Umanistica Digitale – ISSN: 2532-8816 – n. 20, 2025: Special Issue
L. Fintoni, M. Daquino, F. Tomasi – *Rethinking scholarly digital objects as cultural heritage: the KNOT project*
DOI: http://doi.org/10.6092/issn.2532-8816/21188

Rethinking Scholarly Digital Objects as Cultural Heritage: the KNOT Project

Laurent Fintoni

Università di Bologna, Italia
laurent.fintoni2@unibo.it

Marilena Daquino

Università di Bologna, Italia
marilena.daquino2@unibo.it

Francesca Tomasi

Università di Bologna, Italia
francesca.tomasi@unibo.it

Abstract

The KNOT project is a three-year pilot tasked with investigating ways to integrate the digital cultural heritage (DCH) of Italian universities within the national infrastructure being developed by the Ministry of Culture. To do this, KNOT argues for rethinking the digital objects produced by academic research projects as interesting and, so far, unexplored examples of this DCH deserving of more attention and worthy of integration into the national infrastructure. This article discusses the key steps in the development of a conceptual framework in support of this argument starting with 1) the definition of these objects and identification of the potential heritage values they hold, 2) the selection of the humanities, and in particular the digital humanities, as the academic field from which to select them, 3) the development and publication of a data model and catalogue, and 4) some of the key issues that arose during this work around the classification, documentation, and visibility of these objects.

Keywords: Digital Humanities, Digital Cultural Heritage, Italian Universities, Documentation, AIUCD2024

KNOT è un progetto pilota triennale che ha il compito di studiare modi per integrare il patrimonio culturale digitale degli atenei italiani all'interno dell'infrastruttura nazionale sviluppata

dal Ministero della Cultura. Per fare ciò, KNOT sostiene la necessità di ripensare gli oggetti digitali prodotti dai progetti di ricerca accademica come esempi interessanti e, finora, inesplorati di questo patrimonio meritevoli di maggiore attenzione e degni di essere integrati nell'infrastruttura nazionale. Questo articolo discute i passaggi chiave nello sviluppo di un quadro concettuale a supporto di questa argomentazione a partire da 1) la definizione di questi oggetti e l'identificazione dei potenziali valori patrimoniali che detengono, 2) la selezione delle discipline umanistiche, e in particolare delle digital humanities, come campo accademico da cui selezionarli, 3) lo sviluppo e la pubblicazione di un modello di dati e di un catalogo e 4) alcune delle questioni chiave emerse durante questo lavoro sulla classificazione, documentazione e visibilità di questi oggetti.

Parole chiave: umanistica digitale, patrimonio culturale digitale, atenei italiani, documentazione, AIUCD2024

1. Introduction

Italian cultural heritage (CH) is diverse, omnipresent, and connected to the many histories found across its territory [10]. Among the institutions tasked with preserving and managing this heritage, universities stand out due to how their holdings are often intertwined across the three missions they must adhere to – teaching, research, and the dissemination of knowledge into wider society. However, while universities hold an important part of Italy's CH it has remained partially hidden within the national picture due to, in part, the late development of national discussions around its role and function, which only began at the turn of the 21st century [25]. This situation has since been further complicated by the rise and impact of digital technology, with universities having to contend with both analog holdings and newer digitized and born-digital objects that represent a digital cultural heritage (DCH) that has only grown in importance as the years have passed.

The challenge of how to integrate this particular segment of DCH at the national level has taken on new life recently within the National Plan for the Digitalization of Cultural Heritage (NPD),[1] which sets out the strategic vision and guidelines from the Ministry of Culture for the transformation of the country's CH over a five-year period (2022-2026) as part of the National Recovery and Resilience Plan. Drafted by the Central Institute for the Digitalization of Cultural Heritage – Digital Library (ICDP), a new body setup by the Ministry in 2020, the NPD seeks to create a digital ecosystem for CH with the ICDP guiding institutions and places of culture, such as universities, in this process.[2]

As part of this national effort, an agreement was set up between the University of Bologna and the Ministry to establish a research infrastructure to support the ICDP's efforts with parallel and connected research initiatives in three departments (Department of Cultural Heritage - DBC, Department of Classical Philology and Italian Studies - FICLIT, and Department of Computer Science and Engineering - DISI) tasked with investigating ways to integrate and showcase the DCH of Italian universities into this new digital ecosystem of culture [28].

The KNOT project, undertaken within the Digital Humanities Advanced Research Center (/DH.ARC), part of FICLIT, is a three-year pilot (2022-2025) that is part of this agreement.

[1] https://docs.italia.it/italia/icdp/icdp-pnd-docs/it/v1.1-febbraio-2023/index.html

[2] https://digitallibrary.cultura.gov.it/chi-siamo

Tasked specifically with investigating ways to showcase the DCH of Italian universities within the context of the ICDP's national effort to create a new digital ecosystem of culture, the project was faced from the start with a challenging proposition: how to effectively represent this DCH at a national level considering that to date there has been little consensus within institutions or across them as to what constitutes it?[3] This challenge led to an early decision to sidestep the more common understandings of DCH, such as digitized objects, which are already subject to research [6][17][29] and well supported by the infrastructure the ICDP is developing,[4] and instead consider how universities are already in possession of a multifaceted and interesting, and so far unexplored, example of DCH in the form of multiform digital objects created by academic research projects. To be clear, in the context of our research and this paper, these objects – which we term scholarly digital objects – are intended as those that may give access to or hold instances of aforementioned existing digitized CH objects, so for example a digital library that holds digitized copies of existing manuscripts created as part of an academic research project would be an example of such a scholarly digital object. Because these scholarly digital objects are most commonly evaluated and understood in Italy within a scientific context, as products of research, their CH dimension is often ignored or undervalued at the expense of the dimension of the CH objects they might make use of. Yet, we believe that they represent a valuable addition to the DCH of Italian universities deserving of attention. To the best of our knowledge this presents us with a research gap that is the starting point for the KNOT project, from which we will work to develop a data model, a web application, and guidelines for the collection, management, enrichment, and reuse of these scholarly digital objects.

In this article we present the first results of the KNOT project which focused on the development of a conceptual framework and data model to support the description and valorization of these scholarly digital objects as DCH and the implementation of this data model in a small-scale catalogue intended to show how these objects could be considered for integration in the ICDP's infrastructure.

The article is structured as follows: Section 2 outlines the definition of scholarly digital heritage, including the identification of the potential heritage values scholarly digital objects hold and the selection of the humanities, and in particular the Digital Humanities (DH), as the academic field from which to select these objects for the project based on related work; Section 3 presents a census of existing projects and objects that was used to evaluate the validity of these choices as well as provide data for the development of the data model; Section 4 details the KNOT Data Model, which is intended as a guideline for the description of digital scholarly objects and activity as cultural heritage using a common vocabulary; Section 5 presents the KNOT Catalogue, a small-scale catalogue that implements the data model and census data; and Section 6 discusses the issues that arose during this work around the classification, documentation, and visibility of these objects.

[3] The research conducted by the DBC, which ran from 2022 to 2024, was tasked with conducting a census of the DCH of Italian universities and the remark about the lack of consensus stems from the results of their investigation, which are still awaiting publication.

[4] https://ipac.cultura.gov.it/

2. Defining Scholarly Digital Cultural Heritage

The definitions of DCH put forward by the United Nations Educational, Scientific and Cultural Organisation (UNESCO) at the international level and by the National Plan for the Digitalization of Cultural Heritage at the national level provide a starting point for our rethinking of scholarly digital objects as DCH. UNESCO adopted its Charter on the Preservation of Digital Heritage in October 2003, defining the scope of digital heritage as consisting of "unique resources of human knowledge and expression", whether digital-born or converted from analogue, stemming from a plurality of backgrounds (educational, legal, scientific, and others), and represented by a multiplicity of formats, from text to databases, software to web pages [32]. The NPD meanwhile defines DCH in a more focused way, honing on "the set of digital objects produced by the modeling of data or by the organization of digitally-native content to achieve more advanced knowledge objectives through the development of the relational potential that characterizes its dissemination" which taken together and understood within an ecosystem logic contribute to the formation of a CH similar to the one assigned to tangible and intangible assets [21].

