applied sciences

MDPI

Article

A Framework for the Evaluation of the Cultural Heritage Information Ontology

Andrej Tibaut * and **Sara Guerra de Oliveira**

Faculty of Civil Engineering, Transportation Engineering and Architecture, University of Maribor, 2000 Maribor, Slovenia; sara.guerraoliveira@um.si

* Correspondence: andrej.tibaut@um.si; Tel.: +386-2294382

Abstract: The intelligent management of built cultural heritage, including heritage buildings, requires common semantics in the form of standardized ontologies to achieve semantic interoperability. Foundational ontologies should be reused when building new ontologies, as they provide high-level terms; however, candidate foundational ontologies should be evaluated for quality and pitfalls. Simple metrics (e.g., number of concepts) are easy to obtain with existing tools. Complex metrics such as quality of ontology structure, functional adequacy, transferability, reliability, compatibility, maintainability, and operability, are defined in recent ontology evaluation frameworks; however, these do not evaluate interoperability features. The paper proposes an improved framework for an automated ontology evaluation based on the OQuaRE framework. Our approach improved some of the metrics of the OQuaRE framework and introduced three metrics for assessing the interoperability of the ontology in question (Externes, Composability, and Aggregability). In the experimental section, the framework is validated in an evaluation of cultural heritage information ontology (CIDOC CRM—ISO 12217:2014) with the use of new software for ontology evaluation. The detailed results reveal that the ontology is minimally acceptable and that the improved evaluation framework efficiently integrated interoperability metrics. Recommendations for the improvement of the cultural heritage information ontology are described in the Discussion and Conclusions section.

Keywords: ontology evaluation; interoperability metrics; cultural heritage; heritage buildings

check for updates

Citation: Tibaut, A.; Guerra de Oliveira, S. A Framework for the Evaluation of the Cultural Heritage Information Ontology. *Appl. Sci.* **2022**, *12*, 795. https://doi.org/10.3390/app12020795

Academic Editor: Cristiano Riminesi

Received: 19 November 2021
Accepted: 10 January 2022
Published: 13 January 2022

1. Introduction

The cultural heritage domain can be divided into tangible and intangible heritage. Historical buildings, monuments, and archaeological sites can be categorized as tangible cultural heritage. The justified interest in built heritage, combined with the technological advances in the construction field, emphasizes the need for reliably structured information in this domain. The associated size and multidisciplinary components justify the development of specific detailed ontologies, and an important and valuable step in this process is their evaluation. In essence, ontologies provide the backbone for the complete formal description of the concepts connected to a domain, their naming, definitions, properties, and relationships between them.

Generally, ontologies for the broader AECOO domain where cultural heritage and heritage buildings belong are rare, and this paper presents an original systematic evaluation. The presented improved evaluation framework is content agnostic and therefore applicable to cultural heritage domain ontologies.

Construction informatics [1,2] have grown in appeal in the highly structured knowledge datasets for the architecture, engineering, construction, owner–operator (AECOO) domain because the built environment has several knowledge-intensive subdomains, namely: technical plans, quantity take-off, construction contracts, construction safety, data exchange, and linked data [3–8]. The digitalization of the (to-be and as-is) built environment, as enabled by building information modelling (BIM) [9,10], sets the foundation for research and applications of knowledge engineering on the growing number of BIM models for existing buildings. The conversion of BIM models to a web ontology language (OWL) representation of the industry foundation classes (IFC) schema, ifcOWL [11], upgrades them to building knowledge models. The link between heritage buildings (HB) and BIM is highlighted in recent research literature focused primarily on sustainable retrofitting [12–16] and on-site interventions [17–19].

Historical buildings, part of tangible heritage, and built cultural heritage are of interest for domain experts in culture, management, tourism, sociology, and natural sciences (e.g., biology), all of which contribute knowledge to tasks connected to their rehabilitation, including interventions such as:

- Preservation: The Burra Charter [20] defines preservation as the maintenance of a building in its current state, retarding deterioration. Aimed at fixing the current state of the building, minor repairs, adjustments, and improvements can be listed as examples of preservation.
- Restoration means "returning a place to a known earlier state by removing accretions or by reassembling existing elements without the introduction of new material" [20]. It is important to note that, according to the Burra Charter, "place" has a broad scope and includes natural and cultural features, such as individual buildings or groups of buildings. According to the CEN EN 15898 (Conservation of Cultural Property—Main General Terms and Definitions), restoration interventions are "actions applied to a stable or stabilized object aimed at facilitating its appreciation, understanding, and/or use, while respecting its significance and the materials and techniques used".
- Renovation focuses on building appearance and tenant facilities. Retrofitting can also be considered as renovation targeted at the improvement of energy and water efficiency—sustainable retrofitting [13].
- Remodeling establishes a new function within a building, which may imply substantial modifications [21].

Often, the information exchanged between experts does not have a clear formal meaning [22], rendering it semantically non-interoperable. A knowledge modelling approach, connected to ontologies, provides a way to include across and between domain semantics, which is necessary for establishing connections and formally defining concepts. Ontologies include the conceptualization of the intended meaning, suggested by domain experts, maintaining the usability of the concepts from the application point of view. Based on ontologies, advanced SPARQL (SPARQL Protocol and RDF Query Language) querying mechanisms can be used to retrieve data and information. Through these mechanisms, knowledge bases can be searched for a specific type of material, or even a particular element that is part of a building, (e.g., stone lintel). The retrieved information can then be used in future interventions, and as a source of knowledge for other projects.

Besides the application of querying mechanisms, ontologies can also be used for semantic reasoning to infer logical consequences from asserted heritage building information. Semantic reasoning requires a knowledge base created from an ontology (schema and data) that contains semantic rules [23,24].

Therefore, the mentioned knowledge modelling, or knowledge engineering, can assist in the creation of consultation systems that contribute greatly to the management of heritage buildings projects.

Tibaut et al., [25] researched methodologies for building ontologies and chose to implement the first four steps of the Methontology [26] (1—purpose of the ontology, 2—collection of the knowledge, 3—conceptualization, 4—reuse, 5—implementation, and 6—documentation). The present paper continues from the fifth step of the methodology and the evaluation of the ontological resources that can be reused for the heritage building domain. Since the heritage building domain can be considered as tangible cultural heritage or built cultural heritage, the ISO 21127:2014: Information and documentation—A reference ontology for the interchange of cultural heritage information, is considered the appropriate candidate ontology for reuse. However, the appropriateness for reuse has not been measured and evaluated in a formal way.

Therefore, the main research focus of this paper is a novel framework for ontology evaluation based on improved OQuaRE [27] metrics. OQuaRE is an ontology quality model based on the SQuaRE standard for software quality evaluation that proposes 14 metrics (e.g., *NOMOnto*—number of properties per class). The framework is then used for the evaluation of the cultural heritage information ontology (CIDOC CRM—ISO 12217:2014). The evaluation results will help to decide to what extent the ontology can be adopted for the HB domain. For the evaluation of the ontology, a framework with a new algorithm for the OQuaRE methodology is presented.

A discussion of results that additionally address the future development of the CIDOC CRM is presented in Section 4.

2. Related Work

In this section, the requirements for knowledge research in heritage buildings are described, including a detailed review of relevant existing knowledge resources and the identification of the foundational ontologies for reuse.

2.1. Built Cultural Heritage

Heritage buildings, as built cultural heritage, have an intrinsic value (architectural, engineering, cultural, aesthetic, etc.), and the necessary information for their proper protection is demanding both in terms of quantity and quality. The specifics and requirements of projects focused on heritage buildings give remarkable insights and inputs regarding the needed characteristics for ontology research and development. Usually, the project data are lacking or unavailable, and in most cases, not digitalized. The information contained in technical drawings is generally presented in a non-digital, less-structured way, making the process of transforming it in a machine-readable form difficult and time consuming. If contemporary technologies included in soft computing can be helpful, there is still no streamlined way to execute this function. Before interventions in a building occur, large volumes of information need to be managed and reasoned.

Due to possible similarities between heritage buildings projects, whether in terms of construction processes or common pathologies for treatments and interventions, this field of study can benefit significantly from the comparison of similar projects. Adding to the advantages that may come from better organization from data, knowledge models can assist in the identification of missing data.

As one of the objectives is to provide a uniform base for the common concepts and their detailed standardized and validated description, ontologies also encompass the machine-readability side, combining semantic meaning with fully digital prepared information [28]. Another advantage of developing dedicated ontologies is the connection of concepts, establishing links and relationships, to specific processes.

Historical data "transformed" into easily accessed information lead to linked databases where interested parties can retrieve said data. Databases can then be connected, benefiting studies, analysis, and research. With ontology engineering in the background providing and structuring the information, a knowledge base is added to the processes connected to the projects. Lytras et al. [29] propose that, due to the vastness of the heritage buildings domain, it is preferable to reuse existing reliable parts and sub-domain knowledge models to create a detailed and necessary ontology.

2.2. Non-Ontological Resources

Comprehensive knowledge-serving ontologies dedicated to heritage buildings can be found in non-ontological resources that should be researched and investigated to validate integration. Among these resources, we find glossaries, vocabularies, taxonomies, classification systems, conceptual schemas, and catalogues/databases devoted to heritage buildings, as well as data models for the industry foundation classes (IFC).

