Library Modelling Report

by Maria Ingold

INTRODUCTION

Local Community-funded Library (LoCL) uses Civica's Spydus Library Management System (LMS) to manage adult and children's fiction and non-fiction books (Civica, N.D.). Due to efficiency concerns, LoCL is developing an Artificial Intelligence (AI)-powered search engine to better serve its patrons (Rout, 2018). This report presents LoCLOnt, a prototype ontology representing LoCL's book concepts and novel relationships to enable AI semantic search (Uschold & Grüninger, 1996).

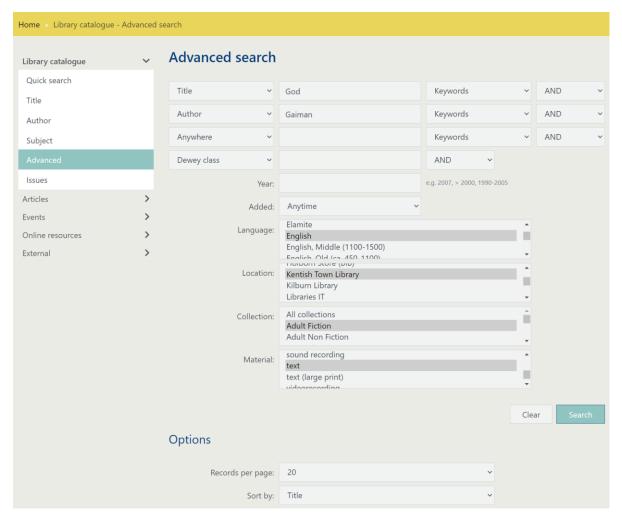


FIGURE 1 | Spydus

BUSINESS CONTEXT

An Online Public Access Catalogue (OPAC), Spydus uses MAchine-Readable Cataloging (MARC) (*MARC Standards*, N.D.). MARC 21 provides bibliographic metadata structure, while OPAC searches author, title, subject, identifiers, and keywords with Boolean logic (Husain & Ansari, 2006).

OPACs require an exact match, resulting in low precision and recall (Rout, 2018). For example, searching for 'God' and 'Gaiman' misses *American Gods*, and returns religious texts if only 'God' is searched (Figure 1).

'Genrefication', like Netflix's navigation, improves book discovery (Hider et al., 2021). Coupling genre with audience-appropriate search empowers patrons to efficiently find books aligning with their interests and reading level (Moeller & Becnel, 2019; Journey to Kidlit, 2021).

APPROACH JUSTIFICATION

OPACs provide syntactic search—without context (Rout, 2018). Conversely, ontologies enable semantics—concepts and relationships—like in Grüninger and Fox's (1995) competency questions used in LoCLOnt's top-down deployment of Noy and McGuinness' (2001) seven-step methodology.

1: Determine Domain and Range

The domain, searching local library books, uses competency questions to informally define the library and patron's scope:

Which book titles...

- ...have 'Science Fiction' as their genre?
- ...are classified under both 'Historical' and 'Fantasy'?
- ...have 'God' in the title?
- ...are authored by 'Gaiman'?
- ...are suitable for Middle Grade children?
- ...are suitable for Middle Grade children who like Fantasy?
- ...were published after a date (new releases)?

Formally, in First Order Logic (FOL):

- 1. $\exists x, y, t (Book(x) \land ScienceFiction(y) \land HasGenre(x, y) \rightarrow Title(x, t))$
- 2. $\exists x, y_1, y_2, t \big(Book(x) \land Historical(y_1) \land Fantasy(y_2) \land HasGenre(x, y_1) \land$ $HasGenre(x, y_2) \rightarrow Title(x, t) \big)$
- 3. $\exists x, t (Book(x) \land Contains(Title(x), 'God') \rightarrow Title(x, t))$
- 4. $\exists x, t (Book(x) \land Author(x, 'Gaiman') \rightarrow Title(x, t))$
- 5. $\exists x, z, t (Book(x) \land MiddleGrade(z) \land IsSuitableFor(x, z) \rightarrow Title(x, t))$
- 6. $\exists x, y, z, t \big(Book(x) \land MiddleGrade(z) \land IsSuitableFor(x, z) \land Fantasy(y) \land HasGenre(x, y) \rightarrow Title(x, t) \big)$
- 7. $\exists x, t (Book(x) \land PublishedAfter(x, 'date') \rightarrow Title(x, t))$

2: Consider Reusing Existing Ontologies

MARC21 is not an ontology, and Semantic Publishing and Referencing (SPAR)

Ontologies lack library concepts (Kruk et al., 2005; Peroni & Shotton, 2018).

Furthermore, fiction genre gaps in Dewey Decimal Classification and the Library of Congress Genre/Form Terms (LCGFT), mean Amazon and some libraries use Book Industry Standards And Communications (BISAC); unfortunately, 'juvenile' is broadly classed as ages 1 to 12 (Hider & Spiller, 2020). With no suitable ontologies, LoCLOnt reuses OPAC search types and BISAC genres.

ONTOLOGY DESIGN RATIONALE

LocLOnt uses Protégé 5.5.0 Web Ontology Language (OWL) editor with SPARQL Protocol and Resource Description Framework Query Language (SPARQL) (DeBellis, 2021).

3: Enumerate Important Terms

Key distinct concepts include book, audience, genre, and searchability.

4: Define Classes and Hierarchy

'Book', 'BookAudience', 'BookGenre', 'Search' and 'SearchType' are disjoint top-level classes (Figure 2) (Uschold & Grüninger, 1996). While 'Book-' classes are static, 'Search' represents action using enumerated class 'SearchType' (Figure 3).