In addition to these definitions, there are valuable insights from the heritage sector that we should take into account when arguing for scholarly digital objects as examples of DCH. One such insight is the need to move our thinking and praxis of DCH beyond existing critical and social boundaries and extend the possibilities of what heritage can be which requires "a shift in thinking from the digital object as data, as information and materiality, to a form that is multi-agential and distributed, spatially and temporally continuous and extended, and that operates in a unique rhizomic fashion involving a proliferating number of coordinates and multiple forms of rationality, all of which has far-reaching implications for digital cultural heritage practice." [5]:279.

With this in mind we can therefore think of the scholarly digital objects we are interested in to comprise not just collections of information, most commonly represented by the catch-all term dataset, but, crucially, also digital forms that enable interaction with information, most commonly represented by software and data services such as search interfaces, Application Programming Interfaces (APIs), or interactive tools that allow a user to, for example, visualize, annotate, or represent digital information. Considering these objects thusly we can then begin to think past the scientific context in which they are most commonly understood, as well as evaluated, and look at what heritage value might be found in their "reservoir of meaning" [5]:34: the activity that produced them; relationships to the information they encode; new contexts they create for this information; and the ways in which they can foster the acquisition of new knowledge from this combination of context and information.

Having established our definition of DCH to specific objects of interest based on our argument, next we needed to narrow the scope of academic research to a specific field from which we could draw these objects. Considering that our research takes place within a DH research center at the University of Bologna, the humanities presented itself as the ideal field of reference for the project, a choice that was strengthened by recent research and previous efforts to engage with the DCH of universities in Italy.

At the national level, there have been a handful of humanities projects focused on the digitized CH holdings of universities, among them MICHAEL (Multilingual Inventory of Cultural Heritage in Europe) in the mid-2000s, a census of digitized collections from 77 Italian universities for inclusion in a European initiative coordinated by the University Library Center of the University of Padua,[5]

[5] https://bibliotecadigitale.cab.unipd.it/biblioteca-digitale/progetti/michael

and POMUI (POrtal MUseums Italian) in the early 2010s, the country's first network of university museums in which 12 institutions participated in an inventory and electronic catalogue of their holdings and the creation of a bilingual web portal to raise awareness of and promote this particular segment of their CH.[6] More recently, a bibliometric analysis of CH research in the humanities, based on data from the Web of Science platform (journal articles published between 2003 and 2022), showed how central the field has become to Italian academia's engagement with CH as a whole: Italian humanities scholars are among the top three in the world for the production of scholarly articles about CH and among the most cited [34]. Meanwhile, an analysis of projects presented at the yearly conference of the Italian Association for Digital Humanities and Culture (AIUCD) between 2018 and 2020 identified disciplines of the text (publishing, philology, literature, and linguistics) and the management of DCH as the dominant areas of focus for Italian digital humanists [29]. The DH in particular offer further opportunities for our project as they promote interdisciplinary approaches [30] and project-oriented practices where existing CH objects are often used as source material for the creation of new digital objects that aim to recontextualize this material to promote engagement from end users both within and outside of academia. In many ways the DH, both in its (relatively young) history and practices, reflect the plurality of approaches to empirical investigation of the human past found within CH, and which should be understood as "not an obstacle to be overcome but a constitutional condition of the domain" itself [4], as well as some of the heritage values that can be found in the digital forms mentioned above. These aspects of the field are also aligned with the goals of I.PaC (Infrastruttura e servizi digitali per il Patrimonio Culturale), the online infrastructure the ICDP is building and within which our results are expected to be included.

3. A Census of Italian Humanities Projects

In order to evaluate the validity of our conceptual approach and choice of the humanities as outlined in Section 2, we conducted a census of existing Italian humanities research projects and analyzed the results. In addition, this census would allow us to extract potential features and characteristics of scholarly digital objects and activity to use in the development of a data model which could be used to describe these objects and their associated activity. This data model is detailed in Section 4.

We began our census by looking at research projects listed on the AIUCD[7] and European Association of Digital Humanities (EADH)[8] websites. We then visited the websites of specific humanities departments, research centers, and laboratories that were included in these lists as well as those of members to look for additional relevant projects. Finally, we also conducted manual searches within the official websites of universities, including digital libraries, for further potential projects and lists of relevant resources. At each step we focused on recording projects, whether finished or ongoing, that had produced one or more digital objects that could be considered as representative of our argument such as catalogues, databases, software, corpora, archives, and scholarly editions. This produced a list of 91 projects for which we then recorded a variety of metadata based on available information including: links to websites and data; the agents involved, including entities (universities, departments, labs, research centers) and

[6] http://www.pomui.unimore.it/site/en/home/project-2012.html

[7] http://www.aiucd.it/progetti/

[8] https://eadh.org/projects

individuals; the year of creation; the primary subject and time period of study; the types of CH objects used and the types of digital objects created; the tools and technologies used in the creation of the digital objects; the research activities (based on TaDiRAH, a taxonomy of digital research activities in the humanities); [9] the licenses; the state of the project's documentation; and the degrees to which projects made their results available via Linked Open Data practices or technologies such as APIs.

The results of this census presented an overall picture that gave us confidence the humanities could offer the necessary variety of scholarly digital objects needed for the project, both to help develop the data model but also to populate a web application, while also offering a representation of the variety of the DCH these objects can present. Alongside what we can consider traditional humanities objects in the Italian context, such as corpora, digital editions, catalogues, and databases, were examples of newer ones that reflect the digital forms of interest mentioned in Section 2 such as visualization tools, machine learning pipelines, software, and the hypernym term "platform". In terms of the humanities disciplines involved, computer science, which is an integral part of the digital in Digital Humanities, was unsurprisingly well represented alongside philology, literature, and linguistics which [29] noted as key areas of interest for Italian digital humanists. In addition to these expected disciplines we also found projects based in or involving design, musicology, archeology, and archival science with some, such as PATHs,[10] making explicit reference to the DH as part of their multidisciplinary approach.

4. The KNOT Data Model

The process of modelling in the DH is key to moving from theory to practice [31], turning critical thinking and interpretation into a potentially practical object such as a data model. Furthermore, modelling, both in the CH and DH fields, is not just about the representation of the source but also about the process itself, in part because this process involves representing and capturing some of the inherent subjectivity and peculiarities of human creations and activities to express principles that are anchored in a specific context rather than a general law [7].

As such, our goal was to create an easy to understand and use data model that could capture the peculiarities of our argument and act as a guideline for the description of digital scholarly objects and activity as digital cultural heritage using a common vocabulary and facilitating the dissemination of good practices. To design the KNOT Data Model (KNOT-DM) we took inspiration from the Simplified Agile Methodology for Ontology Development (SAMOD) [27], which borrows from the field of software development to bring flexibility and ease-of-use to ontology development via data-centric iterative workflows.

Firstly, we used the results of the census to create an initial set of requirements for the data model, based on the features and characteristics we had collected. These requirements included: being able to describe both scholarly digital objects and the activity that created them, including agents, locations, subjects, research activities, technologies used; being able to describe different types of objects (collections of information and digital forms for interacting with information) and different types of access (download, interface, APIs, etc.); being able to describe the CH

[9] https://tadirah.info/pages/Browser.html

[10] https://atlas.paths-erc.eu/

dimension of these objects, including that of the CH objects they might use as source as well as that of the scholarly digital object itself.