Vocabularies and glossaries are good sources of information on terminology. They often provide detailed descriptions and structured, indexed, and simplified access for fast retrieval of information. Among those dedicated to cultural heritage, mainly focused on tangible heritage as buildings, the following can be highlighted: the Building and Construction Glossary [30], the UNESCO World Heritage Center Glossary [31], the Getty Vocabularies [32], and the Forum on Information Standards in Heritage (FISH) [33].

Taxonomies, as hierarchically structured and classified information resources, can be consulted for a more effective organization of the documentation of a project. The WAND Building and Construction Management Taxonomy [34] counts 4565 terms and over 3000 synonyms of concepts, equipment, processes, and types of documents for any building and construction project. It suggests tagging and organizing project documents. It is linked to the WAND Product and Service Taxonomy (materials, equipment, and supplies).

Other resources assisting in the organization of information are classification systems. One example is the ISO 12006-2:2015 Building construction—Organization of information about construction works—Part 2: Framework for classification. The standard aids organizations to develop classification systems applied to the complete life cycle of construction works. The third part of the same standard, EN ISO 12006-3:2016 (Building construction—Organization of information about construction works—Part 3: Framework for object-oriented information, identifies a language-independent information model, which enables applications as the buildingSMART Data Dictionary (bSDD). The bSDD presents classifications and their properties, allowed values, units, and translations, working as a semantic mapping tool for construction-related terms.

Specifically focused on historic buildings, the EU-CHIC project [35], proposes a cultural heritage identity card, supporting sustainable interventions on historic sites and monuments. The suggested database schema advises eight entities: names and references, location, functional type, dating, construction, current physical condition, protection/legal status, and major risks.

As previously noted, digitalization in construction, strongly connected to the use of information models, is advantageous for the heritage domain, and the industry foundation classes (IFC) are a resource that can be used as an exchange format for BIM, specifying a conceptual data schema for the built asset industry. The data schema is specified in the ISO 16739-1:2018 industry foundation classes (IFC) for data sharing in the construction and facility management industries—Part 1: Data schema.

2.3. Core Ontologies for Reuse

Reusing existing ontologies when building a new one is still a practice not established in some domains [36]. Taking already-existing knowledge when creating new ontologies is particularly advantageous when core ontologies are reused. This step needs to be included and analyzed when ontology engineering processes are being implemented. One of the first advantages is the elimination of duplicated components, being sufficient to reference them as already existent in the reused external core ontologies.

According to [37], ontologies are categorized as foundational, core domain, and domain ontologies. Foundational ontologies (FO), also called top-level, reference, high-, upper-, or meta-level ontologies, are domain independent and contain basic and general concepts and relations that can be reused in many arbitrary domains. Basic concept in FOs can be a fundamental higher level philosophical abstraction, e.g., "endurants" or persistent objects such as a person or castle in UFO-A [38], e.g., "perdurants" or temporal events such as a medieval period or reconstruction works in UFO-B, or simpler, lower-level daily business terms, e.g., "Obligation" in FO gist [39].

When the higher- and lower-level approach is combined within a single ontology, this is regarded as hybrid ontology.

The reuse of basic, fundamental, and well-established concepts prevents the duplication of underlying concepts, already well defined by specialized communities and organizations. The result of standardization processes and agreements, these concepts are stable, and their usability and applicability are guaranteed. Several concepts can therefore be identified as core for the heritage buildings domain.

Several core ontologies were identified and considered relevant for reuse:

- The Dublin Core Metadata Element Set establishes a vocabulary for cross-domain resource description [40]. The set includes 15 terms: contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, and type. The ISO 15836-2:2019 describes classes and properties of these metadata terms.
- Temporal concepts are part of the time ontology in OWL [41]. Including 19 classes (*time:DateTimeDescription*, *time:DateTimeInterval*, *time:DayOfWeek*, *time:Duration*, *time:Interval*, etc.) and 58 properties (e.g., *time:after*, *time:before*, *time:hasDuration*, *time:weeks*), the ontology can be reused for the description of the temporal properties of heritage buildings resources.
- For the description of people, the connections between them, and what they create and do, the Friend of a Friend (FOAF) Vocabulary Specification [42] establishes concepts such as (*foaf:name*, *foaf:familyName*, etc.), linking a specific person, as an engineer or an architect, to information about the building; for example, through properties as *foaf:title* and *foaf:currentProject*.
- The ISA Programme Location Core Vocabulary [43] is used for concepts describing a place in terms of three classes: location, address, and geometry, combined with properties such as full address, post code, etc.
- Basic Geo (WGS84 lat/long) vocabulary [44] complements the ISA Programme Location Core Vocabulary for exact geometry (lat/long), e.g., wgs84_pos#Point is composed of wgs84_pos:lat and wgs84_pos:lon.
- The ISA Programme Person Core Vocabulary [45] can be reused due to the simplified description of a person, using three classes: person, code, and identifier, and properties such as family name, patronymic name, gender, birthdate, place of birth, country of birth, citizenship, residency, etc.
- The Simple Knowledge Organization System Reference (SKOS) [46] serves for semantically linking concepts between different web-based knowledge organization systems related to heritage buildings, e.g., *skos:closeMatch*, *skos:exactMatch*, *skos:prefLabel*.
- Schema.org [47], an RDF vocabulary for structured data on many domains, contains concepts that can be reused for the heritage buildings domain, namely *schema:birthDate* and *schema:deathDate* (e.g., birth and death date of a duke, which owned a castle), *schema:alternateName* (e.g., "Correspondance générale" as alternative name to "Emperor of the French"), and *schema:sameAs* (to semantically link different URI resources describing the same HB individual).

2.4. Ontologies for Built Cultural Heritage

The ISO 21127:2014—reference ontology for the interchange of cultural heritage information [48] is currently the furthest developed ontology for the integration of cultural heritage information and is intended to promote a shared understanding of cultural heritage information by providing a common and extensible semantic framework to which any cultural heritage information can be mapped. The standard was reviewed in 2020 and confirmed as current. Originally, it was developed for the international museum community (UNESCO's ICOM) by a team of multidisciplinary experts intended for the exchange of information managed by museums, libraries, and archives. Implicit and explicit concepts in cultural heritage documentation are detailed. The standard was development in cooperation with the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM). The CIDOC Conceptual Reference Model (CRM), which provided a common and extensible semantic framework for the integration of cultural heritage information, was recognized as an ISO standard in 2006.

Being a reference ontology, the CIDOC CRM is often a base of specific ontologies, such as the one developed for the Cantabria's Cultural Heritage Ontology project [49]. In this case, for the Cantabria region and its cultural heritage. Other resources populated the ontology, namely official sources from the Cantabria government. Standardized and proprietary formats, such as the Encoded Archival Description and Guide files (for encoding archival finding aids or guides to primary source material), MARC 21 (set of codes for encoding machine-readable records), Excel spreadsheets, and relational databases, were used.

In [50], heritage domain ontology engineering is applied to build the Yorùbá Cultural Heritage Ontology, devoted, in general, to African heritage museums, monuments, memorabilia, libraries, icons, and archival resources. More than 900 cultural heritage objects were associated with 237 concepts/classes, including sculptures, dialects, ceremonies, and more relevant to the heritage building domain, monuments (ancient, historical, and intentional). The ontology represents, and can be used as, a knowledge resource for digital documentation.

The Built Cultural Heritage (BCH) ontology for preventive conservation [51] merged and expanded the Geneva City Geographic Markup Language (Geneva CityGML), Monument Damage ontology (Mondis), and the CIDOC-CRM, applying it to a case related to the New Cathedral of Cuenca in Ecuador.

The Cultural Heritage Abstract Reference Model [52] is based on some concepts for the ontology of cultural heritage, with an emphasis on archaeology and anthropology. In the literature, authors neglect the latest CIDOC CRM with the argument that it imposes an over-prescriptive ISO standard while proposing their own proprietary conceptual modelling language, ConML. However, it is questionable whether such approach can guarantee the necessary conceptual stability for the HB domain.

In [53], Colucci et al., propose an Ontology-Based Method for the Generation of Parametric Structured Models for Historical Built Heritage, supported by CityGML-Geography Markup Language, the CIDOC-CRM, the IFC, and the Getty Art and Architecture Thesaurus (AAT). A partial example of this conceptual formalization is presented, with reference to the Castellina of Norcia, enhancing the comprehension of fortified structures.

Other ontological resources worthy of mention and analysis are the databases and catalogues of cultural heritage, such as Europeana.eu [54]. Divulging information on tangible, intangible, and natural heritage in a digital, organized, user-friendly, and accessible manner, in collaboration with European archives, libraries, and museums, the Europeana Data Model (EDM) is a core ontology (OWL) of this resource.

3. Evaluation Framework for the Cultural Heritage Information Ontology

The term "ontology evaluation" can be defined as the application of an ontology-evaluation methodology for an ontology. Ontology evaluation gives measures about applicability and reusability of the ontology, which ontology creators usually do not have in mind when encoding usability of the intended meaning of the ontology in question. The evaluation does not use semantic similarity approaches, which usually consider the shortest path between the two concepts and are used for the evaluation of taxonomic knowledge in large ontologies (e.g., medical ontologies). Our approach examines structural characteristics.

An ontology-evaluation methodology [55] must include quantifiable measures (metrics) for which numerical or descriptive values can be obtained or calculated from explicit ontology document content. Evaluation results can be a single value for ranking ontology [56], description of design policy, or description of its usage. Ontology-evaluation results can be stored to an ontology profile [57] or ontology metadata.