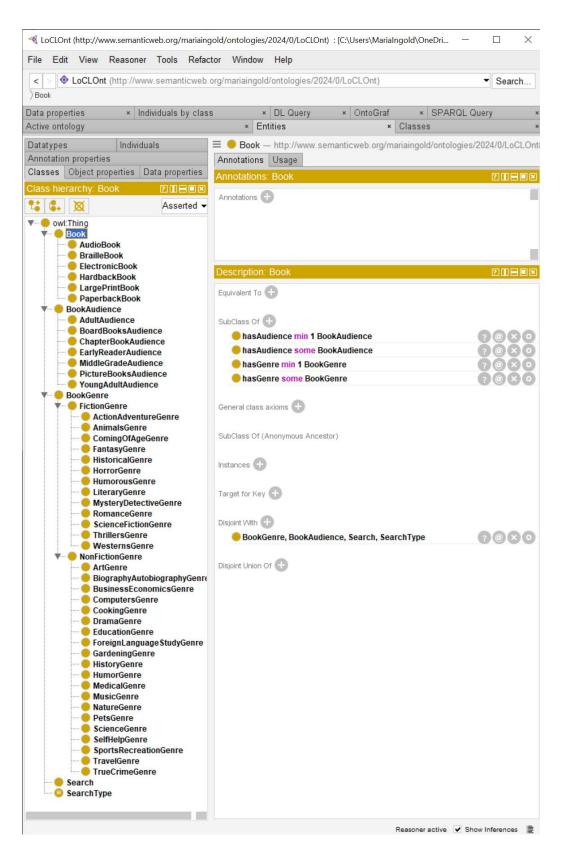


FIGURE 2 | Class hierarchy

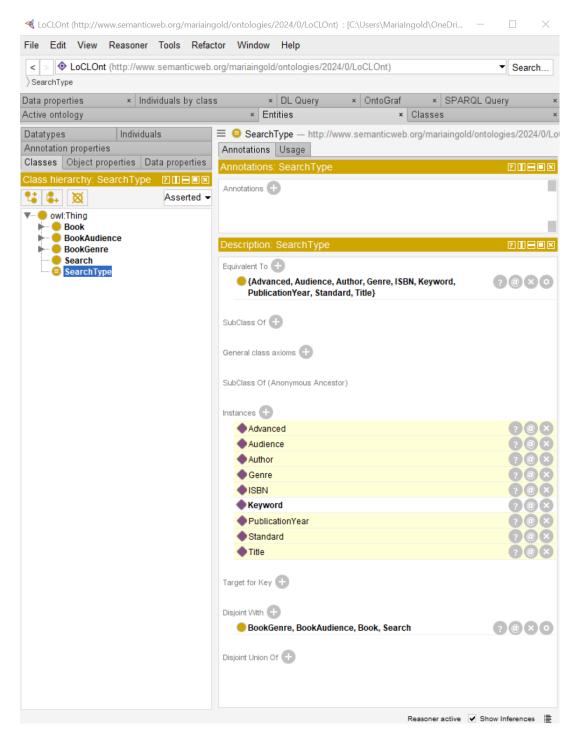


FIGURE 3 | Enumerated class (reasoner infers yellow)

5: Define Class Properties

Class properties denote relationships: Object (two instances), Data (instance and datatype), and Annotation (any entity) (DeBellis, 2021).

Book 'has' genre and audience Object Properties (Table 1, Figures 4, 5). No property characteristics apply (not Functional because a book can have more than one genre or audience). Enumerated type searchType is not Functional as the application layer combines.

TABLE 1 | Object properties

Object Property	Inverse	Domain	Range
hasGenre	isGenreOf	Book	BookGenre
hasAudience	isAudienceOf	Book	BookAudience
searchType	N/A	Search	SearchType

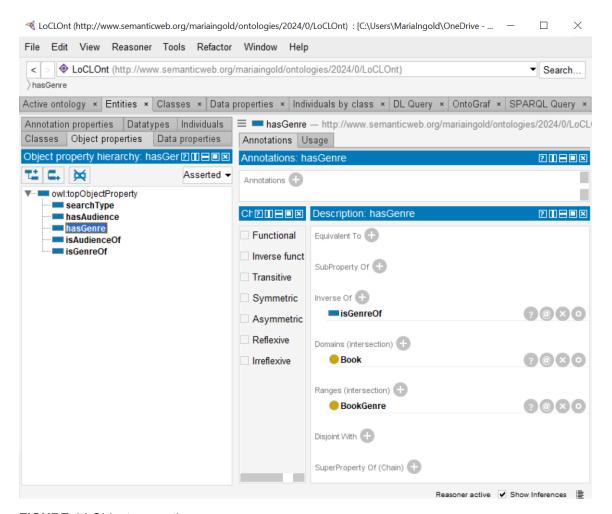


FIGURE 4 | Object properties

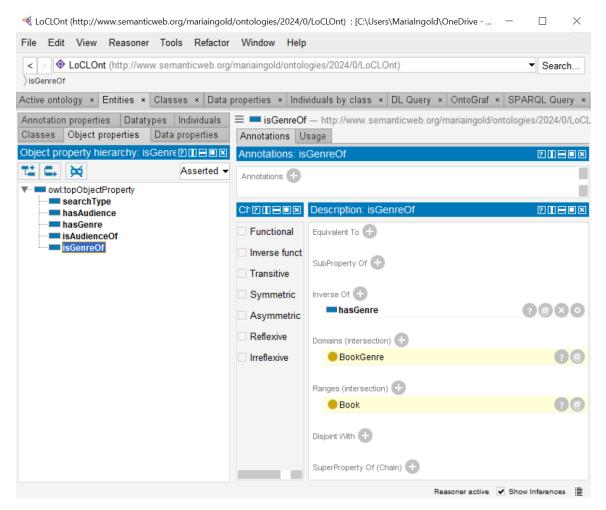


FIGURE 5 | Inverse

Given the simplicity required, author is a data property (Figure 6), whereas a BookAuthor class could have genres and audiences. Because formats are different instances, most properties are Functional, but multiple authors can exist (Table 2).