Alongside these requirements we also made the design decision that KNOT-DM should be focused on reusing and combining existing formal ontologies and Knowledge Organization Systems (KOS) rather than creating new ontological components, in order to allow flexibility and ensure eventual compatibility with the ICDP infrastructure. This infrastructure was at the time we began our work still being finalized and the only confirmed standard was an application profile of METS, called METS-ECOMiC,[11] which the ICDP developed as part of the NPD in order to support ingestion of existing DCH into I.PaC.

This led us to evaluate various existing standards, models, ontologies, and vocabularies with the following criteria in mind: interoperability, to facilitate access to and automated processing of the data created by our model [22]; existing, and ideally ongoing, development support; established uses and practical real-world applications both in Italy and internationally. We evaluated: DCAT,[12] a W3C RDF vocabulary designed to facilitate interoperability between data catalogs published online; the aforementioned TaDiRAH taxonomy and the associated NeDiMAH Methods Ontology,[13] an ontological model of scholarly practice in the arts and humanities; the Common European Research Information Format (CERIF),[14] an information model for scientific research intended to support interchange of research information between and with CRISs that includes a data model and suite of semantic vocabularies; the VIVO ontology,[15] used to represent the expertise of people engaged in the creation, transmission, and preservation of knowledge and creative works; RO-Crate,[16] based on the Research Object ontology, which offers a lightweight approach to packaging research artefacts along with their metadata in a machine readable manner based on JSON-LD and schema.org; the ontology for the Portland Common Data Model,[17] intended to underlie a wide array of repository and DAMS applications and maintained by Lyrasis, a company that supports research infrastructures and open access through a variety of software solutions and community efforts; the European Union's EU vocabularies;[18] Metadata4Ing,[19] an ontology for describing the generation of research data within a scientific activity; the NFDIcore Ontology,[20] focused on describing Digital Information Artifacts, independent autonomous agents, and planned processes and events; the OntoPiA network (or stack) of ontologies and vocabularies,[21] created and maintained by the Agenzia per l'Italia Digitale, which includes metadata, domain-level, and supporting ontologies

[11] https://ipac.cultura.gov.it/2024/04/19/mets-eco-mic/

[12] https://www.w3.org/TR/vocab-dcat-3/

[13] http://nedimah.dcu.gr/index.php?p=home

[14] https://eurocris.org/services/main-features-cerif

[15] https://wiki.lyrasis.org/display/VIVODOC113x/VIVO+Ontology+Domain+Definition

[16] https://www.researchobject.org/ro-crate/

[17] https://pcdm.org/2016/04/18/models

[18] https://op.europa.eu/en/web/eu-vocabularies

[19] https://nfdi4ing.pages.rwth-aachen.de/metadata4ing/metadata4ing/index.html

[20] https://ise-fizkarlsruhe.github.io/nfdicore/

[21] https://github.com/italia/daf-ontologie-vocabolari-controllati/wiki

and vocabularies relevant to our project such as ArCo, the knowledge graph of Italian CH; Nanopublications,[22] small knowledge graphs that act as FAIR containers containing information about assertion and provenance; the PROV Ontology (PROV-O),[23] a W3C ontology designed to represent provenance information generated in different systems and under different contexts; and, lastly, CIDOC Conceptual Reference Model (CIDOC CRM),[24] a well-used and established standard within the field of CH used for information integration [4].

What emerged from these initial steps was the understanding that the domain we were interested in modelling lay at the intersection of three existing domains that represent specific aspects of our argument for scholarly digital objects as examples of the DCH of Italian universities: the domain of academia, where research activity and object production takes place; the domain of public data, which represents the availability of scholarly digital objects to the public, whether specialized such as researchers or broader such as members of the public; and the domain of cultural heritage, which represents both the DCH dimension of scholarly digital objects, as outlined in Section 2, and the CH dimension of existing objects used in research. This reinforced our choice to focus on reusing existing standards and ontologies in order to further support our work and give prominence to this intersection, which is represented as a Venn diagram in Figure 1.

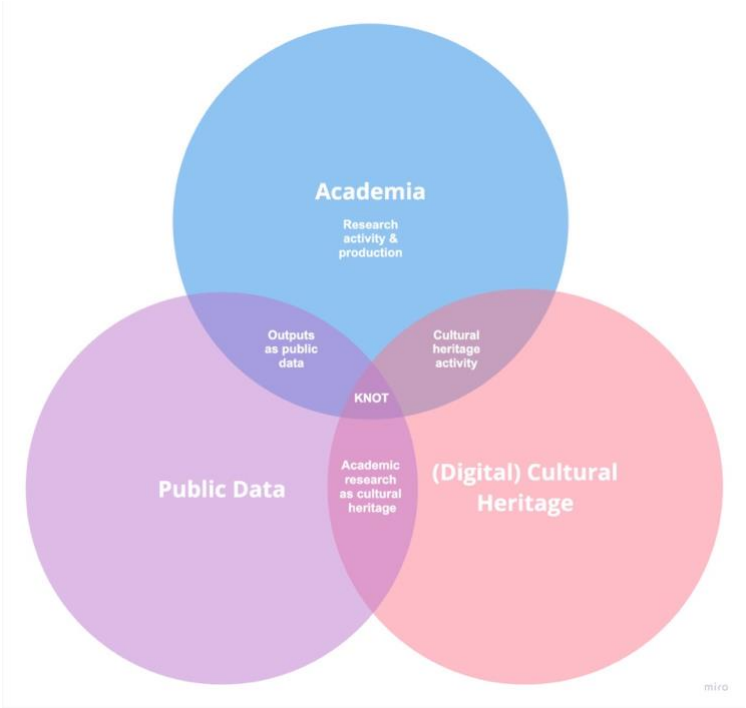

Figure 1 Intersection of the domains covered by the KNOT Data Model

[22] https://nanopub.net/

[23] https://www.w3.org/TR/prov-o/

[24] https://www.cidoc-crm.org/

Next, we mapped this intersection of domains to equivalent segments in our data model (public data, academia, cultural heritage) so that we could test how some of the specific ontologies and KOS we had assessed could be used and connected to each other. To do this we modelled a small subset of data from our census (12 entries, intended to represent the variety of scholarly digital objects hinted at by our census such as databases, digital libraries, knowledge graphs/bases, digital editions, catalogues, corpora, digital storytelling, and digital tools) and evaluated the results against a set of competency questions drawn from our census analysis, grouped into categories aligned with each segment, and intended to answer key who, what, where, when, and how questions (e.g. What type is the digital object? What objects are available from which agent? What research activities did a project use to create the object? When was the object first made available? Where are the agents who produced the object based? How does the digital object use existing CH?). The ontologies and KOS we chose to test included DCAT, Nanopublications, RO-Crates (alongside schema.org), PROV-O, CIDOC CRM, the CERIF Output Types and Activity Sub Types vocabularies, and the TaDiRAH taxonomy. DCAT was used to model the public data segment, RO-Crates and PROV-O were used to model the academia segment, and CIDOC CRM and Nanopublications were used to model the cultural heritage segment while the CERIF and TaDiRAH vocabularies were used to handle the types of objects produced and research activities involved, which emerged early on in this process as a key aspect of our data model.

Following four rounds of testing in which each was applied to the relevant segment and evaluated against the competency questions, we decided to settle on the following to build KNOT-DM: DCAT, to model the public data domain, as this is already included in the OntoPiA network (as an Italian application profile), makes use of standard DCMI Metadata Terms properties which facilities alignment with existing systems, and allows us to describe scholarly digital objects as both datasets and data services, a central element of our argument, as well as describe associated digital objects such as the websites of research projects; PROV-O, to model the academia domain, chosen because it is already integrated with DCAT to describe the provenance of digital objects, giving us an easy way to connect these two domains while being flexible enough to cover many aspects of academic research activity; CIDOC-CRM to model the CH domain, as it widely used and offers extensions that are relevant to our project; and, lastly, TaDiRAH to describe the research activities, as it is most aligned with the humanities from which our objects are drawn and is well-known and extensively used in the field [18].