The evaluation concentrates on an automated, domain- and task-independent evaluation of the CIDOC CRM. Ontology metrics are focused on the structure and quality of the ontology. Calculation of the ontology metrics is achieved programmatically with our Java code, which can be shared on request.

In fact, the CIDOC CRM is a hybrid ontology because it models the cultural heritage domain concepts (i.e., concept E38_Image: distributions of form, tone and color that may be found on surfaces such as photos, paintings, prints and sculptures or directly on electronic media), and also includes its own version of higher-level concepts (i.e., concept

E77_Persistent_Item: this class comprises items that have a persistent identity, sometimes known as "endurants" in philosophy).

New Ontology Metrics for Interoperability

Simple ontology metrics can be obtained from ontology-authoring tools (e.g., Protégé, etc.) or programmatically (e.g., OntoMetriCalc Java API, https://sourceforge.net/projects/ontometricalc/, accessed on 10 January 2022). A collection of simple ontology metrics for the CIDOC CRM, (version 6.2, encoded in OWL, https://cidoc-crm.org/sites/default/files/cidoc_crm_v6.2-2018April.rdfs, accessed on 10 January 2022) shows that it contains 2455 axioms, 84 classes, 277 object properties, 8 data properties, and 0 individuals.

These simple metrics give only a very general quantitative perception of an ontology. In addition, there is ambiguity in naming across different ontology software. For example, number of levels, maximal depth, and maximal length may all expresses maximal depth in the ontology concepts tree (e.g., in the CIDOC CRM the concept *owl:Thing* is at level 0, the concept E1_CRM_Entity is at level 1, while the concept E34_Inscription is at the level 9).

A more complex metric system is needed for further detailed ontology evaluation and comparison.

For our research, complex metrics were created as suggested in the ontology-evaluation framework OQuaRE (http://miuras.inf.um.es/oquarewiki/, accessed on 10 January 2022). We studied the OQuaRE metrics and have developed a new evaluation framework, which we used for the evaluation of the CIDOC CRM. The framework includes our Java software based on the Apache Jena Ontology API. During the research, some errors and inconsistencies were detected in the definition and formulas of the original OQuaRE metrics, which were corrected. Corrected and improved definition and formulas for the framework, along with the results of the evaluation of the CIDOC CRM, are presented in Section 3.

Interoperability of ontologies can be regarded as an integration (reuse) of a conceptualization subset of an ontology A (external) inside an ontology B (internal). Integration is guaranteed by design because the to-be-integrated ontologies use the same mechanisms (e.g., syntax XSD, RDF, RDFS, OWL) to express axioms. The external conceptualization subset must be semantically aligned with the internal conceptualization set. Such interoperability occurs within complementary domains (e.g., wine and food) and at the level of resources (concepts and properties). External concepts are immutably contextualized within the maintained ontology, and this enables new properties relating internal (base) and external concepts. Imported ontologies (e.g., <owl:imports rdf:resource="http://xmlns.com/foaf/0.1/index.rdf"/>) are also considered as external.

For a better understanding of the quality and quantity of the interoperability between ontologies, we assume the following:

- O_b—Maintained (base) ontology (in evaluation);
- O_k—External ontology, $1 \leq k \leq K$, K is number of all external ontologies;
- N—Number of all resources (internal + external) in the O_b;
- N_e—Number of all resources from all external ontologies in the O_b;
- ENS_i—Namespace of an (external) ontology;
- n—Number of base resources from the O_b (excluding external),
- p—Percentage of base resources in O_b, $p = n/N$.

If $1 \leq i \leq K$ (K is number of all external ontologies):

- n_i—Number of resources from external ontology O_i used in the O_b;
- p_i—Percentage of resources from O_i used in the O_b, $p_i = n_i/N_i$.

N is the number of all resources in O_b:

$$N = n + \sum_{i=1}^{K} n_i, \tag{1}$$

P_i is the percentage of resources from the O_i in the O_b:

$$P_i = \frac{n_i}{N} = \frac{p_i N_i}{N}, \tag{2}$$

Based on this assumption, the following new metrics are proposed:

1. Externes (*EXOnto*)

The metric expresses the number of external ontologies used in the maintained ontology. The ontology is considered external if it contributes at least one resource from its external domain conceptualization (concepts, properties). Externality is always considered positive because the usage of external ontologies increases the value of the maintained ontology more than it costs to create (reinvent) its own conceptualization.

$$\begin{gathered} EXOnto = \textstyle\sum_{i=1}^{K} ENS_i, \\ ENS_i \neq [\ XSD,\ RDF,\ RDFS,\ OWL], \\ EXOnto \in [0, \infty] \end{gathered} \tag{3}$$

Extreme values are:

Minimum value *EXOnto* = 0 means that there are no external ontologies used in the maintained ontology.

2. Composability (*CPOnto*)

The metric measures the composability of the maintained ontology by exploiting the quality of the composure when external ontologies are used through the name-spacing mechanism. It expresses a percentage of the resources from external ontologies in the maintained ontology not considering the standard namespaces (XSD, RDF, RDFS, OWL).

CPOnto is the coverage of the maintained ontology with external ontologies:

$$\begin{gathered} CPOnto = \sum_{i=1}^{K} P_i = \tfrac{\sum_{i=1}^{n} n_i}{N} = \tfrac{N-n}{N} = 1 - p, \\ K = EXOnto \\ CPOnto \in [0, 1] \end{gathered} \tag{4}$$

Extreme values are:

- Monolithic ontology, no coverage from external ontologies: *CPOnto* = 0, if $p = 1$ or $n = N$,
- Total coverage from external ontologies: *CPOnto* = 1, if $p = 0$ or $n = 0$.

Ontology is monolithic when the composability equals to 0 (no external ontologies). If composability is high, then external ontologies are well embedded and are well contextualizing the maintained ontology.

Typically, fundamental ontologies have low composability, core ontologies ideally have medium composability, and domain ontologies ideally have higher composability.

Initially, the metric was tested with the ontology of bioscientific data analysis and data management (EDAM, version 1.20, http://edamontology.org/, accessed on 10 January 2022), as well as the Description of a Project (http://usefulinc.com/ns/doap, accessed on 10 January 2022). EDAM (an ontology for bioinformatics and computational biology) uses the following external namespaces:

- *dc = http://purl.org/dc/elements/1.1/;*
- *edam = http://purl.obolibrary.org/obo/edam#;*
- *foaf = http://xmlns.com/foaf/0.1/;*
- *oboInOwl = http://www.geneontology.org/formats/oboInOwl;*
- *oboOther = http://purl.obolibrary.org/obo/;*
- *doap = http://usefulinc.com/ns/doap.*

The *CPOnto* value for the EDAM ontology is 0.38. It contains 42,628 internal resources in axioms and 25,946 external resources in axioms. The DOAP vocabulary uses the following external namespaces:

- xmlns:vs = http://www.w3.org/2003/06/sw-vocab-status/ns#;
- xmlns:foaf = http://xmlns.com/foaf/0.1/#;
- xmlns:dc = http://purl.org/dc/elements/1.1/#.

It imports FOAF (*owl:imports rdf:resource = "http://xmlns.com/foaf/0.1/index.rdf"*), which adds all FOAF concepts to the DOAP during evaluation but then only a small number of the concepts are used. The *CPOnto* value for the DOAP ontology is 0.06. It contains 1610 internal resources in axioms and 103 external resources in axioms. The composability metrics for EDAM and DOAP are indicated in Table 1.

Table 1. Composability metrics for EDAM and DOAP.

Ontology	Internal Resources	External Resources	*CPOnto*
EDAM	42,628	1610	0.38
DOAP	25,946	103	0.06

3. Aggregability (*AGOnto*)

The metric relates to the aggregation size of the sets of the external ontologies in the maintained ontology. It calculates the relative size (percentage) of the conceptualization for the maximal and minimal conceptualization set according to the total size of all external resources in the maintained ontology. The size of the maximal set is then divided by the size of the minimal set to obtain the multiplier.

$$\begin{gathered} AGOnto = \frac{Max\left(\frac{n_1}{N_e},\ldots,\frac{n_K}{N_e}\right)}{Min\left(\frac{n_1}{N_e},\ldots,\frac{n_K}{N_e}\right)}, \\ N_e = \sum_{i=1}^{K} n_i \\ K = EXOnto \\ AGOnto \in [0,\infty] \end{gathered} \tag{5}$$

Extreme values are:

Minimum *AGOnto* = 1 means that all external ontology sets are equally sized or that there is only one external set in the maintained ontology. Maximum value is achieved when there is a great difference in the sizes of the two extreme external conceptualization sets. Lower value of the *AGOnto* means that external conceptualization sets are of comparable importance, while a bigger value demonstrates a large dispersion of external ontologies contributions used in the maintained ontology.

The contribution from external ontologies is more important if the aggregability value *AGOnto* is lower, meaning that they contribute more equally to the maintained ontology.

The conclusion is that a low dispersion of external ontologies, and thus external conceptualization sets of comparable importance, is more desirable. However, this metric should be analyzed in combination with the composability metric (*CPOnto*) to draw conclusions about the interoperability, always taking the mapping scoring table in consideration.

The interoperability metrics discourage a high number of external ontologies; in practical terms, it is advisable to reduce this number because it is easier to maintain the quality of the ontology (e.g., keeping up with possible high number of updates of the used external ontologies is expensive).

