenge for any approach has to do with usability. Given a set of ontologies and alignments, tractability and operational usability can easily be lost. Fortunately, stakeholders are usually interested in only a specific subset of the network. Therefore, it has to be possible to access and use only a fragment of the network, *i.e.,* the network has to be modular. Due to stakeholders' diverse needs and desiderata, flexibility is also necessary when it comes to alignments' required expressiveness, and the availability of alignments differently balancing informativeness and error. Finally, the harmonized network ought to be *plastic, update-friendly and reusable*, allowing for the introduction of new ontologies, and adjustments to the harmonized ontologies and the relative knowledge bases.

2.4. Core uptakes

To summarize: **(i)** Even with richly axiomatized foundational ontologies, the encoded meaning will never truly be fully transparent, especially to machines. **(ii)** Ontologies' (and ontology entities') adequacy should be evaluated based on whether they reduce ambiguity below break-points determined by contextual factors. **(iii)** Let the *formal characterization* of an ontology entity be the subset of its set of axioms involving said entity and ontological entities recursively related to the entity, and the *informal characterization* of an ontology entity encompass all the related annotations and documentation (including labels, descriptions, comments), and even contextual factors related to actual usage of the ontology; **(iii.a)** informal characterizations are at least as important as (and complementary to) formal characterizations for usability, and **(iii.b)** ultimately more salient when it comes to ontology harmonization, as the ontologies to be connected are often built referring to different goals/aims/stakeholders/use cases. **(iv.a)** Harmonization procedures must fully account for both formal and informal aspects to produce adequate results in non-trivial scenarios, and **(iv.b)** it shouldn't be expected there to be a single "correct" alignment independently of context and pragmatic choices, but rather a plurality of approximations with different pros and cons.[2] **(v)** In establishing networks of ontologies, *scalability*, *modularity* and *flexibility* are of the utmost importance. **(v.a)** Scalability requires mappings' partiality, and a strategy to ensure that the number of alignments does not depend exponentially on the number of ontologies involved, without an excessive loss of informativeness. **(v.b)** The resulting framework has to be *modular* and *flexible* as a precondition for its practical adequacy, given computational requirements and the diverse plethora of involved stakeholders.

Overall, and in a motto, both ontology engineering, and ontology harmonization are a matter of "fit, rather than match". Under this respect, they can arguably be assimilated to other forms of representation and communication, where negotiation and friction-reduction occupy the center stage.

3. Bridge-Concepts

The proposed tool, and related ontology network harmonization methodology, has been designed around the points listed in Section 2.4. It revolves around the creation of a limited set of *FAIR, well-documented, standalone ontology entities* (*Bridge-Concepts*), establishing *pinpointed mediated alignments* among a set of ontologies at *core network junctures*, individuated *bottom-up* from the network's overall structure (at a given time), taking into due account knowledge bases and use-cases' salience.

Table 1
Bridge-Concept template: General Information.

Concept Name	The label, preferred label, or Internationalized Resource Identifier (IRI) title used to identify the Bridge-Concept.
IRI	Suggested Bridge-Concept IRI.
OWL Type	A value between: Class OR ObjectProperty OR DataProperty OR Individual.
Domain	The domain(s) the Bridge-Concept was built for. More can be included, possibly organized taxonomically. This serves as a first source of disambiguation for domain experts and users in general.
Concept Elucidation	This provides a natural language, informal definition of the concept, intended to be easily understood by domain experts. Elucidations should align with common knowledge and domain resources, avoiding references to other ontology entities (*i.e.,* they should be *ontology neutral*). Ideally, they should also remain *ontologically neutral* (avoid commitments beyond the domain they pertain to) and be concise, with the inclusion of diverse usage *examples* (a plurality of them, to avoid prototyping effects) and the explicit *addressing of potential ambiguities*, focusing on *contextually* salient cases relevant for (the expected) ontology usage.
Labels	Labels used to refer to the concept, categorized as follows: **(i)** *preferred label* – the primary label for referring to the concept, combining intuition and informativeness; **(ii)** *alternative labels* – multiple labels commonly used to address the concept, even if they have narrower or wider meanings; **(iii)** *deprecated labels* – labels that may be misleading or encourage misuse, but which are used in practice. **(iv)** *Hidden labels* can also be included to support *queries*.