TABLE 2 | Data properties

Data Property	Functional	Domain	Range
title	Υ	Book	xsd:string
author	N	Book	xsd:string
publicationYear	Υ	Book	xsd.integer
isbn	Υ	Book	xsd:string
searchDate	Υ	Search	xsd:dateTime

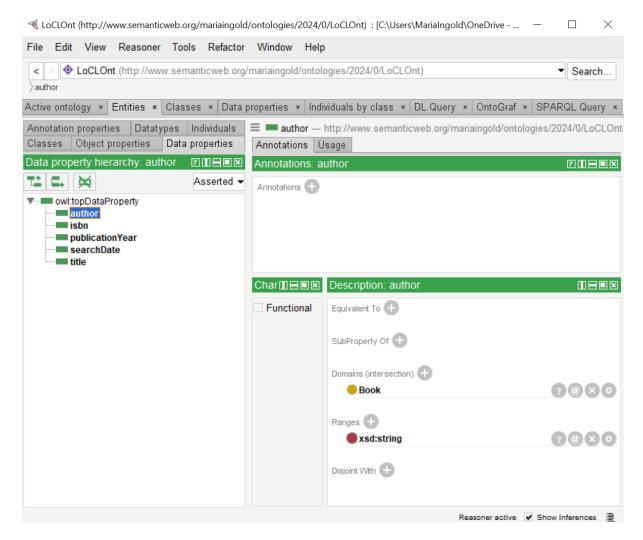


FIGURE 6 | Data properties

6: Define Property Restrictions

Class instance properties are restricted by quantifier (existential or universal), cardinality (number), or value (specific) (DeBellis, 2021). FOL \exists requires 'Some (existential)' quantifier, with cardinality 'min 1' ensuring books have at least one audience and genre. Universal is too restrictive and value is inapplicable.

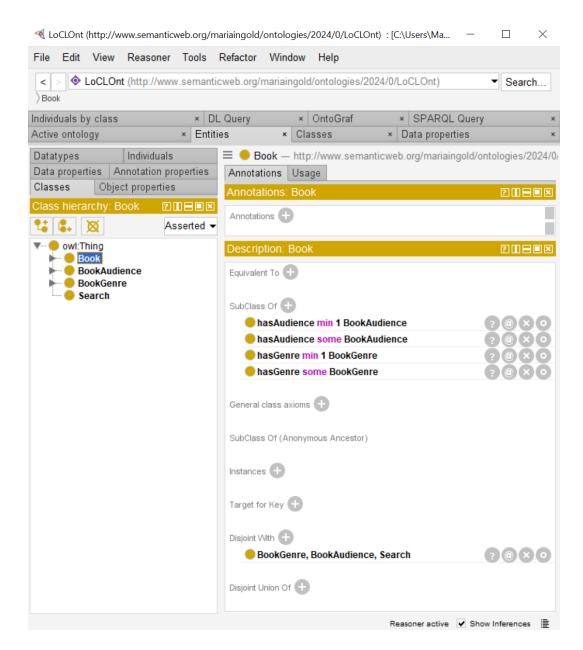
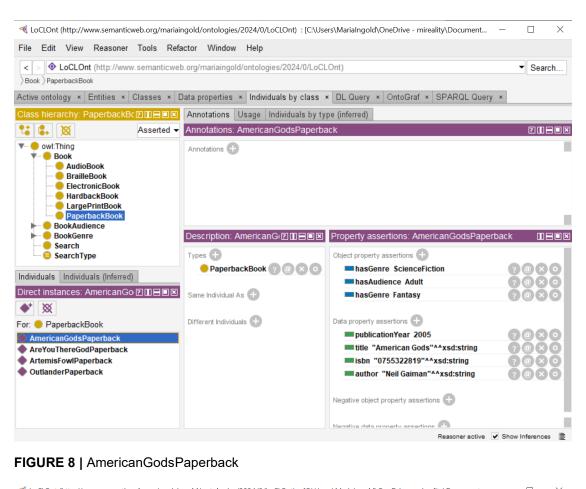


FIGURE 7 | Property restrictions

7: Create Instances

BookGenre and BookAudience instances were created and four PaperbackBook instances populated (Figures 8-11).



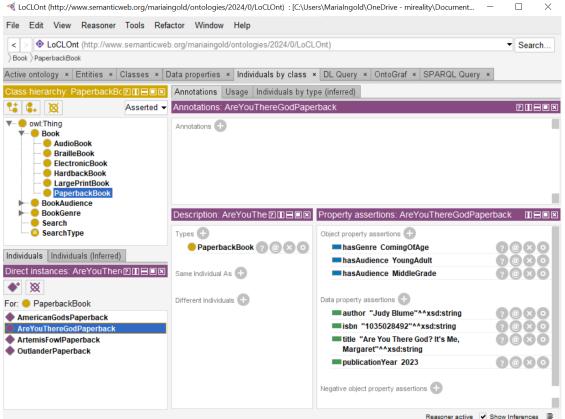


FIGURE 9 | AreYouThereGodPaperback

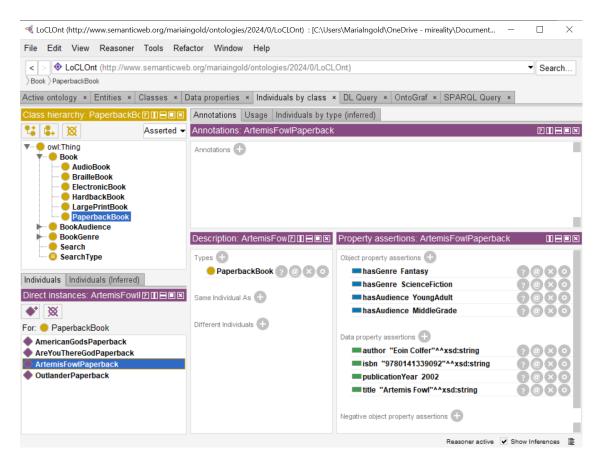


FIGURE 10 | ArtemisFowlPaperback

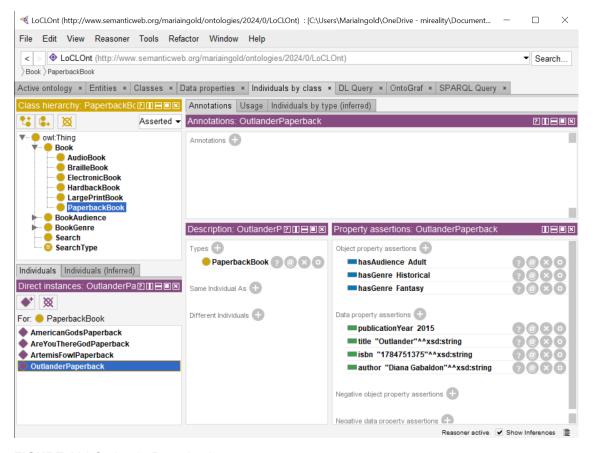


FIGURE 11 | OutlanderPaperback

OUTPUT ANALYSIS

The Pellet reasoner internally evaluated consistency, accuracy, and inferences.

SPARQL validated LoCLOnt's FOL-compliant competency questions, demonstrating effective simple and combined searches.