The *AGOnto* value for the EDAM ontology is 13,425, which means that there is an extreme difference in the size of external conceptualization sets. The external ontology *oboInOwl = http://www.geneontology.org/formats/oboInOw* contributes 52%, (13,425 used resources), while the *doap = http://usefulinc.com/ns/doap* only contributes 0.00385% (one used resource).

Table 2 presents the definitions and formulas for the framework and the results of the evaluation of the CIDOC CRM. The scoring tables for ontology-evaluation metrics, as defined in the OQuaRE framework, are indicated in Table 3, and the calculated ontology metrics based on the corrected OQuaRE framework are presented in Table 4.

Table 2. Complex ontology metrics obtained from the improved OQuaRE framework.

Metric (Definition and Formula)	Score Table	Improvement
LCOMOnto2 (Lack of cohesion in methods): Semantic and conceptual relatedness of classes. It can be used to measure the separation of responsibilities and independence of components of ontologies. Formula: *LCOMOnto2* = *∑PathLength(Cthing,LeafCi)/∑PathLeafCj*, where *PathLength* is the function that calculates length (number of edges) between the i-th leaf concept *LeafCi* and the *Cthing* (*owl:Thing*) and *PathLeafCj* is the j-th path between *Cthing* and a leaf.	MT1	Improved definition and corrected formula because the previous formula was not correct in saying that the total length of all paths between Thing and leaves must be divided by "total number of paths in the ontology". In that case, the result would always be lower than 1.
WMCOnto2 (Weight method per class): Arithmetic mean number of path (number of edges between concepts) from Thing (*owl:Thing*) to a leaf class per leaf concepts. Formula: *WMCOnto2* = *∑PathLeafCi/∑LeafCj*, where *PathLeafCi* is the i-th path between *Cthing* (*owl:Thing*) and a leaf and *LeafCj* is the j-th leaf.	MT2	Improved definition and improved formula.
DITOnto (Depth of subsumption hierarchy): Measures ontological depth, which is the length (number of links between concepts) of the longest path from Thing (*owl:Thing*) to a leaf concept. Formula: *DITOnto* = *Max(PathLength(CThing, LeafCi))*, where *PathLength* is the function that calculates length between the i-th leaf concept *LeafCi* and the *CThing* (*owl:Thing*). The maximum path length (function *Max*) is then selected as result.	MT2	Improved definition and improved formula.
NACOnto (Number of ancestor concepts): Arithmetic mean number of direct ancestor concepts per leaf concept. Formula: *NACOnto* = *∑AncLeafCi/∑LeafCj*, where *AncLeafCi* is the i-th direct ancestor of a leaf and *LeafCj* is j-th leaf concept.	MT2	Improved definition and improved formula.
NOCOnto (Number of children concepts): Number of the direct subconcepts divided by the number of concepts minus the number of leaf concepts. Formula: *NOCOnto* = *∑Ci∑SubCj/(∑Ci − ∑LeafCk)*, where *Ci* is the i-th concept and *SubCj* is its j-th direct subclass and *LeafCk* is k-th leaf concept.	MT2	Improved definition and improved formula.

Table 2. *Cont.*

Metric (Definition and Formula)	Score Table	Improvement
CBOnto (Coupling between objects): Number of direct ancestors of all concepts divided by the number of concepts not counting concepts with *owl:Thing* as direct ancestor. Formula: $CBOnto = \sum Ci \sum AncCj / (\sum Ci - \sum CTk)$, where *Ci* is the i-th concept, *AncCj* is its j-th direct ancestor, and *CTk* is k-th concept with *owl:Thing* as direct parent.	MT3	Improved definition and improved formula.
RFCOnto (Response for a concept): Number of direct properties (*owl:ObjectProperty*, *owl:DatatypeProperty*) and direct ancestor concepts that can be directly accessed from a concept divided by the number of all concepts. Formula: $RFCOnto = (\sum Ci \sum ProCj + \sum Ci \sum AncCk) / \sum Ci$, where *Ci* is the i-th concept and *ProCj* is its j-th property and *AncCk* is its k-th ancestor.	MT3	Improved definition and formula. Our definition includes direct properties and direct ancestors (super-classes) while the original definition is not clear. Without that restriction, the RFCOnto value in our case would be 49.86.
NOMOnto (Number of properties): Arithmetic mean number of properties (*owl:ObjectProperty*, *owl:DatatypeProperty*) per concept. Formula: $NOMOnto = \sum Ci \sum ProCj / \sum Ci$, where *Ci* is the i-th concept and *ProCj* is its j-th property.	MT1	Improved definition and formula. The previous definition and formula were not correct because they allowed the number of all properties including those of super-classes.
RROnto (Relationship richness): Number of subconcepts (*rdfs:subClassOf*) divided by the sum of subconcepts (*rdfs:subClassOf*) plus object and data properties (*owl:ObjectProperty*, *owl:DatatypeProperty*) of the concepts. Formula: $RROnto = \sum Ci \sum SubCj / (\sum Ci \sum SubCj + \sum Ci \sum ProCk)$, where *Ci* is the i-th concept and *SubCj* is its j-th subconcept and *ProCk* is its k-th property.	MT4	Improved definition and formula.
PROnto (Properties Richness): Sum of usages of object and data properties in axioms for concepts (also including *owl:intersectionOf*, *owl:unionOf*) and individuals divided by the sum of all direct subconcepts (*rdfs:subClassOf*) plus number of properties defined as *owl:ObjProperty* and *owl:DataProperty* in the ontology. Formula: $PROnto = (\sum Ci \sum ProCj + \sum Ij \sum ProCk) / (\sum Ci \sum SubCl + \sum Ci \sum ProCl)$, where *Ci* is the i-th concept and *ProCj* is its j-th property, *Ij* is the j-th individual and *ProcCk* is its k-th property, *SubCl* is its l-th subconcept, and *ProcCl* is the l-th object or data property defined for the concept *Ci*.	MT4	Improved definition and formula because the previous definition did not specify exactly the usage of object and data properties and, therefore, the formula was unclear. The enhanced formula explicitly includes the usage of properties inside axioms for *owl:Class* constraints (*owl:intersectionOf*, *owl:unionOf*) and *owl:Individuals*. An ontology with only object and data properties defined but never used, has this metric evaluated to 0.

Table 2. *Cont.*

Metric (Definition and Formula)	Score Table	Improvement
AROnto (Attribute richness): Number of property restrictions (*owl:Restrictions* (*owl:someValuesFrom*, *owl:allValuesFrom*, *owl:hasValue*, *owl:minCardinality*, *owl:maxCardinality*)) nested inside of *rdfs:subClassOf* per concept in the ontology. Formula: $AROnto = \sum Ci \sum RestCj / \sum Ci$, where Ci is the i-th concept and $RestCj$ is its j-th restriction.	MT4	Improved definition and formula.
INROnto (Relationships per concept): Arithmetic mean number of subconcepts (*rdfs:subClassOf*) per concept. Formula: $INROnto = \sum Ci \sum SubCj / \sum Ci$, where Ci is the i-th concept and $SubCj$ is its j-th subconcept. Remark: if the value is greater than 1, multiple inheritance is contained in the ontology, which means that the ontology in place is not normalized.	MT4	Improved definition and formula.
CROnto (Concept richness): Arithmetic mean number of direct individuals per concept (excluding individuals of its subconcepts) Formula: $CROnto = \sum Ci \sum IndCj / \sum Cj$, where Ci is the i-th concept and $IndCi$ is its j-th individual.	MT4	Improved definition and formula.
ANOnto (Annotation richness): Arithmetic mean number of annotation properties (existing in OWL: *owl:versionInfo*, *rdfs:comment*, *rdfs:label*, *rdfs:seeAlso*, *rdfs:isDefinedBy*) per concept (*owl:Class*). Formula: $ANOnto = \sum Ci \sum ApCj / \sum Ci$, where Ci is the i-th concept and $ApCj$ is its j-th annotation property.	MT4	Improved definition and formula.
TMOnto2 (Tangledness): Mean number of direct ancestors (super-classes) of concepts with more than 1 direct ancestor (multiple inheritance). Formula: $TMOnto2 = \sum Ci \sum AncCj / \sum Ci$, where Ci is the i-th concept with more than one direct ancestor and $AncCj$ is its j-th direct ancestor.	MT2	Improved definition and formula. Based on the formula, the value for *TMOnto2* cannot be lower than 2, which contradicts recent published work showing an example with $TMOnto2 < 2$ [39].
EXOnto (Externes): The metric measures the number of external namespaces that the maintained ontology consumes. The namespace is considered external if any of its resources are used at least one time in the maintained ontology. Standard namespaces (XSD, RDF, RDFS, OWL) and base namespace are not counted. Formula: $EXOnto = \sum ENSi$, where $ENSi$ is the i-th external namespace (ontology).	-	New metric.

Table 2. *Cont.*

Metric (Definition and Formula)	Score Table	Improvement
CPOnto (Composability): The metric describes the composure of the ontology on the scale from a monolithic, self-sufficient ontology to a highly composed and interconnected ontology. It is calculated as the usage of resources in axioms from external namespaces divided by the usage of all resources (base namespace and external namespaces). Formula: *CPOnto* = ∑*ResEi*/(∑*ResEi* + ∑*ResBj*), where *ResEi* is the i-th external resource and *ResBj* is the j-th resource from the base namespace.	MT4	New metric.
AGOnto (Aggregability): The metric relates to the aggregation size of the sets of the external ontologies. It calculates the percentage of the used resources for the maximal and minimal sets according to the total size of all external sets in the maintained ontology. The maximal percentage is then divided by the size of the minimal percentage to get the multiplier. Formula: *AGOnto* = *Max*(*ResEi*/*Ne*, . . . , *ResEn*/*Ne*)/*Min*(*ResEi*/*Ne*, . . . , *ResEn*/*Ne*) where *ResEi* is the size of the i-th external namespace (i between 1 and number of external namespaces), Ne is the sum of used external resources in the maintained ontology.	MT5	New metric.