Being engineered as standalone ontology entities, Bridge-Concepts initially lack a formal characterization. This feature, which might initially appear puzzling (especially considering the principles of semantic technologies), finds an immediate explanation in their role of "mediators" among diverse formal conceptualizations, as well as support in the discussion above concerning the limits of ontology entities' formal characterizations, tractability and context-sensitivity. The emphasis is thus on Bridge-Concepts' informal characterization, which pivots on stakeholders' domain expertise while avoiding non domain-specific commitments (unless strictly necessary), referring to salient gold standards, and

[2]The difficulties encountered in setting up effective automatic harmonization tools focusing on structural and terminological approaches can find an explanation in the points just discussed. It is thus important to investigate scalable manual alternatives.

Table 2
Bridge-Concept template: Knowledge Domain Resources.

Knowledge Domain Resources	It lists existing domain resources, such as standards, books, articles, and dictionaries considered during the development of the Bridge-Concept. The template includes static references to these resources and quotations of relevant content. Multiple resources can be reported; renown resources that have a high likelihood of having influenced users' conceptualizations are given priority. These resources act as reference points in the engineering phase and (together with the related comments) help domain experts better understand the Bridge-Concept, enhancing conceptual clarity.
Comments	Comments in this section explain the motivations underlying engineering choices in the elucidation, drawing from domain resources and highlighting similarities and differences; the discussion should aim at solving possible ambiguities not fully addressed in the elucidation.

Table 3
Bridge-Concept template: Alignment to Existing Ontologies

Target Ontology	This section includes the IRI of one of the ontologies encompassing ontology entities which are aligned to the Bridge-Concept. This part of the template is replicated for each ontology.
Related Ontology Entities	A list of IRIs of specific ontology entities (belonging to the target ontology) to which the Bridge-Conceptis connected to.
Mapping Comments	This section provides an extensive discussion (in natural language) of the mapping (between the Bridge-Concept and the target ontology entities) choices and the underlying rationale. It includes contextual information concerning the intentionally adopted trade offs between informativeness and error, considerations about possible alternative mappings considered and the evidence gathered in support of the choices made, facilitating third-party evaluation and validation of the proposed connection, and contributing to the clarification of the Bridge-Concept.
Type of Alignment	A description of the kind of mapping established. E.g., strongly hierarchical (such as owl:EquivalentClass or rdfs:SubPropertyOf), weakly hierarchical (*e.g.,* skos:narrower), of similarity (*e.g.,* skos:related). The latter can be employed to enhance the framework's querying capabilities, and for advanced analytic graph-based approaches.
Mapping Axioms	Proposed mapping axiom(s) between the Bridge-Concept and the ontology entities are provided in an OWL2 compliant syntax, such as Turtle, Manchester, RDF/XML, Functional-Style, or OWL/XML. Notably, the mappings can be *complex* [2], but, as a maxim, formulas expressible in weaker languages should always be preferred, to increase interoperability and reusability for a diverse set of stakeholders. In some cases, it might be beneficial to provide different sets of (consistent) axioms, given different OWL profiles, depending on the specific scenarios.

addressing pragmatically (ontology usage-wise) salient ambiguities. The core aim of the characterization is, in fact, allowing users to situate Bridge-Concepts with respect to their own conceptualizations. This grounded approach aims to go ways towards avoiding common failures in usability due to lack of documentation, the impossibility of providing complete formal characterizations, and preconceptions related to the need to capture entities' essences in conceptualization. Hence, absolute priority is given to the goal of engineering useful concepts capable of connecting a network of ontologies, understandable by relevant users, and capable of reducing the emergence of frictions in practice under contextually determined acceptable thresholds, following the communicative principle of "fit rather than match". Bridge-Concepts play two roles:

(1) they establish scalable mediated alignments among a plurality of ontologies, functioning as *data pipelines*;

(2) they double as practical concept vocabulary-entries tailored for ontology implementation, acting as a user-friendly interface for stakeholders, including both end-users and ontologists, thereby improving usability and understanding.