Title or Author Search (Figures 14,15)

Improving accuracy over Spydus, a regex search for 'God' returns 'American Gods' and 'Are You There God...'. Similarly, 'Gaiman' finds author 'Neil Gaiman'. To return titles, not word subsets, case insensitivity is not used. However, for production, case and exact matches should be addressed.

Genre Search (Figures 12, 13, 17)

Unavailable in Spydus, genrefication enables quickly finding books that match interests. Further genres can be added.

Audience and Genre Search (Figures 16, 17)

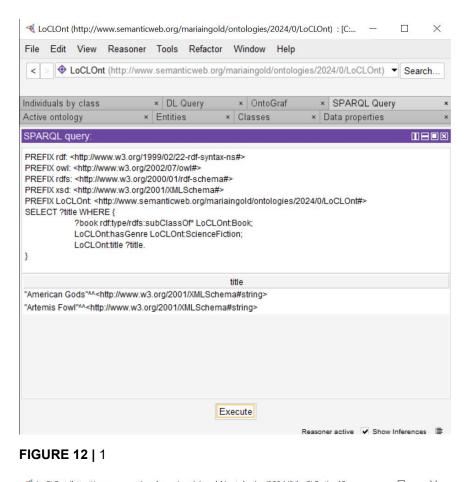
A reading-level audience combined with genre helps parents and children efficiently find interesting age-appropriate books.

New Releases (Figure 18)

Netflix-like new releases encourages discovery. While an integer year made search simple, a full date would enable more refinement.

Limitations

SPARQL Query does not support updates or now() so neither Search instance nor searchDate can be created. A fully scoped deliverable, testing, and validation trial must be completed prior to deployment.



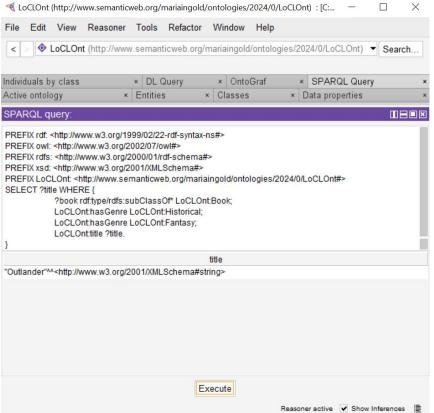
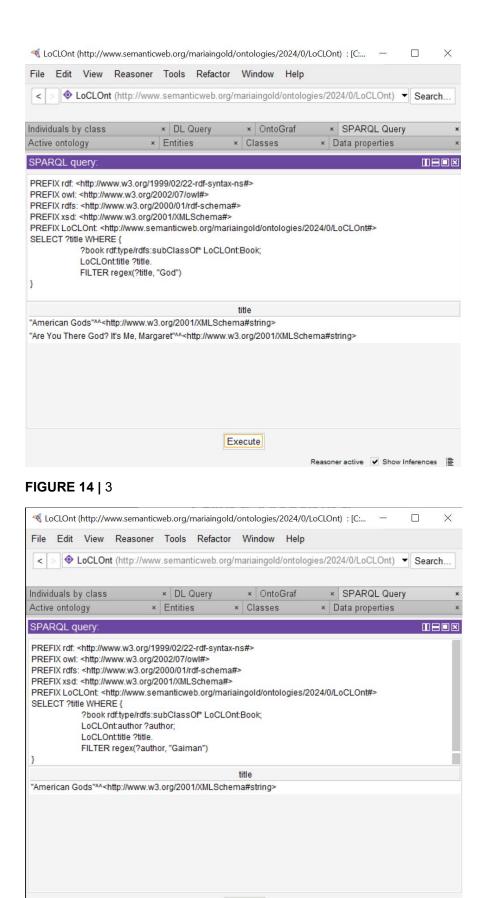


FIGURE 13 | 2



Execute

Reasoner active Show Inferences

FIGURE 15 | 4

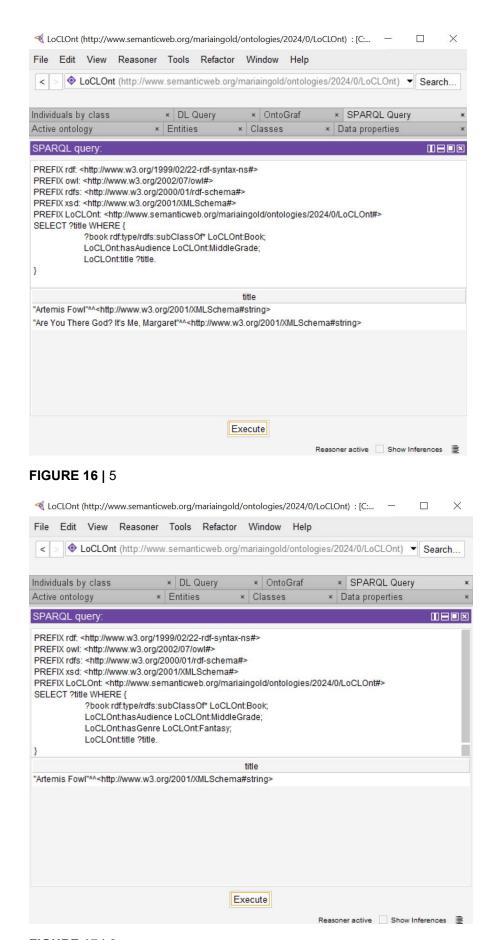


FIGURE 17 | 6

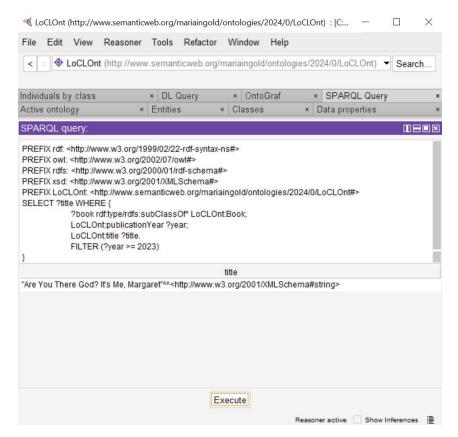


FIGURE 18 | 7

APPLICATION

Integration

LoCLOnt maps to Spydus' standard OPAC search types and can be easily extended.