With these choices made we then spent some time further refining which elements of each the model should use and how they should connect in order to enable the best representation of our domain intersection. This step led us to decide to use DCAT classes and properties from both the W3C specification[25] and the European[26] and Italian[27] Application Profiles, which offer some additional granularity, such as distinctions between data service and distribution, and practical adaptations at the national and European level, such as the use of EU-Vocabularies and Italian government vocabularies. For CIDOC CRM we chose to focus our usage on two

[25] 2.0 at the time of development, but capable of using the recently released 3.0 as well.

[26] https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/solution/dcat-application-profile-data-portals-europe/releases

[27] https://www.dati.gov.it/content/dcat-ap-it-v10-profilo-italiano-dcat-ap-0

extensions, CRMdig[28] and LRMoo[29] (formerly FRBRoo), which we believe allows our model to better represent our argument. The former models the production and provenance of digitized objects and digital representations, giving us the ability to describe specific events that led to the creation of a scholarly digital object, while the latter represents semantics about bibliographic information and is used to enable relations with the WEMI (Work, Expression, Manifestation, Item) concept of the Functional Requirements for Bibliographic Records model. We should note here that there exists another CIDOC extension, called CRMpe, which emerged from the Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies (PARTHENOS) project in the 2010s and was designed to capture and represent the knowledge generation process within the CH, digitization, and provenance domains by making use of both CIDOC CRM and CRMdig [13] [24]. However, the extension is not included in the official list provided by the CIDOC CRM website and its existing published RDF version[30] offers limited implementation examples or scope notes, indicating that it is not yet ready for use though it does offer something similar to what our data model aims to do but contained within the wider CIDOC CRM standard. Lastly, we also decided at this stage to create two controlled vocabularies as part of the KNOT-DM which could be used to describe the types of scholarly digital objects created by research projects (a taxonomy) and the technologies used in their creation (a thesaurus). The need to create these vocabularies as part of our data model became apparent as none of the existing vocabularies we had tested or evaluated were able to capture some of the nuances of the types of objects created by humanities research nor provide adequate authority control for the technologies used by humanities research. Further details of these vocabularies are given in Section 6.

Using insights acquired from each of these steps we finalized the first version of the KNOT-DM and published it in late 2023 in RDF as the KNOT Ontology (KNOT-O), which collects all terms and concepts, and the KNOT Taxonomy and KNOT Technology Thesaurus, two controlled vocabularies based on the Simple Knowledge Organization System (SKOS)[31] data model. The use of existing standards based on, or adapted to, RDF enables the KNOT project to publish its metadata as Linked Open Data (LOD) to ensure machine readability and maximize its interoperability with other existing data available online.

Figure 2 shows a summary of the core of KNOT-DM and its intersection of the three domains by highlighting key classes (yellow boxes) and properties (arrows) from each standard and their interconnection as well as relationship to examples of class instances (pink dots) central to the project.

[28] https://cidoc-crm.org/crm-dig/

[29] https://cidoc-crm.org/frbroo/

[30] https://ontome.net/namespace/216

[31] https://www.w3.org/TR/skos-reference/

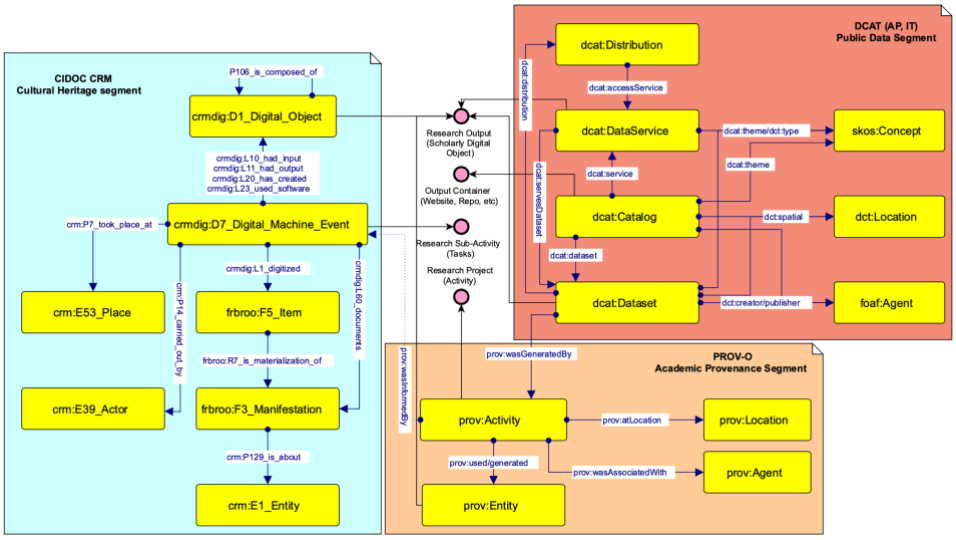

Figure 2 Main classes and properties in KNOT-DM, their associated segments, and examples of class instances.

The key classes in the KNOT-DM are as follows: for the public data segment `dcat:Dataset` and *`dcat:DataService`* represent scholarly digital objects, *`dcat:Distribution`* represents the downloadable form of an object, and *`dcat:Catalog`* represents the container for scholarly digital objects created by a project and most commonly represented in the real world by an official project website; for the academia segment *`prov:Activity`* represents the research project as activity and is connected to DCAT via *`prov:wasGeneratedBy`* to indicate the provenance of a *`dcat:Dataset`* or *`dcat:DataService`*; for the cultural heritage segment *`crmdig:D1_Digital_Object`* (a subclass of CIDOC's E73 Information Object) represents scholarly digital objects within a cultural heritage context, which is not explicit in DCAT, while *`crmdig:D7_Digital_Machine_Event`* (a subclass of CIDOC's E65 Creation) represents specific activities within the research project (connected to PROV-O via *`prov:wasInformedBy`* whose range in KNOT-DM is extended to include the CRMdig class), while *`frbroo:F5_Item`* and *`frbroo:F3_Manifestation`* are used to represent physical objects that may be digitized as part of a research project and their potential relation to real world entities.

The KNOT-DM allows the following information about scholarly digital objects and activities to be recorded:

- Research project activity: who, what, where, when, and how.
- Research project outputs: the scholarly digital objects created by research project activities.
- Research project inputs: the entities used by a research project such as physical items, people, places, and themes.

- Agents: involved in research projects and the creation of outputs both as contributors and publishers.[32]
- Relationships: between activities, agents, entities, locations, and concepts such as the influence of external concepts or the reuse of scholarly digital objects between projects.
- Spatial and temporal information: including that of the projects and their agents but also that recorded in scholarly digital objects, such as for example the places mentioned in a text or the time period covered by the data.
- Concepts: such as the type of objects produced, the technologies used, the research activities, the technical specifications, the academic disciplines involved, and the subjects of research. These concepts are described using a selection of controlled vocabularies and authority controls.

Lastly, we should note that the three standards used within KNOT-DM each include classes which refer to similar concepts, in particular entities, activities, location, and agents. While this could be dealt with by using the `owl:equivalentClass` axiom or the `rdfs:subClassOf` property (choosing a top-level class or pointing to a specifically created class), these classes are respectively defined within each standard with enough difference that we decided to not consider them as being equivalent within KNOT-DM or create a new class for them even though this adds a degree of potential complexity to the use of the data model. As such individuals described with KNOT-DM can be assigned each class as appropriate. A full documentation of the data model as well as further conceptual maps are available on our website.[33]

5. The KNOT Catalogue

Following the publication of the first version of the KNOT-DM, we began work on the KNOT Catalogue, a web application intended to put the data model into practice by using it to describe the results of our census, create a knowledge graph that can be accessed via both a traditional record interface and a more advanced SPARQL endpoint, and experiment with semantic enrichment and services.