Before delving into the presentation of Bridge-Concepts' two core roles, it might be beneficial to get acquainted with the tool. A template, divided in three parts depending on the core stakeholders addressed, is proposed for documenting bridge concepts. Notably, the template acts as a human interface, while being implementation-friendly in .owl file format. Examples produced in the context of the OntoCommons project in cooperation with practitioners are available at https://github.com/OntoCommons/OntologyFramework, and the implementation schema, as well as practices to ensure FAIR-ness, are described in the already cited OntoCommons' D 2.9.

The first part of the template (Table 1) is relevant for all users, and contains the core elements constituting the informal characterization of a Bridge-Concept. The second part (Table 2) holds particular significance for domain experts as it encompasses links with gold standards and other knowledge domain resources which better situate the concept. The third part (Table 3) delineates the formal mappings with existing ontology entities, along with pertinent information regarding the latter.

3.1. Bridge-Concept-Mediated Alignments

In order to fulfill their formal role, *i.e.,* establish mediated alignments among ontologies, Bridge-Concepts themselves have to be semantically connected to the target ontologies through axioms. In the best case scenario, for each relevant ontology in the network, there should be one ontology entity equivalent to a given Bridge-Concept. In practice, this would only be feasible if all the involved ontologies covered approximately the same domain(s) and revolved around the same core concepts, with operationally and semantically consistent formal (and informal) characterizations. Usually, given an effective choice of Bridge-Concepts, it is possible to individuate both a super and a sub class/relations, without the need to make use of complex alignment axioms to establish sufficiently informative mediated connections.

Let us consider a practical example (see Fig. 1), involving a Bridge-Concept developed in the context of the OntoCommons project, as well as two salient industrial ontologies part of the OntoCommons EcoSystem (OCES) connected to it, SAREF and IOF-Core.[3] Specifically, we consider $C_{SAREF} = \{Device, Sensor, SmokeSensor, TurbididtySensor, TemperatureSensor\}$ and relative axioms A_{SAREF}, and $C_{IOF} = \{PieceOfEquipment, MaterialArtifact\}$ and relative axioms A_{IOF}. The engineered Bridge-Concept, with preferred label "Equipment", was aligned to SAREF and to IOF-Core as follows: $Device \subseteq Equipment$, $Equipment \subseteq PieceOfEquipment$, $PieceOfEquipment \subseteq Equipment$. Given the ontology Network $\mathcal{N}$, such that $C_{\mathcal{N}} = C_{SAREF} \cup C_{IOF} \cup \{Equipment\}$ and $A_{\mathcal{N}_1} = A_{SAREF} \cup A_{IOF} \cup A_{Equipment}$, it trivially follows that *e.g.,* $Sensor \subseteq MaterialArtifact$, whereas the first class belongs to SAREF, and the second to IOF-Core. Thus, SAREF and IOF-Core's ontology entities are semantically linked through OntoCommons' Bridge-Concept "Equipment", allowing the exchange of data, and individual constants to be imported. As per the example, Bridge-Concepts effectively function as a data pipeline. Intuitively, data "flows upwards": all the data covered by the sub-ontology entities connected to the Bridge-Concept is made available to the entirety of the network. Conversely, reasoning "flows downwards": the axioms characterizing super-ontology entities connected to the Bridge-Concept can be exploited by the entirety of the network: *e.g.,* in the example all the axioms A_{IOF} involving $PieceOfEquipment$, $MaterialArtifact$, as well as BFO's superclasses, constrain SAREF's $Device$, and its subclasses, following the alignment.

In principle, two kinds of Bridge-Concepts could be distinguished, depending on whether the focus is establishing "vertical" connections among ontologies at different levels, or "horizontal" connections among ontologies at the same level – provided that the two functions are not mutually exclusive, and the classification is contextually dependent on the set of ontologies part of the network considered, as well as use-case-related factors. Vertical connections can be particularly useful for validation and

[3]See https://saref.etsi.org/ and https://spec.industrialontologies.org/iof/, respectively.