Validation

While OPAC search format simplifies adoption, Netflix-style interfaces could improve efficiency. Output analysis demonstrates relevant simple or advanced searches with age-related interests and recency. Ultimately, validation by library management and patrons is required.

Considerations

LoCLOnt is reusable with OPAC, and prefixing 'Book-' establishes relationship and enables extension, such as a Film class using FilmRating not BookAudience.

Similarly, FilmGenre could include DocumentaryGenre not NonfictionGenre.

However, going global requires changes like searchTime using xsd:dateTimeStamp. Furthermore, simplifications, like ISBN, require addressing ISBN-10 and ISBN-13 and not mandating use for digital books.

Future

The ontology enables Al-driven expansion, like recommendations or personalisation.

CONCLUSION

Surpassing Spydus for precision and relevance, LoCLOnt uses semantic search to help patrons efficiently discover relevant books. Building on Spydus' OPAC structure, competency questions evaluated age-appropriate, genre, and date searches.

LoCLOnt demonstrates effective knowledge representation and reasoning in improving patron services and enabling further Al-based search, like films or recommendations.

REFERENCES

Civica (N.D.) Spydus Integrated Library Management System Software. Available from: https://www.civica.com/en-gb/product-pages/spydus-integrated-library-management-solution/ [Accessed 23 January 2024].

DeBellis, M. (2021) A Practical Guide to Building OWL Ontologies Using Protégé 5.5 and Plugins. Available from:

https://www.researchgate.net/publication/351037551_A_Practical_Guide_to_Building _OWL_Ontologies_Using_Protege_55_and_Plugins [Accessed 15 December 2023].

Grüninger, M. & Fox, M.S. (1995) *Methodology for the Design and Evaluation of Ontologies*. Available from: https://www.researchgate.net/publication/2288533.

Hider, P. & Spiller, B. (2020) Fiction Genres in Bookstores and Libraries: A Comparison of Commercial and Professional Classifications, *Cataloging & Classification Quarterly* 58(8): 664–682. DOI: https://doi.org/10.1080/01639374.2020.1850590.

Hider, P., White, H. & Barlow, P. (2021) Film Genres through Different Lenses: Mapping Commonly Used Film Vocabularies onto the Library of Congress Genre/Form Terms, *Library Trends* 69(3): 630–645. DOI: https://doi.org/10.1353/LIB.2021.0007.

Husain, R. & Alam Ansari, M. (2006) From Card Catalogue to Web OPACs, *DESIDOC Bulletin of Information Technology* 26(2): 41–47. Available from: https://publications.drdo.gov.in/ojs/index.php/djlit/article/view/3679/2091 [Accessed 23 January 2024].

Journey to Kidlit (2021) *The Age Levels for Children's Books You Should Know - Journey to Kidlit*. Available from: https://journeytokidlit.com/age-levels-for-childrens-books/ [Accessed 25 January 2024].

Kruk, S.R., Synak, M. & Zimmerman, K. (2005) MarcOnt - Integration Ontology for Bibliographic Description Formats, *International Conference on Dublin Core and Metadata Applications* 231–234. Available from:

https://dcpapers.dublincore.org/pubs/article/view/829 [Accessed 24 January 2024].

MARC Standards (N.D.) Network Development and MARC Standards Office, Library of Congress. Available from: https://www.loc.gov/marc/ [Accessed 23 January 2024].

Moeller, R.A., & Becnel, K.E. (2019) 'Why On Earth Would We Not Genrefy the Books?': A Study of Reader-Interest Classification In School Libraries. DOI: 10.5771/0943-7444-2019-3-199.

Noy, N.F. & McGuinness, D.L. (2001) Ontology Development 101: A Guide to Creating Your First Ontology. Available from:

https://protege.stanford.edu/publications/ontology_development/ontology101.pdf [Accessed 8 January 2024].

Peroni, S. & Shotton, D. (2018) The SPAR Ontologies. DOI: https://doi.org/10.1007/978-3-030-00668-6_8.

Rout, R. (2018) Enhancement of subject searching in OPAC using ontology, *Espacios* 39(44). Available from:

https://es.revistaespacios.com/a18v39n44/a18v39n44p15.pdf [Accessed 29 January 2024]

Uschold, M. & Grüninger, M. (1996) Ontologies: Principles, methods and applications, *Knowledge Engineering Review* 11(2). DOI: https://doi.org/doi:10.1017/S0269888900007797.

APPENDIX

The appendix includes additional research detail (via screenshots) that was compressed due to the limited wordcount. These screenshots demonstrate a portion of the research undertaken for this paper and exemplify pictorially what has been discussed. These are available at the links shown as of 10/02/2024, but are not all individually referenced and are purely supplementary information. However, this paper links to relevant references for the research area described by each appendix section.

Spydus: Camden Library Example

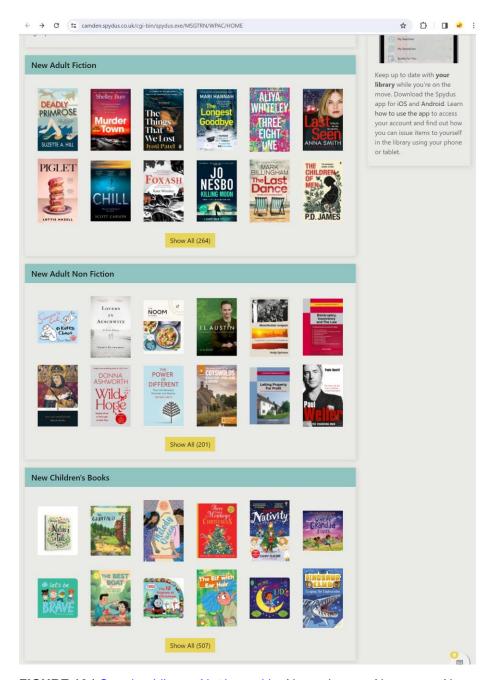


FIGURE 19 | Camden Library. Not logged in. New releases. No genres. No age ranges for children.