Similarly to how we had evaluated existing ontologies and KOS for use in the KNOT-DM, we evaluated software solutions for digital catalogues, repositories, and Content Management Systems (CMS) against some basic requirements: does it support Linked Open Data? What services are included or available for inclusion, such as reconciliation with authority controls, visualization, or API endpoints? Is it open source and how much work is needed for upgrade and maintenance? The solutions we evaluated included established and popular options such as

[32] KNOT-DM also makes use of some classes and properties from the W3C's Organization Ontology to help further detail agents from the individual to the organization via units (such as departments or laboratories) as this ontology is connected to FOAF which is used by DCAT to represent agents.

[33] https://icdp-digital-library.github.io/KNOT/website/ENG/data_model.html

Omeka,[34] a CMS for digital collections, Islandora,[35] a Digital Repository System based on Drupal, and DSpace,[36] repository software most commonly used for academic publications, alongside more specific options such as ResearchSpace,[37] a platform focused around using knowledge graphs and CIDOC-CRM to contextualize collections, Paradisec,[38] a catalogue solution based on RO-Crates and OCFL developed by a consortium of Australian researchers, the Academy extension for Typo3,[39] a CRIS framework for a popular CMS solution developed by the DH center at the Academy of Sciences and Literature Mainz, and lastly Crowdsourcing Linked Entities via web Form (CLEF),[40] a Linked Open Data-native cataloguing and crowdsourcing platform developed at the University of Bologna as part of the Polifonia H2020 project.

While most of these solutions met the potential needs of our project, the decision was made to use CLEF as it included some key services such as reconciliation and a SPARQL endpoint while also offering the possibility of allowing external users to submit new scholarly digital objects for inclusion as well as review and update existing records. Furthermore, CLEF was developed by and for digital humanists and supports and encourages the reuse of external ontologies [9], which aligns with the focus of the KNOT project.

Despite this, CLEF did also have some limitations with regards to a practical implementation of our data model. In CLEF, resources are described using a system of templates, with each template mapped to a class and its various fields mapped to predicates, that are then connected to input forms and used to generate named graphs for each record with an editorial process that can consist of up to three phases (record creation, record modification, and review and publication) [9]. As a result of this data management approach, the objects of triples cannot be described as instances of specific classes that are not themselves mapped to a template. Furthermore, at the time we decided to use CLEF it had some specific technical limitations, such as the lack of support for temporal datatypes or the use of external controlled vocabularies and authority controls aside from Wikidata and Geonames. These limitations meant that it was not possible to make use of the full KNOT-DM, as for example we would not be able to describe agents, locations, or classes that represent more abstract concepts such as downloads without creating entire records for them. This led us to use a simplified version of the data model, aligned with these constraints yet still able to describe scholarly digital objects and activities meaningfully and with the goals and specificities of our project in mind.

Therefore, the KNOT Catalogue makes use of three key KNOT-DM classes to create records: `prov:Activity`, to represent research projects and connect them to the objects they have created and referred to in the front end as "Research Project"; `dcat:Dataset`, to represent scholarly digital objects when these objects represent collections of information that has already been structured in some way (rather than raw, unstructured data), and referred to in the front

[34] https://omeka.org/

[35] https://www.islandora.ca/

[36] https://dspace.org/

[37] https://researchspace.org/

[38] https://language-archives.services/about/modpdsc/

[39] https://github.com/digicademy/academy

[40] https://github.com/polifonia-project/clef

end as "Digital Object"; and `dcat:DataService`, to represent scholarly digital objects when these objects represent digital forms that enable interaction with information via user interfaces, and referred to in the front end as "Web Service". As per the full data model, these three classes can be connected to each other and these connections are represented in the catalogue, allowing for navigation between records. This focus on three classes does however limit the use of properties and introduces some new usages, such as for example the addition of the `foaf:homepage` property to `prov:Activity` and an additional property for status of the project which makes use of `dbpedia:currentStatus` which CLEF already uses to note the status of records within the system. Full documentation of how the KNOT-DM is adapted to and used in CLEF is available on the website.[41]

The results of our census were used to populate the catalogue. Templates for each class were created in CLEF, following our simplified data model, alongside equivalent spreadsheets into which census data was transferred and then transformed into RDF using the excel2rdf Python package.[42] The resulting RDF files created were then added to the catalogue directly via Blazegraph, which is the triple store implementation that CLEF includes by default, and used to generate records. Each record was also given a URI attached to a Persistent Uniform Resource Locator via purl.org.

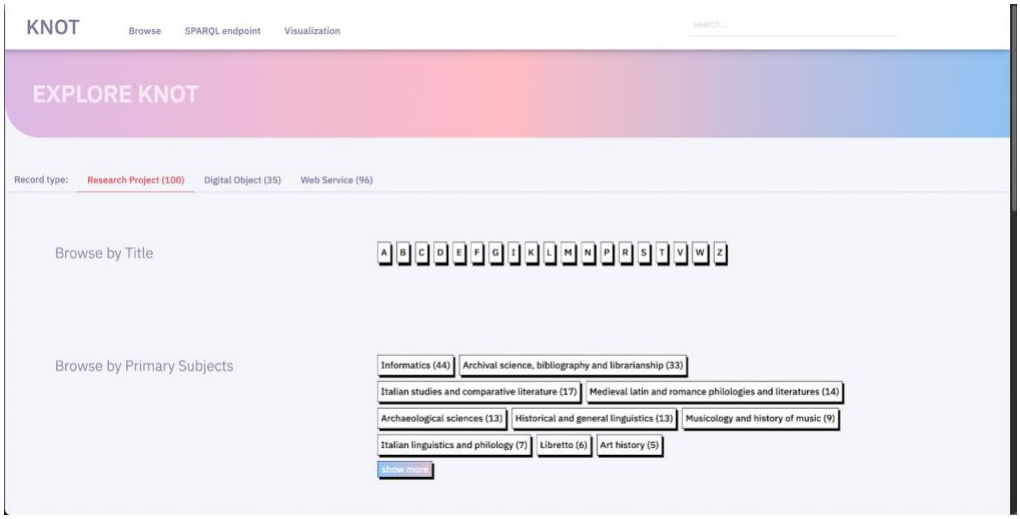

Figure 3 KNOT Catalogue index page.

Records in CLEF can be browsed via an index page, as shown in Figure 3, which allows users access to each type of record alongside a selection of filters. These filters are set via the templates

[41] https://projects.dharc.unibo.it/knot/documentation

[42] https://pypi.org/project/excel2rdf/

and were chosen to reflect specificities of each record type aligned with our data model: for example, research projects can be filtered by primary subject or research activity while digital objects can be filtered by types and license. For more advanced users, all records can be accessed via a GUI to query the application's SPARQL endpoint.

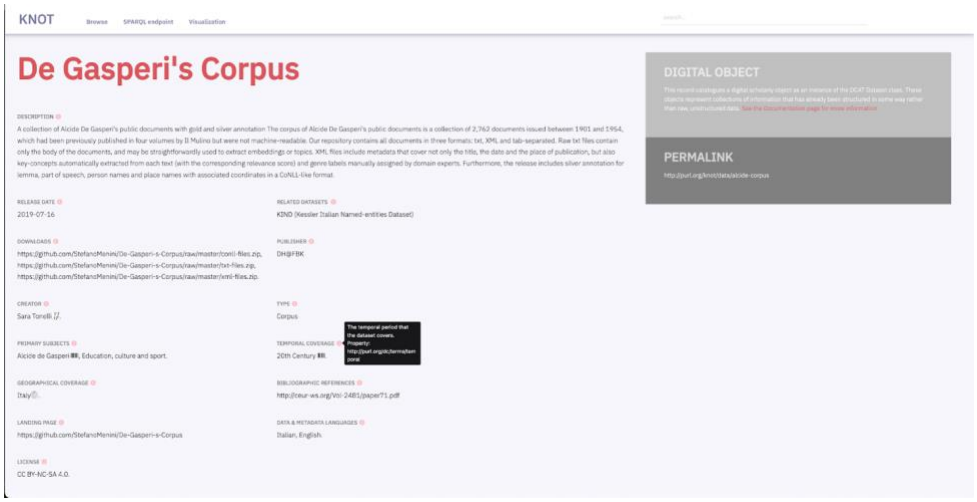

Figure 4 Example record page for a digital object in the KNOT Catalogue.