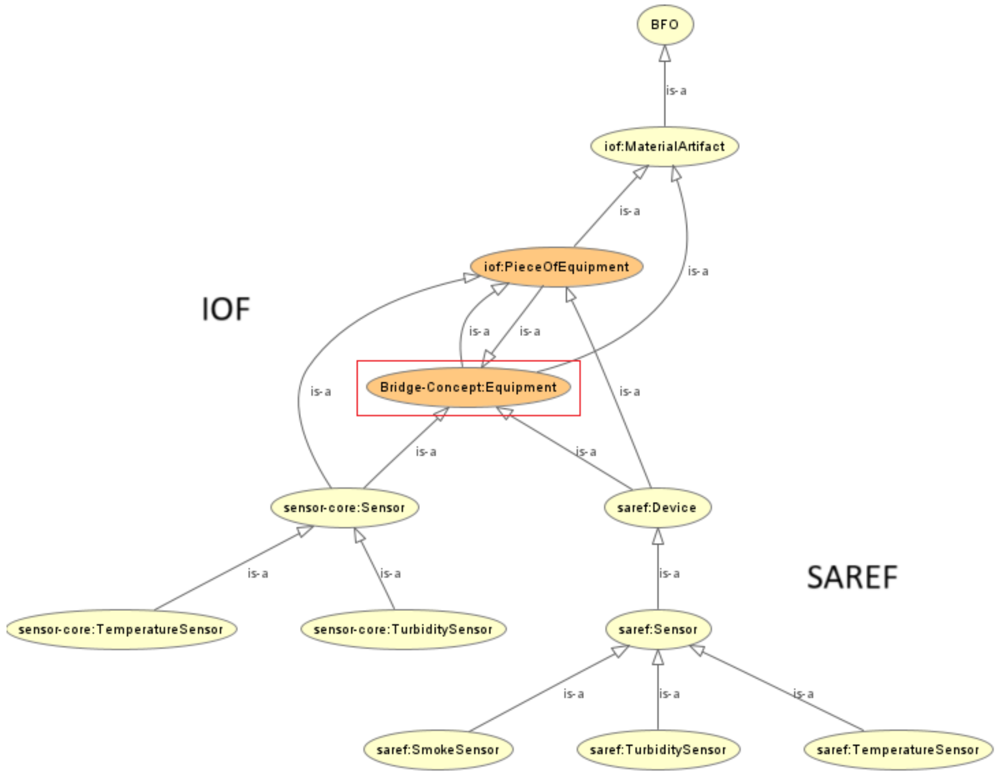

Figure 1: Establishing connections among ontologies through Bridge-Concepts. OWLViz visualization.

avoiding double-counting, especially if High-Level ontologies are included in the network (as in the example). Notably, given standard High-Level ontologies' architectures, few Bridge-Concepts could acts as seeds to ground a domain-level ontology on a top-level ontology capable of providing foundations [13]. Conversely, horizontal connections play a crucial role in ensuring efficient data sharing and the integration of specialized modules and related reasoning. Consequently, they are often regarded as the cornerstone of interoperability by stakeholders.

Notably, the semantic alignments supported by Bridge-Concepts facilitate the reduction of various kinds of heterogeneity. If informativeness and error is properly considered in the alignment process, they significantly contribute to addressing semiotic heterogeneity among the involved ontologies. Semantic heterogeneity is managed through the sharing of reasoning and the establishment of mediated correspondences between ontology entities, resulting in more comprehensive and multifaceted formal characterizations. Moreover, even terminological heterogeneity is addressed to the extent that entities formally linked through a mediated semantic correspondence collapse, rendering labeling/IRI variants inconsequential within the integrated ontology network (likewise, "false friends" are made readily discernible through a lack of mediated correspondences). Thus, in line with the outlined desiderata, these indirect solutions accommodate the pluralistic needs of stakeholders without necessitating any changes to the original ontologies, while establishing interoperability at the network level.