Table 3. Scoring mapping table for ontology-evaluation metrics as defined in the OQuaRE framework.

Mapping Table						
Metric score		**1**	**2**	**3**	**4**	**5**
Metric value	MT1	>8	(6–8]	(4–6]	(2–4]	≤2
	MT2	>8	(6–8]	(4–6]	(2–4]	[1–2]
	MT3	>12	(8–12]	(6–8]	(3–6]	[1–3]
	MT4	[0–20]%	(20–40]%	(40–60]%	(60–80]%	>80%
	MT5	>99	(9–99]	(6–9]	(3–6]	[1–3]

Table 4. Calculated ontology metrics based on the corrected OQuaRE framework.

Metric	Value	Score Table
LCOMOnto2	782/114 = 6.86	MT1
WMCOnto2	114/49 = 2.32	MT2
DITOnto	10	MT2
NACOnto	56/49 = 1.14	MT2
NOCOnto	98/(85–49) = 2.72	MT2
CBOnto	99/(85–2) = 1.19	MT3
RFCOnto	(287 + 99)/(85) = 4.54	MT3
NOMOnto	287/85 = 3.38	MT1
RROnto	98/(98 + 287) = 0.25 (25%)	MT4
PROnto	0/(98 + 287) = 0%	MT4
AROnto	0	MT4
INROnto	98/85 = 1.15 (115%)	MT4
CROnto	0	MT4
ANOnto	640/85 = 7.53	MT4
TMOnto2	30/15 = 2	MT2

The calculated OQuaRE metrics (Table 4) are real values and are, firstly, mapped to 1—"Not Acceptable", 2—"Not Acceptable—Improvement Required", 3—"Minimally

Acceptable", 4—"Acceptable", and 5—"Exceeds Requirements", according to the mapping table (see Table 3); secondly, they are associated with the ontology quality characteristics and sub-characteristics (Table 5).

Table 5. Evaluation results for the CIDOC CRM obtained from the corrected OQuaRE framework.

Characteristics/Sub-Characteristics	Metrics	Calculated Value	Score
1. Structural			3.5
Formal relations support	*RROnto*	0.25	2
Cohesion	*LCOMOnto*	6.86	2
Tangledness	*TMOnto2*	2	5
Redundancy	*ANOnto*	7.53	5
2. Functional adequacy			2.92
Controlled vocabulary	*ANOnto*	7.53	5
Schema and value reconciliation	*RROnto*	0.25	2
	AROnto	0	1
Consistent search and query	*ANOnto*	7.53	5
	RROnto	0.25	2
	AROnto	0	1
	INROnto	1.15	5
Knowledge acquisition	*ANOnto*	7.53	5
	RROnto	0.25	2
	NOMOnto	3.38	4
Similarity	*RROnto*	0.25	2
	AROnto	0	1
Indexing and linking	*RROnto*	0.25	2
	AROnto	0	1
	INROnto	1.15	5
Results representation	*AROnto*	0	1
	CROnto	0	1
Guidance and decision trees	*INROnto*	1.15	5
	AROnto	0	1
Knowledge use	*ANOnto*	7.53	5
	AROnto	0	1
	INROnto	1.15	5
	NOMOnto	3.38	4
	LCOMOnto	6.86	2
3. Transferability			3.5
Adaptability	*WMCOnto2*	2.32	1
	DITOnto	10	1
	RFCOnto	4.54	4
	CBOnto	1.25	5

Table 5. *Cont.*

Characteristics/Sub-Characteristics	Metrics	Calculated Value	Score
4. Reliability			2.6
Recoverability	*WMCOnto2*	2.32	1
	DITOnto	10	1
	NOMOnto	3.38	4
	LCOMOnto	6.86	2
Availability	*LCOMOnto*	6.86	2
5. Compatibility			3.25
Replaceability	*WMCOnto2*	2.32	1
	DITOnto	10	1
	NOCOnto	2.72	4
	NOMOnto	3.38	4
6. Interoperability			
Composability	*CPOnto*	0	/
Aggregability	*AGOnto*	0	/
7. Maintainability			3.56
Modularity	*WMCOnto2*	2.32	1
	CBOnto	1.25	5
Reusability	*WMCOnto2*	2.32	1
	DITOnto	10	1
	NOCOnto	2.72	4
	RFCOnto	4.54	4
	NOMOnto	3.38	4
	CBOnto	1.25	5
Analyzability	*WMCOnto2*	2.32	1
	DITOnto	10	1
	LCOMOnto	6.86	2
	RFCOnto	4.54	4
	NOMOnto	3.38	4
	CBOnto	1.25	5
Changeability	*WMCOnto2*	2.32	1
	DITOnto	10	1
	LCOMOnto	6.86	2
	RFCOnto	4.54	4
	NOMOnto	3.38	4
	CBOnto	1.25	5
	NOCOnto	2.72	4
Modification stability	*WMCOnto2*	2.32	1
	CBOnto	1.25	5
	LCOMOnto	6.86	2
	RFCOnto	4.54	4
	NOCOnto	2.72	4
Testability	*WMCOnto2*	2.32	1
	DITOnto	10	1
	LCOMOnto	6.86	2
	RFCOnto	4.54	4
	NOMOnto	3.38	4
	CBOnto	1.25	5
8. Operability			3.83
	WMCOnto2	2.32	1
	LCOMOnto	6.86	2
	RFCOnto	4.54	4
	NOMOnto	3.38	4
	CBOnto	1.25	5
	NOCOnto	2.72	4
Total ontology quality (characteristics' average)			3.31

Quality characteristics are the quality of ontology structure, functional adequacy, transferability, reliability, compatibility, maintainability, and operability. Sub-characteristics for the characteristics' maintainability are modularity, reusability, analyzability, changeability, modification stability, and testability. An individual sub-characteristic is a combination of different metrics; for example, maintainability is determined by the *WMCOnto2* (score 1) and the *CBOnto* (score 5), which, together with other sub-characteristics-metrics, contribute to an average score for maintainability of three.

The final evaluation is presented in Table 5 and summarized in Figure 1.

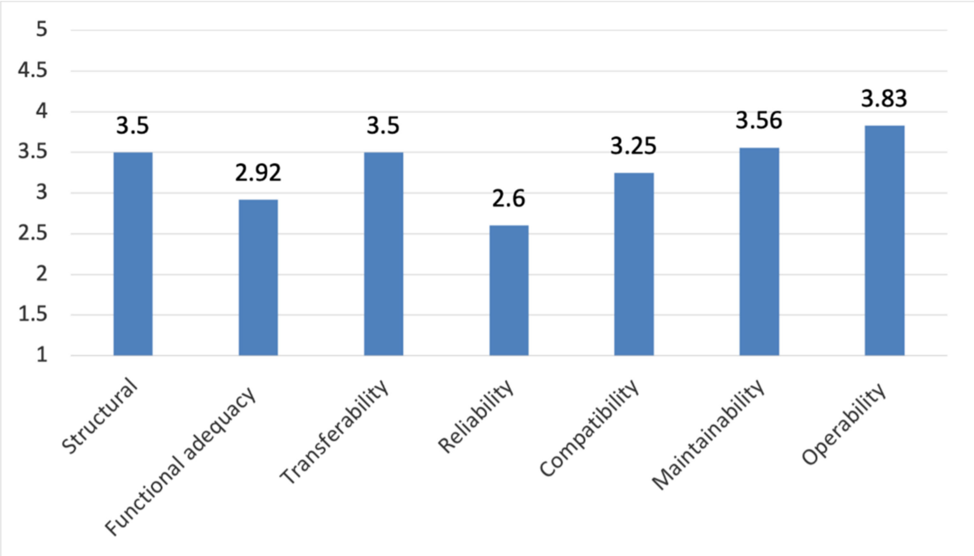

Figure 1. Quality of the CIDOC CRM according to the original OQuaRE evaluation framework.

In comparison with other published OQuaRE evaluations results, such as [27], 11 ontologies for unit measurement were compared, and the lowest total ontology quality average was 2.20. The total quality average of 3.31 (out of 5) for the CIDOC CRM indicates that it could be improved. Recommendations are described in the Discussion and Conclusions section.

Since the original CIDOC CRM does not include external ontologies, the result of the evaluation with the three new metrics (*EXOnto*, *CPOnto*, *AGOnto*) integrated into the OQUARE framework yielded a value of 2.89. To prove the usefulness of the extended OQUARE framework, we cloned the CIDOC CRM and modified it with the reuse of the core ontologies (time, foaf, schema) from the section “Core ontologies for reuse”. The modification included the replacement of the concepts “E2_Temporal_Entity”, “E52_Time-Span”, “E74_Group”, “E40_Legal_Body”, and the property “P98i_was_born”:

http://www.cidoc-crm.org/cidoc-crm/E2_Temporal_Entity →
http://www.w3.org/2006/time#TemporalEntity
http://www.cidoc-crm.org/cidoc-crm/E52_Time-Span →
http://www.w3.org/2006/time#DateTimeInterval
http://www.cidoc-crm.org/cidoc-crm/E74_Group →
http://xmlns.com/foaf/0.1/Group
http://www.cidoc-crm.org/cidoc-crm/E40_Legal_Body →
http://xmlns.com/foaf/0.1/Organization
http://www.cidoc-crm.org/cidoc-crm/P98i_was_born →
http://schema.org/birthDate

The evaluation of the cloned and modified CIDOC CRM presented in Figure 2 resulted in a final value of 3.15. The integration of additional external ontologies to the CIDOC CRM can increase the result for the interoperability sub-characteristic, improving the final evaluation.