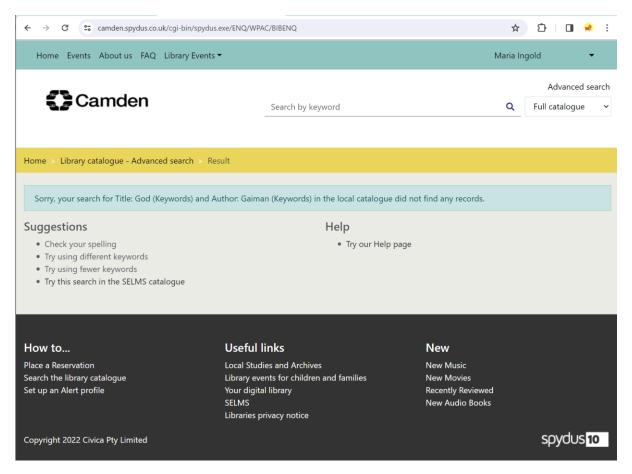


FIGURE 20 | Search for 'God' and 'Gaiman' fails to find 'American Gods'

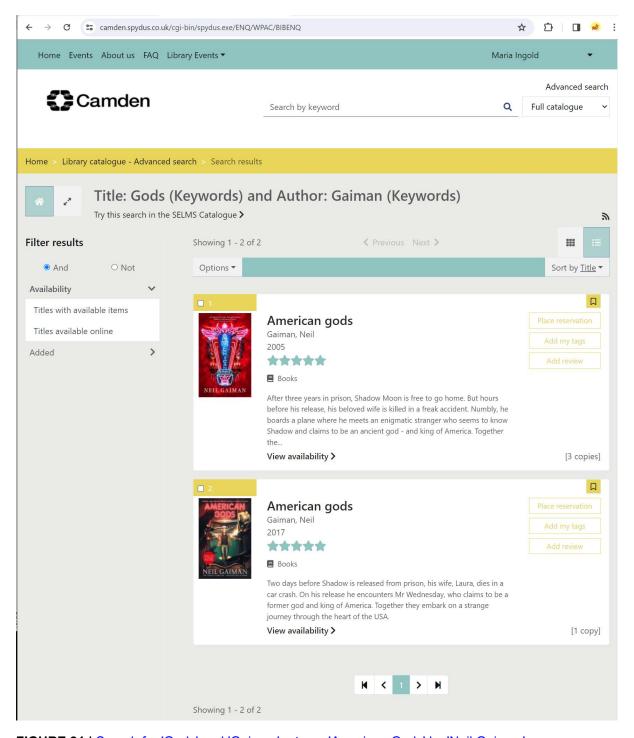


FIGURE 21 | Search for 'Gods' and 'Gaiman' returns 'American Gods' by 'Neil Gaiman'

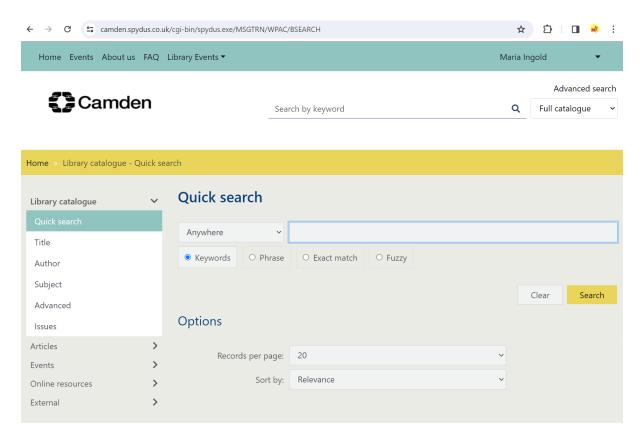


FIGURE 22 | Types of Search

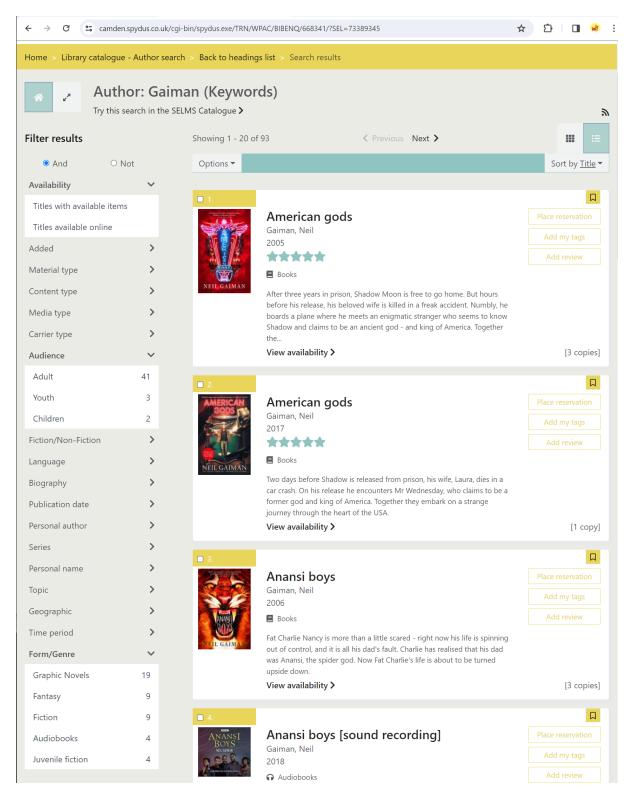


FIGURE 23 | Genre and audience only appears when search on author. Can't search on Genre or Audience.