It is worth reiterating that, if necessary, Bridge-Concepts can be linked to ontology entities through complex axioms, in pursuit of optimal trade-offs between informativeness and error. Additionally, while a simplified case involving classes was presented, in scenarios revolving around knowledge bases, focusing on object and data properties might prove beneficial, although this might pose greater challenges in the alignment phase. Furthermore, it is evident that any number of ontologies can be linked to a given Bridge-Concept. This aspect goes ways towards addressing two core issues in establishing a harmonized network of ontologies, namely scalability (specifically, the number of alignments) and flexibility. These points will be elaborated on in Sec. 3.3; however, it is worth anticipating that Bridge-Concepts' standalone nature allows different stakeholders to select different sets of Bridge-Concepts (and ontologies), supporting spot connections on demand. Hence, through alignments, Bridge-Concepts are ultimately provided with network-specific, extendable and plastic formal characterizations. Moreover, similar to Reference Ontologies, sets of Bridge-Concepts can serve as hubs in a hub-and-spoke structure, ensuring that the number of alignments scales linearly with the number of ontologies to be harmonized, providing advantages with respect to standard approaches already with three ontologies involved.

3.2. A FAIR Vocabulary for Ontology Use

As per the discussion above, informal characterizations are crucial when it comes to usability (and thus reusability) and the ultimate hinges for alignment procedures. This aspect arguably underscores the demand among stakeholders for controlled, shared vocabularies, alongside a renewed emphasis on documentation and the adoption of FAIR principles. With their focus on informal characterizations, and being FAIR-by-design, Bridge-Concepts can serve as vocabulary-entries tailored for ontology usage.

Bridge-Concepts are chiefly characterized through their elucidations, but also via labels, the specification of the more relevant domains, as well as through the connections with domain knowledge resources and existing ontologies' concepts, which are explicitly discussed in their documentation. Elucidations of Bridge-Concepts are crafted to be easily comprehensible by domain experts, leveraging their expertise. They are not strictly definitions, as they do not provide necessary and sufficient conditions referring to other concepts; instead, they are intended to guide stakeholders in making accurate intensional judgments. In essence, they must:

> **(1)** strike a balance between **(1a)** *flexibility* (*i.e.,* they have to be intuitive given the assumed background knowledge, rather than a set of formal constraints that might seem obscure or non-core from the viewpoint endorsed by different stakeholders) and **(1b)** *rigidity* (*i.e.,* they have to provide pragmatically well-defined boundaries, necessary for effective machine implementation);
>
> **(2)** maintain explicit, detailed connections **(2a)** with primary knowledge domain resources and **(2b)** ontology entities they are meant to be semantically connected with, depending on the importance of the relevant ontology in the network and/or for a use-case;
>
> **(3)** align with common sense, that being the ultimate foothold for interpretation negotiation.

To attain the desired level of detail, Bridge-Concepts' elucidations should specifically address ambiguities relevant to their prospective usage. Unlike standard definitions, it can be informative for ontology use to specify that certain traits are not discriminatory, particularly when a concept is more coarse-grained. For example, when engineering a Bridge-Concept centered on atoms for chemical ontologies, it's pertinent to specify whether both standalone and bonded entities are included, and whether they may have an unbalanced number of electrons with respect to their atomic number: in fact, core stakeholders take divergent stances on this specific matter, as testified both by knowledge domain resources and the characterizations (both formal and informal) of relevant ontology entities. At times, offering specific examples and counterexamples can be effective. However, priority should be given to general principles to avoid prototyping effects, especially if the concepts are coarse-grained. Finally, brevity is a desirable characteristic, although achieving a harmonious balance among these requirements is a complex endeavor. A lengthy elucidation might increase the risk of stakeholders overlooking core points. It should be emphasized that the pragmatic aim is to effectively guide stakeholders, rather than precision itself. This necessitates an iterative process of refinement and adjustment. Elucidations should follow this standard internal organization: **(1)** introduction leveraging domain experts' background knowledge; **(2)** informal description with implicit references to selected gold standards and ontology entities; **(3)** notes on the use of adjacent concepts in the domain; **(4)** resolution of ambiguities through the explicit individuation of traits and values commonly cited; **(5)** possible examples and counterexamples.

Moving on, the selection of the preferred label holds particular significance as it constitutes the initial and most prominent element influencing a user. Therefore, preferred labels should be designated as the final step in the Bridge-Concept engineering process. In certain instances, prioritizing clarity over immediacy by making labels explicit might be advantageous and prevent potential misunderstandings, especially if the relevant Bridge-Concepts were developed with specific objectives in mind. Finally, the annotations discussing connections and discrepancies with respect to knowledge domain resources and existing ontologies' concepts (once a Bridge-Concept is aligned to the latter) should be comprehensive. Standards serve as reference points; thus, stakeholders might find these annotations more enlightening than the elucidations themselves.