As shown in Figure 4, each record page includes a short explanation of the class it is attached to and its definition and intended usage (with a link to the data model documentation) as well as an explanation of what each field represents and its connected RDF property, accessible via a hover pop-up mechanism. Information such as agents, locations, and concepts is linked to external authority controls and vocabularies via a secondary page type that also displays where else this information appears in the catalogue (as shown in Figure 5) thus allowing users an additional way to navigate the catalogue as well as visualize the usage of this information across projects and objects.

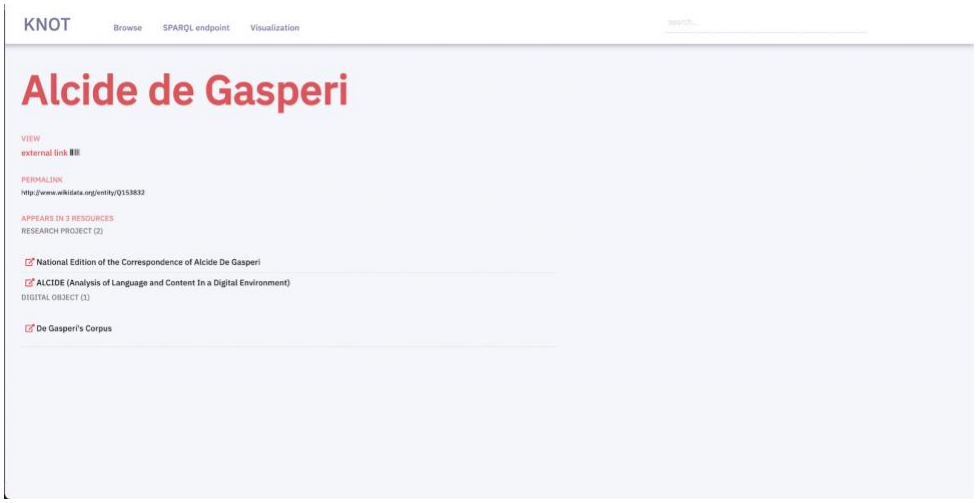

Figure 5 Example page for a historical subject in the KNOT Catalogue.

The KNOT Catalogue was published in spring of 2024 and currently includes records for 100 research projects (the 91 original projects from the census and some additions collected at regular intervals), 35 digital objects, and 96 web services. A dedicated page in the catalogue offers a selection of visualizations through which to navigate the contents of the catalogue and some of the more interesting aspects of the knowledge graph.[43] These visualizations are inspired by some of the competency questions used in the development of the data model and offer insights into which universities are represented (as well as the departments, faculties, or research centers and laboratories involved), the types of digital objects and web services created by humanities research in Italy, which academic disciplines and subjects are involved, and how scholarly digital objects and activities are described.

6. Issues and Challenges

The development of the KNOT-DM and Catalogue evidenced a number of issues for us to consider and confront regarding the classification of scholarly digital objects, their documentation, and their visibility within the Italian academic system.

As noted in Section 4, the first of these issues we had to confront relates to classification. Many of the projects we included in our census referred to the objects they had created in their documentation using different, sometimes complimentary, sometimes contradictory, terms. For example, Archivio della Latinità Italiana del Medioevo (ALIM)[44] uses the term archive in its title and project description but the term digital library in its interface. As a scholarly activity, ALIM is creating a digital archive in that it preserves items because of their perceived value to one or more communities (in this case Latin texts produced in Italy during the Middle Ages and their value to scholars), yet the digital form created by this activity, the scholarly digital object through

[43] https://projects.dharc.unibo.it/knot/numbers

[44] http://alim.unisi.it/

which this archive is accessed, more closely resembles a digital library with content classified into collections, available for retrieval, and with additional services for users [3]. Another example is DanteSources,[45] which produced an online application with which to visualize Dante's primary sources through semantic representation. In its documentation and publications, the project refers to this scholarly digital object as a digital library, tool/application, and knowledge base. The latter two more closely reflect the affordances of the object (especially when you consider that the interface only indicates to the user the ability to search) while digital library feels less appropriate as a term in that there are no practical resources to be found and retrieved aside from information extracted from the analysis of resources. Based on the results of our census and development of the KNOT Catalogue we found that typological terms such as archive, library, catalogue, and database were most commonly treated as interchangeable in documentation and interfaces while digital edition was the most consistently, and accurately, used term, which is unsurprising considering the much narrower range of such an object.

As our work progressed it became apparent that we needed a way to classify scholarly digital objects produced by humanities research using a reliable, empirical system such as a taxonomy. However, while there have been efforts within the international DH community to create useful taxonomies of research methods, tools, and activities – notably the DiRT Directory of digital research tools [26], the TAPoR gateway for text analysis tools,[46] and the aforementioned TaDiRAH (the latter two integrating elements of or knowledge from DiRT [1] [16]) – less focus appears to have been given to a taxonomy for the objects such methods, tools, and activities might produce. Early versions of TaDiRAH did include both research methods and objects, though these objects were intended to reflect a range of digital research objects that included both data types associated with methods as well as potential inputs rather than the specific objects created by research methods [2]. The current version of TaDiRAH, 2.0, published in 2020, no longer includes objects.

Parallel to this issue of taxonomy we also noticed that many of the common technologies used in the creation of scholarly digital objects, such as programming languages like Python, formats and standards like JSON or TEI, and software packages like Omeka, were not always reliably included or detailed in existing authority controls and thesauri, whether well-established ones such as the Library of Congress Subject Headings and the Getty Art & Architecture Thesaurus or more specific ones related to the humanities and CH such as those available in the OntoPiA network or the DHA Taxonomy from the Austrian Centre for Digital Humanities and Cultural Heritage.[47]

This led us to develop two SKOS-based controlled vocabularies as part of the KNOT-DM, as explained in Section 4, aimed at answering these specific classification needs while also providing the necessary controls for integration of this information into retrieval systems [20] such as the KNOT Catalogue. The KNOT Taxonomy is intended to provide a correspondence between information needs and objective reality while taking into account the fluidity of the types of scholarly digital objects that emerges from practice and theory while the KNOT Technology Thesaurus offers a way to explore terminology related to technology while enabling the production of standardized data for reuse [4]. In both cases, we have drawn from existing research and definitions to support our own while using SKOS properties for semantic

[45] https://dantesources.dantenetwork.it/en/index.html

[46] https://tapor.ca/home

[47] https://vocabs.acdh.oeaw.ac.at/dha_taxonomy/en/

(*skos:related*) and mapping (*skos:relatedMatch*, *skos:closeMatch*) relations to express internal connectivity between our concepts and external connectivity to concepts maintained by the LoC's Subject Headings, the Getty AAT, and the ACDH's DHA taxonomy. Furthermore, these vocabularies are intended to act as a base for ongoing reflection regarding definitions and meanings of scholarly digital objects and their associated technologies. Taking a cue from the rethinking of methods and activities as scholarly primitives that motivated the creation of TaDiRAH [1] [33] we can see how thinking about the definition of a set of primitives for scholarly digital objects, reflecting the basic functions that these objects fulfill for scholarly activity in the humanities alongside their practical functions, which may push against accepted definitions, could be useful both within and outside of the field. These insights and reflections are a key aspect we intend to include in the guidelines for the collection, management, and enrichment of scholarly digital objects as DCH at the national level that the KNOT project is expected to produce at the end of its duration. Documentation of the development and use of both vocabularies are available on our website.