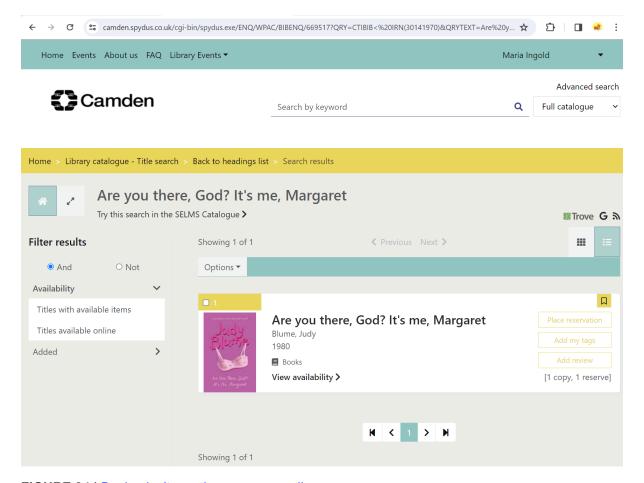


FIGURE 24 | Books don't mention genre or audience

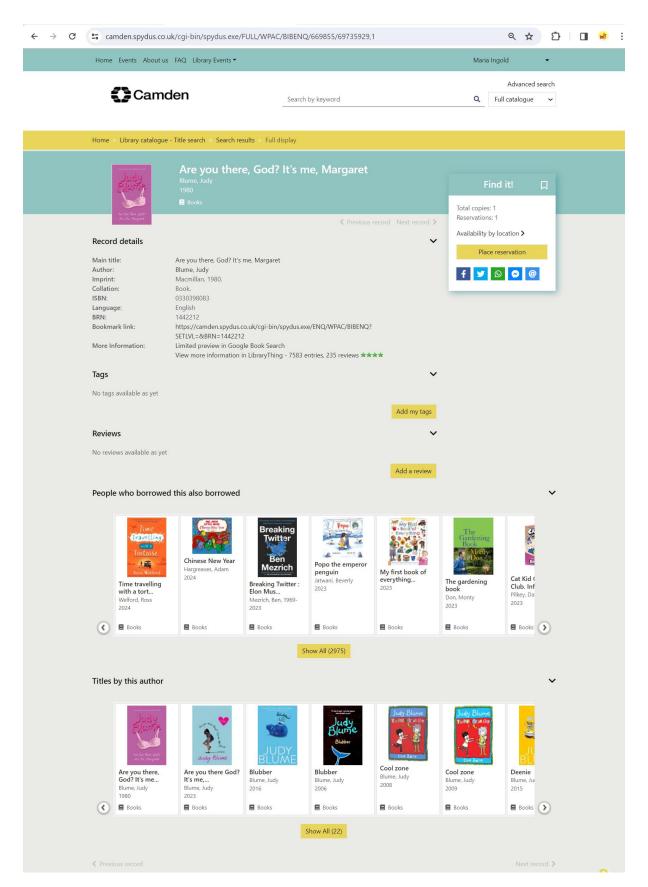


FIGURE 25 | Books don't list genre or audience. 'People who borrowed...' is not age-appropriate.

Gen Z and Libraries



Books

'Reading is so sexy': gen Z turns to physical books and libraries

Book sales boom as readers escape the 'oversaturation and noise of the wild west digital landscape'



Chloe Mac Donnell









■ The model Kaia Gerber has launched a book club, Library Science. Photograph: Instagram

hey have killed skinny jeans and continue to shame millennials for having side partings in their hair. They think using the crying tears emoji to express laughter is embarrassing. But now comes a surprising gen Z plot twist. One habit that those born between 1997 and 2012 are keen to endorse is reading - and it's physical books rather than digital that they are thumbing.

This week the 22-year-old model Kaia Gerber launched her own book club, Library Science. Gerber, who this month appears on the cover of British Vogue alongside her supermodel mum, Cindy Crawford, describes it as "a platform for sharing books, featuring new writers, hosting conversations with artists we admire – and continuing to build a community of people who are as excited about literature as I am".



Ellen E Jones: Hollywood, race and the power of storytelling

Thursday 15 February, 8pm-9pm GMT

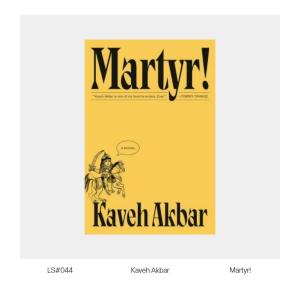
"Books have always been the great love of my life," she added. "Reading is so sexy."

Gerber isn't alone. Last year in the UK 669m physical books were sold, the highest overall level ever recorded. Research from Nielsen BookData highlights that it is print books that gen Z favour, accounting for 80% of purchases from November 2021 to 2022. Libraries are also reporting an uptick in gen Z users who favour their quiet over noisy coffee shops. In the UK in-person visits are up 71%.

FIGURE 26 | Gen Z is reading more, including at libraries

BOOKS CHANNEL ABOUT

LIBRARY SCIENCE







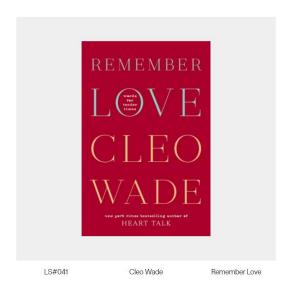






FIGURE 27 | Gen Z's Kaia Gerber's Library Science looks more like Netflix than Spydus

Netflix

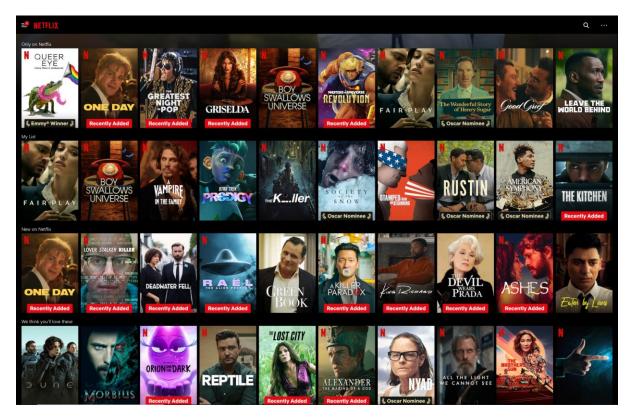
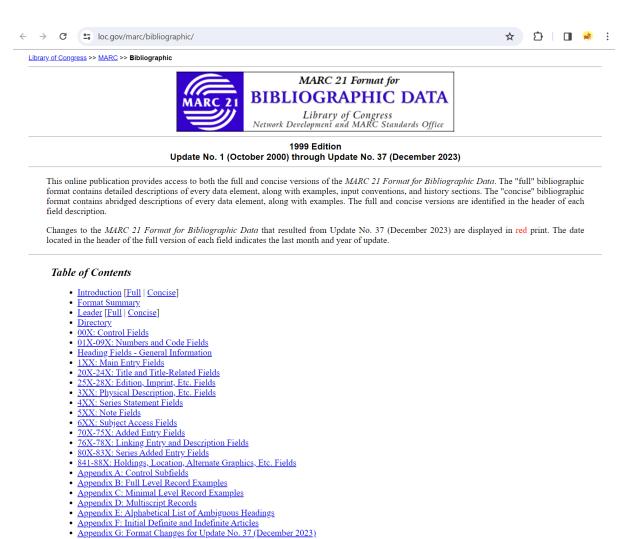


FIGURE 28 | Netflix. With My List (explicitly added), New on Netflix, We think you'll love these (personalised recommendations).