A noteworthy aspect of Bridge-Concepts is their potential to serve as a standardized vocabulary. While primarily designed for pragmatic purposes, they offer the prospect of evolving into recognized standards themselves, if they prove effective. It is essential to acknowledge that Bridge-Concepts aim to establish unifying connections rather than supplanting other resources, a fundamental concern within the domain of standards and meta-ontologies' ecology, undermining long-term reusability and interoperability. Finally, like Bridge-Concepts are formally characterized through the alignment to ontology entities belonging to a set of ontologies making up the core of a network, said ontology entities' documentation is indirectly enriched by their connection to Bridge-Concepts, addressing one of the core issues outlined in Section 2.4. Notably, Bridge-Concepts are FAIR-by-design, with all the sections in the template above being directly implementable in a machine-readable environment (.owl), either as elements of an ontology, or as annotations, following a standardized schema in line with W3C recommendations. They are meant to be associated with IRIs, and made available in maintained repositories and commonly employed portals. Thus, Bridge-Concepts can significantly enhance the FAIR-ness level of individual ontologies and the overall network.

In their role as ontology-specific vocabulary-entries, Bridge-Concepts can be assimilated to (degenerate) content ontology design patterns [14], and can fulfill some of their functions to enhance ontology design and reusability. Notably, an ontology deliberately incorporating an ontology entity equivalent to a Bridge-Concept will seamlessly integrate into the relevant ontology network. However, ontology design content patterns appear to be more effective for ontology design, offering standardized, modular, and formal solutions that exemplify best practices. Conversely, Bridge-Concepts are arguably more appropriate for ontology harmonization, being engineered bottom-up for that very purpose. In fact, being standalone entities, they are more easily connectable and less formally committed.[4]

3.3. Establishing Harmonized Networks and selecting Bridge-Concepts

As shown by the example, individual Bridge-Concepts can serve as the foundation for mediated alignments among sets of ontologies, facilitating data and reasoning sharing while potentially significantly enhancing local clarity and FAIR-ness. However, establishing a fully harmonized network of ontologies typically necessitates a multitude of links. This leads to considerations regarding the selection of which Bridge-Concepts to engineer for a given set of ontologies (and given practical applications), as well as related issues concerning the framework's maintenance. Indeed, scalability has been identified as one of the core challenges in establishing harmonized networks. While the target ontologies need only be partially aligned to support effective interoperability, the effort required to engineer a single Bridge-Concept (including both characterization and alignments) makes it mandatory to keep their number contained. Once again, an abundance of (potentially low-quality) Bridge-Concepts would be counterproductive, potentially diminishing their findability and reusability.

Delving into the selection procedure in detail exceeds the scope of this introductory paper. Nevertheless, several options can be outlined. One approach involves conducting a statistical analysis of the terms present in the ontologies to be harmonized, or in a subset forming the core of a potentially expandable network. Although purely terminological analysis is susceptible to the limitations outlined throughout the discussion, the frequency of terms can serve as a reliable indicator of the salience of underlying concepts, provided a sufficiently large sample of ontologies is available. If the framework is to remain open and expandable, the results of the analysis can be supplemented with candidates directly selected by experts to mitigate deviations stemming from the idiosyncrasies of the initial set of core ontologies. A similar strategy has been employed within the context of the OntoTrans and

[4]It's worth noting that specific use-cases might benefit from the utilization of semantically connected and inter-defined clusters of concepts or outright concept patterns to establish connections among multiple ontologies. However, exploring this topic further is beyond the scope of this brief, general introduction to Bridge-Concepts.