Connected to this issue of classification is the inconsistent quality of the documentation of the scholarly digital objects we recorded and catalogued. Inconsistent documentation makes it more difficult to accurately describe scholarly digital objects using a data model by obscuring their narrative [11]. Previous studies of documentation of DH resources, going back to the late 2000s, have found that documentation is often insufficient or lacks clear visibility [12][15][35] and this is something we have also seen in our work to date with projects often either lacking or obscuring important information such as clear summaries of the project or general descriptions, their technology stack, their status (whether ongoing or concluded), and the licenses available for use of the objects created (whether data or service). One notable exception is the handful of projects we have recorded that openly strive to meet the FAIR principles for scientific data management and stewardship [36], which often results in clearer, more accurate, and extensive documentation. Returning to the previous examples, ALIM offers a fairly complete documentation of the project on their website while DanteSources does not detail its technology stack nor indicate the status of the project or licenses for use of its data and services. LiLa: Linking Latin[48] is a project that openly strives to meet FAIR requirements and provides one the most complete documentation of all the research projects we have catalogued. A recent checklist for documentation of humanities data that builds on previous studies recommends 13 components of consideration [23] that match many of the same problematic areas we have encountered in our research (such as general information, scope, functionality, data provenance, data access and reuse, and publications), underlining the importance of the issue while also providing a useful point of reference for our own guidelines.

Lastly, our census also brought us to consider issues of visibility for the digital objects we are interested in. While Italy boasts an active DH community, there are no official positions for the discipline, as exist in other countries, leading to difficulties in both professional recognition for academics as well as for the products of their research [19]. This is further compounded by the ways in which the National University Council organizes disciplinary areas and how it and the National Agency for the Evaluation of the University and Research Systems evaluate scientific production. This leaves interdisciplinary research, central to the DH as noted previously, subject to evaluation under a specific area (rather than across the different areas of the different disciplines involved) and many of the digital objects produced by such research as deemed not always worthy of scientific recognition [30]. While the growth of European research

[48] https://lila-erc.eu/

infrastructures focused on the arts and humanities, such as DARIAH[49] and CLARIN,[50] has helped in this regard by providing other avenues for the dissemination of digital scholarly objects and collaboration at the international level, at the national level visibility for the scholarly digital objects we are interested in remains a critical issue. Another aspect of this visibility issue that emerged from our research, and which is also related to the issue of documentation, is that within the current Italian system these objects and their various components are often scattered across institutional and public repositories and websites without clear connections thus further obfuscating them. A common example of this is as follows: a digital object produced by a research project will have an accessible public interface on an official site hosted on an institutional sub-domain, yet this official site may not be linked to from the primary institutional domain; in turn, publications related to the creation of the object may be included in an institutional repository without a clear link to the official site where the online interface can be accessed in the text or in the metadata; and lastly, some of the components of the object itself, such as the underlying data, may be available via an external repository like GitHub but not clearly connected to the official site or other documentation. This has a direct impact on the amount of work necessary to both evaluate and describe a project and the objects it created, and it likely also has an impact on the visibility of these digital scholarly objects for other potential users. While it is beyond the scope of the KNOT project to address issues of visibility caused by the specifics of the Italian academic system, we believe that arguing for digital scholarly objects to be considered as examples of DCH worthy of integration and preservation within the national infrastructure has the potential to bring further attention to issues of access and connectivity between humanities research projects and the objects they produce by raising their visibility outside of academia while also providing an updated perspective on their CH value that goes beyond that of the objects they may use as source material.

7. Conclusion and Future Work

The development of a national infrastructure for DCH such as the one currently being undertaken by the ICDP offers a unique opportunity for Italian universities to engage with this segment of their CH holdings in new ways. To do this will require a broadening of the understanding of DCH that takes into consideration how scholarly digital objects can be more than data and how meaningful heritage value lies in digital forms that enable interaction with information. By focusing on the scholarly digital objects produced by humanities research as relevant examples of such DCH already held by universities, the KNOT project aims to provide a useful base from which to let new insights emerge around how to valorize these objects both within and outside of academia.

The challenges we have highlighted around classification, documentation, and visibility have been key areas of focus in the development of the KNOT-DM and its integration and use in the KNOT Catalogue. Updates to both the data model and vocabularies will continue throughout the duration of the pilot, with the latest updates made following the publication of the Catalogue, to reflect the changes required by the use of CLEF, and publication of the

[49] https://www.dariah.eu/

[50] https://www.clarin.eu/

controlled vocabularies via an online browser, using Skosmos,[51] in late 2024[52] (thus making the vocabularies FAIR based on the rules outlined in [8]). In addition, the use of CLEF by the KNOT project formed the basis of a use case for updates to the application [14], which resolved many of the technical issues we faced during development of the catalogue and introduced new functionality. This work is planned to continue in 2025 alongside ongoing alignment with the ICDP for ingestion of the data produced by KNOT into I.PaC (who published their cross-domain model for the infrastructure in late 2024, focused on many of the same key conceptual entities as KNOT-DM).[53] Alongside this ingestion, which should bring some wider visibility to the project and its outputs, we are also planning to reach out to all representatives of the projects catalogued to invite them to review their entry and share these results with their communities. Lastly, a final set of guidelines and recommendations for how the national infrastructure and Italian universities should consider the integration and valorization of scholarly digital objects which will also be published and shared with the ICDP.

References

[1] Borek, Luise, Quinn Dombrowski, Jody Perkins, and Christof Schöch. 2014. "Scholarly Primitives Revisited: Towards A Practical Taxonomy of Digital Humanities Research Activities and Objects." Zenodo. https://doi.org/10.5281/ZENODO.10866.

[2] Borek, Luise, Quinn Dombrowski, Jody Perkins, and Christof Schöch. 2016. "TaDiRAH: A Case Study in Pragmatic Classification." *Digital Humanities Quarterly* 10 (1). https://www.digitalhumanities.org/dhq/vol/10/1/000235/000235.html.

[3] Borgman, Christine L. 1999. "What Are Digital Libraries? Competing Visions." *Information Processing & Management* 35 (3): 227–43. https://doi.org/10.1016/s0306-4573(98)00059-4.

[4] Bruseker, George, Nicola Carboni, and Anaïs Guillem. 2017. "Cultural Heritage Data Management: The Role of Formal Ontology and CIDOC CRM." In *Heritage and Archaeology in the DigitalAge*, 93–131. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-65370-9_6

[5] Cameron, Fiona R. 2021. *The Future of Digital Data, Heritage and Curation: In a More-than-Human World*. London, England: Routledge.

[6] Carriero, Valentina Anita, Aldo Gangemi, Maria Letizia Mancinelli, Ludovica Marinucci, Andrea Giovanni Nuzzolese, Valentina Presutti, and Chiara Veninata. 2019. "ArCo: The Italian Cultural Heritage Knowledge Graph." In *Lecture Notes in Computer Science*, 36–52. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-30796-7_3

[51] https://skosmos.org/

[52] https://lab.dharc.unibo.it/skosmos/

[53] https://partecipa.gov.it/processes/IPaC

[7] Ciula, Arianna, Øyvind Eide, Cristina Marras, and Patrick Sahle. 2018. "Models and Modelling between Digital and Humanities. Remarks from a Multidisciplinary Perspective." GESIS - Leibniz Institute for the Social Sciences. https://doi.org/10.12759/HSR.43.2018.4.343-361.