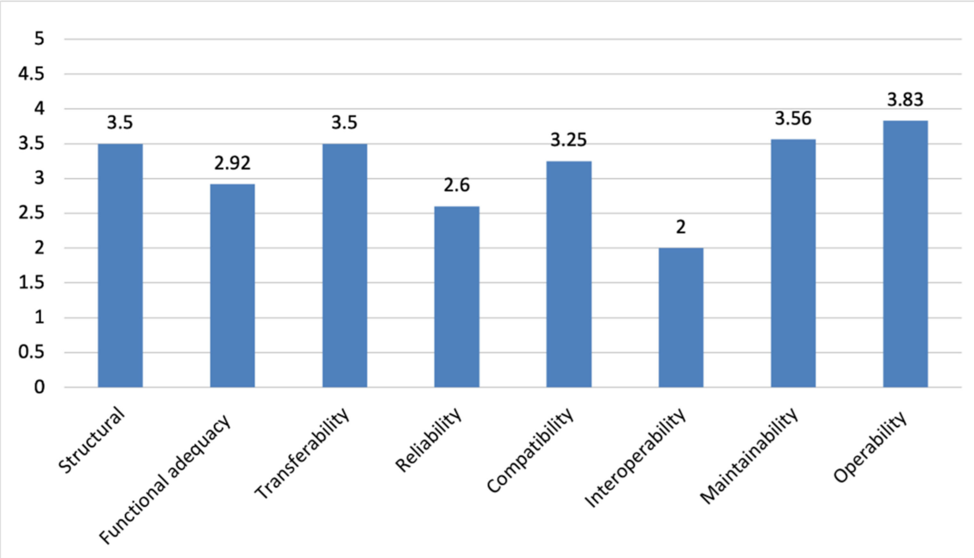

Figure 2. Quality of the CIDOC CRM according to the improved evaluation with application of new metrics for interoperability.

Figure 2 presents the evaluation results of the CIDOC CRM with the application of the new metrics for interoperability.

4. Discussion and Conclusions

The CIDOC CRM is a hybrid ontology (between foundational and domain) mainly focused on museological concepts such as paintings, archaeological literature, cultural phenomena cultural manifestations, and built heritage. It is not committed to reusing well-known core ontologies, e.g., for locations, time, persons.

Our first recommendation for the CIDOC CRM is to reuse the high-level semantics (concepts) from existing and commonly used foundational ontologies. This would shift the CIDOC CRM's generalization level towards the core level and better prepare it for sharing instances of the ontology. The following existing core ontologies semantically intersects with the CIDOC CRM:

- OWL-Time Ontology, because the *crm:E2_Temporal_Entity* is the same as *time:TemporalEntity* (https://www.w3.org/2006/time#TemporalEntity). To support this, the largest repository with 681 biomedical ontologies (https://bioportal.bioontology.org, accessed on 10 January 2022) shows that, out of 17 ontologies that define the concept TemporalEntity, 14 reuse the concept *time:TemporalEntity*.
- Location Core Vocabulary, because the *crm:E53_Place*, which is not complete and clear (see section Ontology Pitfalls), could be the same as locn:location from the Location Core Vocabulary, which actually fully reuses *dcterms:Location* (http://purl.org/dc/terms/Location). To support this, the Bioportal ontologies show that, out of 61 ontologies that define the concept Location, 8 reuse the concept *dcterms:Location*. The Basic Geo (WGS84 lat/long) vocabulary would complement Location Core Vocabulary with GPS positions (latitude and longitude).
- Person Core Vocabulary, because *crm:E21_Person* (comprises real persons who live or are assumed to have lived) is the same as *person:Person* (http://www.w3.org/ns/person) from the Person Core Vocabulary, which is actually a subclass of *foaf:Person* and *schema:Person* with the restriction that it only covers an individual person who may be dead or alive, but not imaginary.
- SKOS (Simple Knowledge Organization System Reference) provides several properties that map concepts between different concept schemes, and it could link the similar concepts from the CIDOC CRM to other similar ontologies in the Semantic Web (*skos:broadMatch*,

skos:narrowMatch, skos:relatedMatch, skos:exactMatch and skos:closeMatch), e.g., *crm:E21_Person skos:exactMatch person:Person*.
- Schema.org can be used to semantically enrich the concepts *crm:E67_Birth* and *crm:E69_Death* with the properties *schema:birthDate* and *schema:deathDate*. Another example is to link the concepts *schema:CreativeWork* and *crm:E65_Creation*.

The proposed modifications to the cloned CIDOC CRM ontology would increase the score for the interoperability characteristic if combined with a high value for Composability (*CPOnto*) while maintaining a low value for the Aggregability (*AGOnto*).

Building on the notion that information not shared is information lost, the new proposed metrics Externes (*EXOnto*), Composability (*CPOnto*), and Aggregability (*AGOnto*) stimulate interoperability and guarantee completeness and sustainable ontology maintenance.

The vastness of the heritage domain, and in particular of built heritage, coupled with the growing necessity for a reliable domain ontology and the existence of several core ontologies that can be reused, justified the inclusion of the Interoperability characteristic (metric) in the evaluation of the cultural heritage information ontology.

For the calculation of the OQuaRE complex metrics, a new software framework for ontology evaluation was developed. The software uses Apache Jena Ontology API (http://jena.apache.org/, accessed on 10 January 2022).

All OQuaRE quality metrics for the CIDOC CRM are ranked between 2.6 and 3.83, and the total quality average is 3.31, which ranks it as a minimally acceptable ontology according to the established scoring table.

The lowest metric scores are for Reliability (2.6) and Functional Adequacy (2.5). Reliability is defined as "capability of an ontology to maintain its level of performance under stated conditions for a given period of time" and is made up of *WMCOnto2*, *DITOnto*, *NOMOnto*, and *LCOMOnto*. *DITOnto* (depth of subsumption hierarchy) expresses the length of the largest path from *owl:Thing* to a leaf class. As already stated, the CIDOC CRM maintains its own high-level and low-level domain concepts that significantly contribute to the low *DITOnto* score of 1 because of the 10 levels of subclass hierarchy (including *owl:Thing*).

LCOMOnto (lack of cohesion in methods) expresses the separation of responsibilities and independence of components of ontologies and is measured as the total length of all paths from leaves to *owl:Thing* divided by number of paths. *LCOMOnto* has a value of 6.86. Simulation shows that separation of the two hierarchy levels containing high-level concepts ('*CRM Entity*', *Dimension*, '*Persistent Item*', *Place*, '*Spacetime Volume*', '*Temporal Entity*', *Time-Span*) would increase Reliability to 3.8 and total ontology quality to 3.73. Therefore, the recommendation is to flatten the ontology.

The Structural characteristic is influenced by the Tangledness sub-characteristic calculated by *TMOnto2* (2). An increase in the value has a negative effect on the total evaluation. For the *TMOnto2*, the formula was improved, and it shows that its value cannot be lower than 2, which contradicts published work showing an example with *TMOnto2* < 2 [57].

Functional adequacy characteristic is "the capability of the ontologies to provide concrete functions" to applications that consume relations and properties from an ontology. The score (2.92) is low because of the low scores for *RROnto* (2), *AROnto* (1), and *CROnto* (1). *RROnto* (relationship richness) is calculated as number of subconcepts (*rdfs:subClassOf*) divided by the sum of subconcepts (*rdfs:subClassOf*) plus object and data properties (*owl:ObjectProperty*, *owl:DatatypeProperty*) of the concepts. The CIDOC CRM has 176 *subclassOf* relationships usages and 286 object and data properties. As an example, Figure 3 explains that the concept *E81_Transformation* contributes two relationships (*rdfs:subClassOf*) and four object properties (2 *rdfs:domain* and 2 *rdfs:range*).

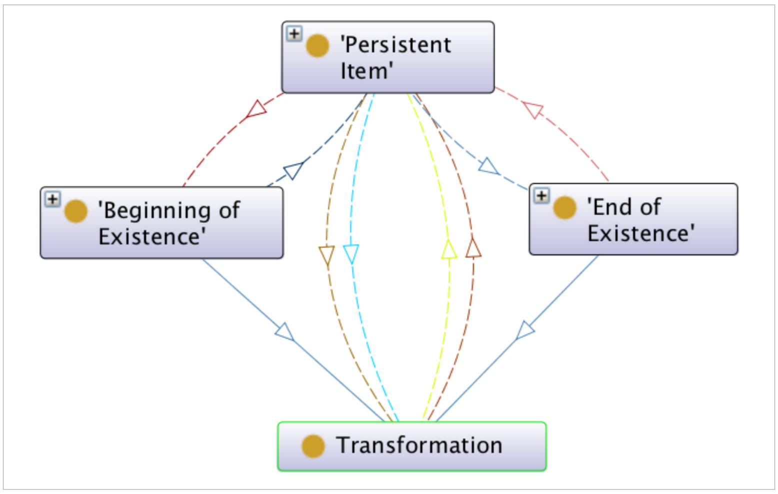

Figure 3. Concept transformation with relationships and properties.