OntoCommons European Projects as detailed in [4]. This approach can be immediately refined in two ways: first, by employing automatic alignment tools in candidate selection (largely circumventing core issues related to the precision of the tools and the inherent opacity of ontologies); second, by incorporating weighting based on an analysis of the architecture/structure of the involved ontologies, utilizing mathematical techniques from graph theory and network sciences, and possibly considering relevant knowledge bases, with informativeness, error, and the number of Bridge-Concepts serving as evaluation metrics.[5] In principle, following the introduction of an initial set of Bridge-Concepts to establish the network, new ones can be created to meet stakeholders' specific needs, progressively refining the pool of reusable tools. Indeed, Bridge-Concepts are designed to be decentralized, both in terms of engineering (requiring domain experts' knowledge) and alignment (allowing each stakeholder to connect their ontology to relevant Bridge-Concepts), making crowd-sourcing an option. Active stakeholder participation has the potential to significantly alleviate scalability issues and enhance the overall network's quality. Notably, Bridge-Concepts are theoretically reusable across networks and could serve as a tool for universal ontology interoperability.

When considering a plurality of Bridge-Concepts, issues related to network consistency arise: let Δ be a set of ontologies $\mathcal{O}_1, \mathcal{O}_2, \ldots\ \mathcal{O}_n$ and $A_\Delta = A_{\mathcal{O}_1} \cup A_{\mathcal{O}_2} \cup \ldots\ \cup A_{\mathcal{O}_n}$ its axioms, and α and β standalone Bridge-Concepts, A_α and A_β being the related alignment axioms. While the consistency of $A_\Delta \cup A_\alpha$ and $A_\Delta \cup A_\beta$ is ensured by the consistency of the single ontologies, and of the Bridge-Concept-specific sets of alignment axioms (provided that the ontologies are not otherwise connected), nothing guarantees the consistency of $A_\Delta \cup A_\alpha \cup A_\beta$. While this inconsistency could suggest issues in Bridge-Concepts' alignments, it could also stem from conceptual or architectural mistakes in the ontologies to be harmonized, which cannot, nor should, be rectified through the alignment itself. Ideally, a network of ontologies should strive for full consistency. However, since Bridge-Concepts are standalone, the resulting network is inherently modular, allowing for ways to manage inconsistency. Specifically, let $\mathcal{N}$ be a network composed of a set of ontologies $\mathcal{O}$ and a set of Bridge-Concepts with related mapping axioms $\mathcal{BC}$; it is possible to select a *consistent* sub-network $\mathcal{S}$ including only a certain set of ontologies $\mathcal{O}_\mathcal{S}$ such that $\mathcal{O}_\mathcal{S} \subseteq \mathcal{O}$ and a certain set of Bridge-Concepts $\mathcal{BC}_\mathcal{S}$ such that $\mathcal{BC}_\mathcal{S} \subseteq \mathcal{BC}$, picking the most relevant ontologies and connections. Likewise, stakeholders can leverage the resulting network's modularity to suit their use cases, focusing on the part of the network that interests them. Nonetheless, extra caution should be exercised in data imports to prevent error escalation.

Networks can be improved through the inclusion of one of more High-Level ontologies, possibly independently *formally aligned* with each other. Aligning Bridge-Concepts to them first, can facilitate further alignments and prevent misalignments, providing a first form of validation. The benefits of leveraging foundational High-Level ontologies for ontology alignment are well-documented in the literature [15], and the prospect of including them through Bridge-Concepts accessible to domain experts can arguably be considered an additional advantage of the proposed tool. This strategy has already yielded positive results in the context of the OntoCommons project, with the creation of the OntoCommons EcoSystem (OCES). Among other things, in this context the inclusion of High-Level ontologies served to facilitate and partially validate Bridge-Concept-mediated alignments between SAREF and IOF-Core (which were presented as an example). In general, alignments among superclasses can facilitate the establishment of alignments among leaf classes. Consequently, it is worth mentioning that the links among ontology entities established by Bridge-Concepts might in turn be exploited as constraints for automatic alignment tools, with the potential of greatly increasing informativeness, with limited errors, further improving scalability [16][13].

[5] Some of these options are explored in the already cited D 2.9 https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e503ae3f85&appId=PPGMS, including a tentative workflow and weighting formulas. Notably, it might be possible to leverage AI-based approaches to enhance the methodology and tailor it to specific contexts.