[8] Cox, Simon J. D., Alejandra N. Gonzalez-Beltran, Barbara Magagna, and Maria-Cristina Marinescu. 2021. "Ten Simple Rules for Making a Vocabulary FAIR." *PLoS Computational Biology* 17 (6): e1009041. https://doi.org/10.1371/journal.pcbi.1009041.

[9] Daquino, Marilena, Mari Wigham, Enrico Daga, Lucia Giagnolini, and Francesca Tomasi. 2023. "CLEF. A Linked Open Data Native System for Crowdsourcing." Journal on Computing and Cultural Heritage. Association for Computing Machinery (ACM). https://doi.org/10.1145/3594721.

[10] Donato, Fabio, and Enrica Gilli. 2011. "Un Approccio 'Multi-Scala' per La Gestione Del Patrimonio Culturale Italiano." https://doi.org/10.13138/2039-2362/118.

[11] Fafinski, Mateusz. 2022. "Facsimile Narratives: Researching the Past in the Age of Digital Reproduction." *Digital Scholarship in the Humanities* 37 (1): 94–108. https://doi.org/10.1093/llc/fqab017.

[12] Franzini, Greta, Melissa Terras, and Simon Mahony. 2019. "Digital Editions of Text: Surveying User Requirements in the Digital Humanities." *Journal on Computing and Cultural Heritage* 12 (1): 1–23. https://doi.org/10.1145/3230671.

[13] Frosini, Luca, Alessia Bardi, Paolo Manghi and Pasquale Pagano. 2018. "An Aggregation Framework for Digital Humanities Infrastructures: The PARTHENOS Experience." *SCIRES-IT* 8 (1). https://doi.org/10.2423/i22394303v8n1p33.

[14] Giacomini, Sebastiano, Marilena Daquino, Francesca Tomasi, and Laurent Antoine Fintoni. 2025. "CLEF 2.0. Solutions for Native Linked Data Cataloguing of Italian Digital Cultural Heritage." *JLIS.It* 16 (1): 108–26. https://doi.org/10.36253/jlis.it-611.

[15] Gibbs, Fred, and Trevor Owens. 2012. "Building Better Digital Humanities Tools: Toward Broader Audiences and User-Centered Designs." *Digital Humanities Quarterly* 6 (2). https://www.digitalhumanities.org/dhq/vol/6/2/000136/000136.html.

[16] Grant, Kaitlyn, Quinn Dombrowski, Kamal Ranaweera, Omar Rodriguez-Arenas, Stéfan Sinclair, and Geoffrey Rockwell. 2020. "Absorbing DiRT: Tool Directories in the Digital Age." Digital Studies 10 (1). https://doi.org/10.16995/dscn.325.

[17] Guccio, Calogero, Marco Ferdinando Martorana, Isidoro Mazza, and Ilde Rizzo. 2016. "Technology and Public Access to Cultural Heritage: The Italian Experience on ICT for Public Historical Archives." In *Cultural Heritage in a Changing World*, 55–75. Cham: Springer International Publishing.

[18] Guiliano, Jennifer, and Laura Estill. 2023. "What Gets Categorized Counts: Controlled Vocabularies, Digital Affordances, and the International Digital Humanities Conference." *Digital Scholarship in the Humanities*. https://doi.org/10.1093/llc/fqac091.

[19] Hall, Crystal. 2019. "Digital Humanities and Italian Studies: Intersections and Oppositions." *Italian Culture* 37 (2): 97–115. https://doi.org/10.1080/01614622.2019.1717754.

[20] Harping, Patricia. 2010. *Introduction to Controlled Vocabularies: Terminology for Art, Architecture and Other Cultural Works*. Santa Monica, CA: Getty Research Institute.

[21] Istituto centrale per la digitalizzazione del patrimonio culturale – Digital Library. 2023. "Piano nazionale di digitalizzazione del patrimonio culturale - Parole Chiave." Docs Italia. 2023. https://docs.italia.it/italia/icdp/icdp-pnd-docs/it/v1.1-febbraio-2023/parole-chiave.html.

[22] Jansen, Ludger. 2019. "Ontologies for the Digital Humanities: Learning from the Life Sciences?" *Joint Ontology Workshops*. https://ceur-ws.org/Vol-2518/paper-WODHSA5.pdf.

[23] Middle, Sarah. 2023. "A Documentation Checklist for (Linked) Humanities Data." *International Journal of Digital Humanities*. https://doi.org/10.1007/s42803-023-00072-z.

[24] Moraitou, Efthymia, John Aliprantis, Yannis Christodoulou, Alexandros Teneketzis, and George Caridakis. 2019. "Semantic Bridging of Cultural Heritage Disciplines and Tasks." *Heritage* 2 (1): 611–30. https://doi.org/10.3390/heritage2010040.

[25] Mozzoni, Isabella, Simone Fanelli, and Chiara Carolina Donelli. 2018. "Italian University Collections: Managing the Artistic Heritage of the University's Ivory Tower." *European Journal of Cultural Management and Policy* 8. https://doi.org/10.3389/ejcmp.2023.v8iss1-article-3.

[26] Perkins, Jody, Quinn Dombrowski, Luise Borek, and Christof Schöch. 2014. "Project Report: Building Bridges to the Future of a Distributed Network: From DiRT Categories to TaDiRAH, a Methods Taxonomy for Digital Humanities," 181–83. https://dl.acm.org/doi/10.5555/2771234.2771256.

[27] Peroni, Silvio. n.d. "A Simplified Agile Methodology for Ontology Development." W3id.org. Accessed January 16, 2025. https://w3id.org/people/essepuntato/papers/samod-owled2016.html.

[28] Sabba, Fiammetta, Francesca Tomasi, Silvia Tripodi, and Laurent Antoine Fintoni. 2023. "L'Università Di Bologna a Supporto Dell'ecosistema Digitale Della Cultura Presso Il Ministero Della Cultura." In *Visioni Future: Next Generation Library - Vol.1*, 45–50. Editrice Bibliografica.

[29] Salvatori, Enrica. 2020. "Digital Public History inside and Outside the Box." *Magazén*, no. 2. https://doi.org/10.30687/mag/2724-3923/2020/02/003.

[30] Tammaro, Anna Maria. 2014. "Evaluation of Digital Humanities: An Interdisciplinary Approach." In *Communications in Computer and Information Science*, 136–46. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-54347-0_15.

[31] Tomasi, Francesca. 2018. "Modelling in the Digital Humanities: Conceptual Data Models and Knowledge Organization in the Cultural Heritage Domain." https://doi.org/10.12759/HSR.SUPPL.31.2018.170-179.

[32] "UNESCO Charter on the Preservation of the Digital Heritage." 2003. In *Intergovernmental Council for the Information for All Programme, 2nd, Paris, 2003*. https://unesdoc.unesco.org/ark:/48223/pf0000229034.

[33] Unsworth, John. 2000. "Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?" https://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html.

[34] Vlase, Ionela, and Tuuli Lähdesmäki. 2023. "A Bibliometric Analysis of Cultural Heritage Research in the Humanities: The Web of Science as a Tool of Knowledge Management." *Humanities & Social Sciences Communications* 10 (1): 84. https://doi.org/10.1057/s41599-023-01582-5.

[35] Warwick, C., M. Terras, P. Huntington, and N. Pappa. 2007. "If You Build It Will They Come? The LAIRAH Study: Quantifying the Use of Online Resources in the Arts and Humanities through Statistical Analysis of User Log Data." *Literary and Linguistic Computing* 23 (1): 85–102. https://doi.org/10.1093/llc/fqm045.

[36] Wilkinson, Mark D., Michel Dumontier, I. Jsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. 2016. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3: 160018. https://doi.org/10.1038/sdata.2016.18.