The great number of properties creates an imbalance with the number of relationships, which scores very low for *RROnto*.

CROnto measures the mean number of individuals per class, but CIDOC CRM has no individuals; therefore, the score is 0.

AROnto measures the number of restrictions of the ontology per class. Since it is not exactly clear, we assume that restrictions in this context are *owl:Restriction* predicates in ontology. They are not defined in the CIDOC CRM; therefore, the score is 0.

Structural quality of the CIDOC CRM is low (3.25), although the metric *ANOnto* is high because it counts occurrences for *rdfs:comment* and *rdfs:label* in seven different languages. Moreover, in the CIDOC CRM, there are 15 concepts with exactly two direct superclasses. As a rule, this should be avoided because it influences structural quality through the metric *TMOnto* (Tangledness). Concepts should ideally only have one direct superclass.

To obtain additional insight into the quality of the CIDOC CRM (version 6.2 draft), the OOPS!—Ontology Pitfall Scanner [58], was used to detect pitfalls in the ontology. To each pitfall, an importance level (critical, important, and minor) is assigned:

Critical: It is crucial to correct the pitfall. Otherwise, it could affect the ontology consistency, reasoning, applicability, etc.

Important: Though not critical for ontology function, it is important to correct this type of pitfall.

Minor: It is not really a problem but correcting it will make the ontology more effective.

For the evaluation, the OOPS! Restful Web Service [59] was used with the adapted Java code from the DrOntoAPI [60] to automatically detect classification criteria pitfalls as defined in various scientific papers (such as COMPLETENESS, CORRECTNESS, ADAPTABILITY, CLARITY, etc.).

The highest importance level of pitfall found in the CIDOC CRM was "important", as shown in Table 6.

The OOPS!—Ontology Pitfall Scanner detected problems related to CLARITY and COMPLETENESS. The problems are closely connected to the weaknesses identified with OQuaRE. For example, P08: Missing annotations as reported by OOPS! reported 134 properties with a missing *rdfs:comment* predicate. The missing predicate demonstrates a lack of time (or discipline) by the ontology designer. Added properties would also increase the score for *ANOnto*. As OWL uses the open-world assumption, descriptions of classes and properties should be 'closed off' where appropriate.

Table 6. Ontology pitfalls for the CIDOC CRM.

P07 [1 time]: Merging different concepts in the same class (Minor). This pitfall affects: CLARITY • A class whose name refers to two or more different concepts is created. • This pitfall affects the element: http://www.cidoc-crm.org/cidoc-crm/E29_Design_or_Procedure
P08 [135 times]: Missing annotations (Minor). This pitfall affects: CLARITY, COMPLETENESS • This pitfall consists of creating an ontology element and failing to provide human readable annotations attached to it. Consequently, ontology elements lack annotation properties that label them (e.g., *rdfs:label*, *lemon:LexicalEntry*, *skos:prefLabel* or *skos:altLabel*) or define them (e.g., *rdfs:comment* or *dc:description*). • Example: http://www.cidoc-crm.org/cidoc-crm/P70i_is_documented_in, http://www.cidoc-crm.org/cidoc-crm/E93_Place doesn't have *rdfs:comment* defined.
P10: Missing disjointness (Important). This pitfall affects: COMPLETENESS • The ontology lacks disjoint axioms between classes or between properties that should be defined as disjointed. • Example for the http://www.cidoc-crm.org/cidoc-crm/E2_Temporal_Entity is textual (*rdfs:comment*), stating that "this class is disjoint from *E77_Persistent Item*". Instead, *owl:differentFrom* should be used.
P13 [275 times]: Inverse relationships not explicitly declared (Minor). This pitfall affects: COMPLETENESS • This pitfall appears when any relationship (except for those that are defined as symmetric properties using *owl:SymmetricProperty*) does not have an inverse relationship (*owl:inverseOf*) defined within the ontology. • Example: http://www.cidoc-crm.org/cidoc-crm/P94i_was_created_by and should have an inverse relationship, e.g., http://www.cidoc-crm.org/cidoc-crm/P94_has_created
P30 [2 times]: Equivalent classes not explicitly declared (Important). This pitfall affects: COMPLETENESS • This pitfall consists of missing the definition of equivalent classes (*owl:equivalentClass*) in the case of duplicated concepts. When an ontology reuses terms from other ontologies, classes that have the same meaning should be defined as equivalent to benefit the interoperability between both ontologies. • Example: http://www.cidoc-crm.org/cidoc-crm/E69_Death, http://www.cidoc-crm.org/cidoc-crm/E6_Destruction

The purpose of the development of knowledge model (ontology) for heritage buildings projects is its integration into semantic applications. Such applications (i.e., semantic web portal) enable simple, quick, and smart access to construction project information through a user's custom semantic queries. The knowledge model is a prerequisite for the automation of semantic reasoning through semantic links established inside the knowledge model.

The ontology-evaluation solution presented in this paper is based on the aggregated average value of calculated ontology metrics that are general and can be used for any ontology. New research could further investigate it in two different directions:

- To extend the interoperability sub-characteristics with additional simple metrics.
- Research into ontology evaluation better tailored to different ontology types (foundational, core, domain, hybrid).

Using the evaluation framework presented in this paper, future research will be focused on building an interlinked, reusable, and interoperable heritage buildings core ontology (HBCO) that can be used for the management of these assets, including interventions such as the preservation, restoration, and renovation of heritage buildings. Based on the evaluation of the CIDOC CRM, the derived ontology will result in a better organization of knowledge, facilitating sharing and reuse among all personnel involved in heritage building projects.

Author Contributions: Conceptualization, A.T. and S.G.d.O.; methodology, A.T. and S.G.d.O.; software, A.T.; validation, A.T. and S.G.d.O.; formal analysis, A.T. and S.G.d.O.; investigation, A.T. and S.G.d.O.; resources, A.T. and S.G.d.O.; data curation, A.T.; writing—original draft preparation, A.T. and S.G.d.O.; writing—review and editing, A.T. and S.G.d.O.; visualization, A.T. and S.G.d.O.; supervision, A.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Turk, Ž. Construction informatics: Definition and ontology. *Adv. Eng. Inform.* **2006**, *20*, 187–199. [CrossRef]
2. Hartmann, T. Advances in architectural, engineering and construction informatics. *Adv. Eng. Inform.* **2012**, *26*, 654–655. [CrossRef]
3. Zhong, B.; Ding, L.; Love, P.E.; Luo, H. An ontological approach for technical plan definition and verification in construction. *Autom. Constr.* **2015**, *55*, 47–57. [CrossRef]
4. Liu, H.; Lu, M.; Al-Hussein, M. Ontology-based semantic approach for construction-oriented quantity take-off from BIM models in the light-frame building industry. *Adv. Eng. Inform.* **2016**, *30*, 190–207. [CrossRef]
5. Niu, J.; Issa, R.R. Developing taxonomy for the domain ontology of construction contractual semantics: A case study on the AIA A201 document. *Adv. Eng. Inform.* **2015**, *29*, 472–482. [CrossRef]
6. Lu, Y.; Li, Q.; Zhou, Z.; Deng, Y. Ontology-based knowledge modeling for automated construction safety checking. *Saf. Sci.* **2015**, *79*, 11–18. [CrossRef]
7. Lee, Y.-C.; Eastman, C.M.; Solihin, W. An ontology-based approach for developing data exchange requirements and model views of building information modeling. *Adv. Eng. Inform.* **2016**, *30*, 354–367. [CrossRef]
8. Pauwels, P.; Törmä, S.; Beetz, J.; Weise, M.; Liebich, T. Linked Data in Architecture and Construction. *Autom. Constr.* **2015**, *57*, 175–177. [CrossRef]
9. Volk, R.; Stengel, J.; Schultmann, F. Building Information Modeling (BIM) for existing buildings—Literature review and future needs. *Autom. Constr.* **2014**, *38*, 109–127. [CrossRef]
10. Solihin, W.; Eastman, C.; Lee, Y.C. A framework for fully integrated building information models in a federated environment. *Adv. Eng. Inform.* **2016**, *30*, 168–189. [CrossRef]
11. Pauwels, P.; Terkaj, W. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Autom. Constr.* **2016**, *63*, 100–133. [CrossRef]
12. Khaddaj, M.; Srour, I. Using BIM to Retrofit Existing Buildings. *Procedia Eng.* **2016**, *145*, 1526–1533. [CrossRef]
13. Khodeir, L.M.; Aly, D.; Tarek, S. Integrating HBIM (Heritage Building Information Modeling) Tools in the Application of Sustainable Retrofitting of Heritage Buildings in Egypt. *Procedia Environ. Sci.* **2016**, *34*, 258–270. [CrossRef]
14. Khalil, A.; Stravoravdis, S.; Backes, D. Categorisation of building data in the digital documentation of heritage buildings. *Appl. Geomat.* **2020**, *13*, 29–54. [CrossRef]
15. Gigliarelli, E.; Calcerano, F.; Cessari, L. Heritage Bim, Numerical Simulation and Decision Support Systems: An Integrated Approach for Historical Buildings Retrofit. *Energy Procedia* **2017**, *133*, 135–144. [CrossRef]
16. Dias Pereira, L.; Tavares, V.; Soares, N. Up-To-Date Challenges for the Conservation, Rehabilitation and Energy Retrofitting of Higher Education Cultural Heritage Buildings. *Sustainability* **2021**, *13*, 2061. [CrossRef]
17. Biagini, C.; Capone, P.; Donato, V.; Facchini, N. Towards the BIM implementation for historical building restoration sites. *Autom. Constr.* **2016**, *71*, 74–86. [CrossRef]
18. Ma, Y.-P.; Hsu, C.C.; Lin, M.-C. Combine BIM-based and mobile technologies to design on-site support system for the communication and management of architectural heritage conservation works. In Proceedings of the 2017 International Conference on Applied System Innovation (ICASI), Sapporo, Japan, 13–17 May 2017; pp. 307–310. [CrossRef]
19. Jordan-Palomar, I.; Tzortzopoulos, P.; García-Valldecabres, J.; Pellicer, E. Protocol to manage heritage-building inter-ventions using heritage building information modelling (HBIM). *Sustainability* **2018**, *10*, 908. [CrossRef]
20. Australia ICOMOS. *The Burra Charter: The Australia ICOMOS Charter for Places of Cultural Significance*; Australia ICOMOS Incorporated: Burwood, VIC, Australia, 2013.
21. Misic, D.; Nahtigal, N.; Kocutar, S.; Mulec, M.; Mlakar, D.; Dugonik, B.; Hojnik, D. *Baronicina Hisa—Baroness' House: 1902–2015*; Fakulteta za Elektrotehniko, Racunalnistvo in Informatiko, University of Maribor: Maribor, Slovenia, 2015.