4. Concluding remarks

This paper offered a concise introduction to a methodology to establish harmonized networks of ontologies by employing Bridge-Concepts, standalone ontology entities tailored to facilitate mediated semantic alignments at pivotal junctures and address issues related to documentation and FAIR-ness. The proposed framework is tailored for the establishment of large harmonized networks of ontologies for federated distributed systems, supporting spot-connections for data and reasoning-sharing, requiring no changes to the ontologies to be harmonized, and granting stakeholders the possibility to isolate specific network segments pertinent to their use cases. The approach is particularly suitable for industrial settings, involving a plurality of stakeholders and value-chains extending over a number of different domains, requiring high informativeness and reduced errors in data sharing, and where considerations about data control and the reduction of operational interference are of paramount concern.

Still, several points require further exploration: specifically, Bridge-Concepts' reliance on manual alignments makes them susceptible to all the associated issues. Additionally, further extensive field testing is essential to identify potential bottlenecks in the procedures for selecting and engineering Bridge-Concepts. Moreover, potential issues may arise due to diachronic changes in ontologies – while the framework should be flexible enough to deal with them, maintenance costs have to be accounted for. Tentative answers to these challenges are outlined in the documents referenced in the introduction; a more extensive discussion is deferred to future publications.

Acknowledgments

This research would not have been possible without the funding from OntoCommons (GA N^o 958371) and OntoTrans (GA N^o 862136). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the involved Consortia. We thank Claudio Masolo, the members of the Consortia, and FOMI's reviewers & participants for the valuable comments.

References

[1] N. Adamovic, Ontocommons roadmap v1, 2023. URL: https://doi.org/10.5281/zenodo.7544509.
[2] I. Osman, S. Ben Yahia, G. Diallo, Ontology Integration: Approaches and Challenging Issues, Information Fusion 71 (2021) 38–63.
[3] M. Poveda-Villalón, et al., Lot: An industrial oriented ontology engineering framework, Engineering Applications of Artificial Intelligence 111 (2022) 104755.
[4] A. De Baas, et al., Review and alignment of domain-level ontologies for materials science, IEEE Access 11 (2023) 120372–120401.
[5] K. Munn, B. Smith, Applied Ontology: An Introduction, Frankfurt: ontos, 2008.
[6] T. R. Gruber, A translation approach to portable ontology specifications, Knowledge Acquisition 5 (1993) 199–220.
[7] R. Studer, V. Benjamins, D. Fensel, Knowledge engineering: Principles and methods, Data & Knowledge Engineering 25 (1998) 161–197.
[8] N. Guarino, D. Oberle, S. Staab, What is an Ontology, Springer-Verlag, Berlin Heidelberg, 2009.
[9] J. Euzenat, Towards a principled approach to semantic interoperability, in: Proc. IJCAI 2001 workshop on ontology and information sharing, Seattle, United States, 2001, pp. 19–25.
[10] J. Euzenat, P. Shvaiko, Ontology matching, 2nd ed., Springer-Verlag, Heidelberg (DE), 2013.
[11] V. V. Cross, P. Silwal, D. Morell, Using a reference ontology with semantic similarity in ontology alignment, in: International Conference on Biomedical Ontology, 2012.
[12] T. Hahmann, S. Stephen, Using a hydro-reference ontology to provide improved computer-interpretable semantics for the groundwater markup language (gwml2), International Journal of Geographical Information Science 32 (2018) 1–34.

[13] P. Barcelos, et al., Inferring ontological categories of owl classes using foundational rules, in: Proceedings of Formal Ontology in Information Systems 2023, 2023.

[14] V. Presutti, A. Gangemi, Content Ontology Design Patterns as Practical Building Blocks for Web Ontologies, in: Q. Li, et al. (Eds.), Conceptual Modeling, Springer Berlin, 2008, pp. 128–141.

[15] D. Schmidt, C. T. dos Santos, R. Vieira, Analysing top-level and domain ontology alignments from matching systems, in: OM@ISWC, 2016.

[16] P. Lambrix, Q. Liu, Using partial reference alignments to align ontologies, in: L. Aroyo (Ed.), The Semantic Web: Research and Applications, Springer Berlin, Berlin, 2009, pp. 188–202.