22. Wang, H.-J.; Chiou, C.-W.; Juan, Y.-K. Decision support model based on case-based reasoning approach for estimating the restoration budget of historical buildings. *Expert Syst. Appl.* **2008**, *35*, 1601–1610. [CrossRef]
23. Slota, M.; Leite, J.; Swift, T. On updates of hybrid knowledge bases composed of ontologies and rules. *Artif. Intell.* **2015**, *229*, 33–104. [CrossRef]
24. Pauwels, P.; Van Deursen, D.; Verstraeten, R.; De Roo, J.; De Meyer, R.; Van de Walle, R.; Van Campenhout, J. A semantic rule checking environment for building performance checking. *Autom. Constr.* **2011**, *20*, 506–518. [CrossRef]
25. Tibaut, A.; Kaučič, B.; Perhavec, D.D.; Tiano, P.; Martins, J. Ontologizing the Heritage Building Domain. In *Advances in Digital Cultural Heritage*; Ioannides, M., Martins, J., Žarnić, R., Lim, V., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 141–161. [CrossRef]
26. Baccigalupo, C.; Plaza, E. Poolcasting: A Social Web Radio Architecture for Group Customisation. In Proceedings of the Third International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution (AXMEDIS'07), Barcelona, Spain, 28–30 November 2007; pp. 115–122. [CrossRef]
27. Duque-Ramos, A.; Fernández-Breis, J.T.; Stevens, R.; Aussenac-Gilles, N. OQuaRE: A SQuaRE-based approach for evaluating the quality of ontologies. *J. Res. Pract. Inf. Technol.* **2011**, *43*, 159–176.
28. Saab, D.J.; Riss, U.V. Information as ontologization. *J. Am. Soc. Inf. Sci. Technol.* **2011**, *62*, 2236–2246. [CrossRef]
29. Lytras, M.; Sicilia, M.; Davies, J.; Kashyap, V. Digital libraries in the knowledge era: Knowledge management and Semantic Web technologies. *Libr. Manag.* **2005**, *26*, 170–175. [CrossRef]
30. Glossary of Building Construction Terms, Lexicon and Abbreviations. Available online: https://www.ats-group.net/glossaries/glossary-building-construction.html (accessed on 5 November 2021).
31. Glossary—UNESCO World Heritage Centre. Available online: https://whc.unesco.org/en/glossary/ (accessed on 5 November 2021).
32. Getty Vocabularies (Getty Research Institute). Available online: http://www.getty.edu/research/tools/vocabularies/ (accessed on 5 November 2021).
33. The Forum on Information Standards in Heritage (FISH) | Heritage Data. Available online: https://www.heritagedata.org/blog/about-heritage-data/fish/ (accessed on 5 November 2021).
34. WAND Building and Construction Management Taxonomy. Available online: https://www.wandinc.com/taxonomies/wand-building-and-construction-management-taxonomy (accessed on 5 November 2021).
35. CORDIS | European Commission. Available online: https://cordis.europa.eu/project/rcn/92042_en.html (accessed on 5 November 2021).
36. Trokanas, N.; Cecelja, F. Ontology evaluation for reuse in the domain of Process Systems Engineering. *Comput. Chem. Eng.* **2016**, *85*, 177–187. [CrossRef]
37. Scherp, A.; Saathoff, C.; Franz, T.; Staab, S. Designing core ontologies. *Appl. Ontol.* **2011**, *6*, 177–221. [CrossRef]
38. Guizzardi, G.; Wagner, G.; Almeida, J.P.A.; Guizzardi, R.S. Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story. *Appl. Ontol.* **2015**, *10*, 259–271. [CrossRef]
39. Gist—The Semantic Arts Minimalist Upper Ontology. Available online: https://www.semanticarts.com/gist/ (accessed on 5 November 2021).
40. DCMI: Home. Available online: http://dublincore.org/ (accessed on 5 November 2021).
41. Time Ontology in OWL. Available online: https://www.w3.org/TR/owl-time/ (accessed on 5 November 2021).
42. FOAF Vocabulary Specification. Available online: http://xmlns.com/foaf/spec/ (accessed on 5 November 2021).
43. ISA Programme Location Core Vocabulary. Available online: https://www.w3.org/ns/locn# (accessed on 5 November 2021).
44. W3C Semantic Web Interest Group: Basic Geo (WGS84 Lat/Long) Vocabulary. Available online: https://www.w3.org/2003/01/geo/ (accessed on 5 November 2021).
45. Person Core Vocabulary. Available online: https://www.w3.org/ns/person# (accessed on 5 November 2021).
46. Miles, A.; Bechhofer, S. SKOS Simple Knowledge Organization System Reference. W3C Recommendation. Available online: http://www.w3.org/TR/skos-reference/ (accessed on 5 November 2021).
47. Schema.org—Schema.org. Available online: https://schema.org/ (accessed on 5 November 2021).
48. Biagetti, M.T. An ontological model for the integration of cultural heritage information: CIDOC-CRM. *JLIS.it* **2016**, *7*, 43–77. [CrossRef]
49. Hernández, F.; Rodrigo, L.; Contreras, J.; Carbone, F.; Fundación, F.C.; Botín, M. Building a Cultural Heritage Ontology for Cantabria. In Proceedings of the 2008 Annual Conference of CIDOC, Athens, Greece, 15–18 September 2008; pp. 1–14.
50. Hassan, J.A.; Odéjóbí, O.A.; Ògúnfolákàn, B.A.; Adéjùwón, A. Ontology Engineering in Yorùbá Cultural Heritage Domain. *Afr. J. Comput. ICT* **2013**, *6*, 181–198.
51. Patino, O.P.Z.; Van Orshoven, J.; Steenberghen, T. Merging and expanding existing ontologies to cover the Built Cultural Heritage domain. *J. Cult. Herit. Manag. Sustain. Dev.* **2018**, *8*, 162–178. [CrossRef]
52. Gonzalez-Perez, C. An Ontology for Cultural Heritage. In *Information Modelling for Archaeology and Anthropology*; Springer International Publishing: Cham, Switzerland, 2018; pp. 195–215. [CrossRef]
53. Colucci, E.; Xing, X.; Kokla, M.; Mostafavi, M.; Noardo, F.; Spanò, A. Ontology-Based Semantic Conceptualisation of Historical Built Heritage to Generate Parametric Structured Models from Point Clouds. *Appl. Sci.* **2021**, *11*, 2813. [CrossRef]
54. Collections | Europeana. Available online: https://www.europeana.eu/en/collections (accessed on 5 November 2021).

55. Alani, H.; Brewster, C. Ontology ranking based on the analysis of concept structures. In Proceedings of the 3rd International Conference on Knowledge Capture, Banff, AB, Canada, 2–5 October 2005; ACM: New York, NY, USA, 2005; pp. 51–58. [CrossRef]
56. Vrandečić, D. Ontology Evaluation. In *Handbook on Ontologies*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 293–313. [CrossRef]
57. Duque-Ramos, A.; Quesada-Martínez, M.; Iniesta-Moreno, M.; Fernández-Breis, J.T.; Stevens, R. Supporting the analysis of ontology evolution processes through the combination of static and dynamic scaling functions in OQuaRE. *J. Biomed. Semant.* **2016**, *7*, 1–20. [CrossRef] [PubMed]
58. Poveda-Villalón, M.; Gómez-Pérez, A.; Suárez-Figueroa, M.C. OOPS! (OntOlogy Pitfall Scanner!). *Int. J. Semant. Web Inf. Syst.* **2014**, *10*, 7–34. [CrossRef]
59. OOPS!—OntOlogy Pitfall Scanner! RESTFul Web Service. Available online: http://oops.linkeddata.es/webservice.html (accessed on 27 September 2021).
60. DrOntoAPI Download | SourceForge.net. Available online: https://sourceforge.net/projects/drontoapi/ (accessed on 5 November 2021).