FIGURE 29 | Netflix sign in - still not fully age-appropriate

MARC 21



 $\frac{\text{Library of Congress}}{(12/07/23)} >> \underline{\text{MARC}} >> \underline{\text{Bibliographic}}$

Contact Us

FIGURE 30 | MARC 21 Format for Bibliographic Data

Appendix H: Local Data Elements
 Appendix I: Organization Code Sources
 Appendix J: Data Provenance Subfields







<u>Library of Congress</u> >> <u>MARC</u> >> <u>Bibliographic</u> >> **Introduction**

Introduction

MARC 21 Bibliographic - Concise

2007

MARC 21 Format for Bibliographic Data is designed to be a carrier for bibliographic information about printed and manuscript textual materials, computer files, maps, music, continuing resources, visual materials, and mixed materials. Bibliographic data commonly includes titles, names, subjects, notes, publication data, and information about the physical description of an item. The bibliographic format contains data elements for the following types of material:

- · Books (BK) used for printed, electronic, manuscript, and microform textual material that is monographic in nature.
- Continuing resources (CR) used for printed, electronic, manuscript, and microform textual material that is issued in parts with a recurring pattern of publication (e.g., periodicals, newspapers, yearbooks). (NOTE: Prior to 2002, Continuing resources (CR) were referred to as Serials (SE)).
 Computer files (CF) used for computer software, numeric data, computer-oriented multimedia, online systems or services. Other classes of electronic
- rces are coded for their most significant aspect. Material may be monographic or serial in nature.
- Maps (MP) used for all types of printed, electronic, manuscript, and microform cartographic materials, including atlases, sheet maps, and globes. Material may be monographic or serial in nature.
- Music (MU) used for printed, electronic, manuscript, and microform music, as well as musical sound recordings, and non-musical sound recordings. Material may be monographic or serial in nature.
- · Visual materials (VM) used for projected media, non-projected media, two-dimensional graphics, three-dimensional artifacts or naturally occurring
- objects, and kits. Material may be monographic or serial in nature.

 Mixed materials (MX) used primarily for archival and manuscript collections of a mixture of forms of material. Material may be monographic or serial in nature. (NOTE: Prior to 1994, Mixed materials (MX) were referred to as Archival and manuscript material (AM)).

Kinds of Bibliographic Records

MARC bibliographic records are distinguished from all other types of MARC records by specific codes in Leader/06 (Type of record) which identifies the following bibliographic record types

Language material Nonmusical sound recording Manuscript language material Musical sound recording Computer file Projected medium

Cartographic material Two-dimensional nonprojectable graphic Manuscript cartographic material Three-dimensional artifact or natural objects

Notated music Kit

Manuscript music Mixed material

FIGURE 31 | MARC 21 Types. This fits into Books (BK).

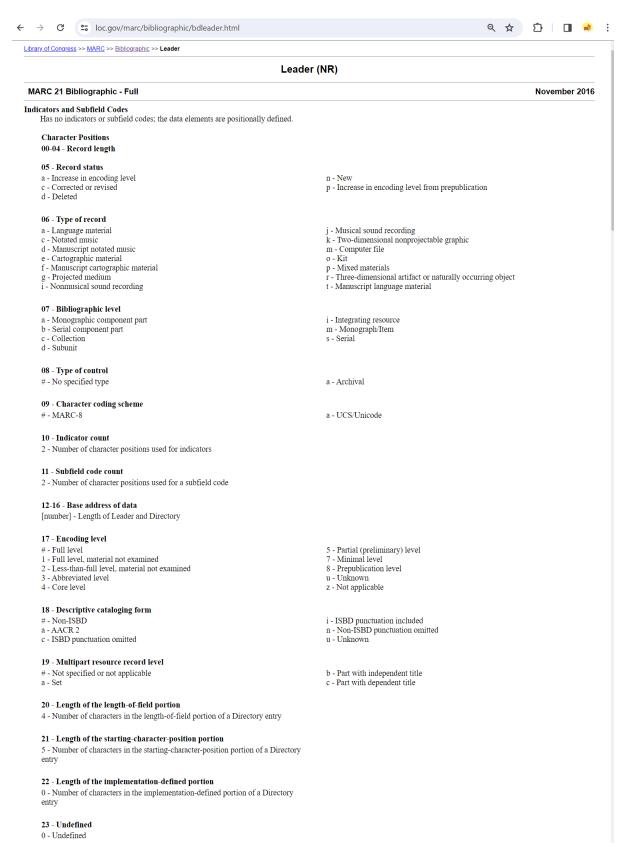


FIGURE 32 | MARC 21 Leader

MARC 21 Bibliographic Format

Full Level Record - Book

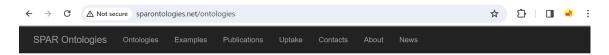
Network Development and MARC Standards Office Library of Congress

This example can be identified as a record for printed language material by code a in Leader/06, and further as a monograph by code m in Leader/07. The record contains the basic bibliographic information fields (fields 100, 245, 260, and 300), as well as additional fields (e.g., fields 020 (International Standard Book Number), 050 (Library of Congress Call Number), 082 (Dewey Decimal Classification Number), 246 (Varying Form of Title), 500 (General Note), and 650 (Subject Added Entry Topical Term)).

LDR		*****nam	#22****#a#4500							
001		<control number=""></control>								
003		<control identifier="" number=""></control>								
005		19920331092212.7								
007		ta								
008		820305	s1991####	nyu####	######	#001#0#	eng##			
020	##	\$a0845348116 :\$c\$29.95 (£19.50 U.K.)								
020	##	\$a0845348205 (pbk.)								
040	##	Sa <organization code="">Sc<organization code=""></organization></organization>								
050	14	\$aPN1992.8.S4\$bT47 1991								
082	04	\$a791.45/75/0973 \$2 19								
100	1#	\$aTerrace, Vincent,\$d1948-								
245	10	SaFifty years of television: Sba guide to series and pilots, 1937-1988 /ScVincent Terrace.								
246	1#	\$a50 years of television								
260	##	SaNew York: SbCornwall Books, Scc1991.								
300	##	\$a864 p. ;\$c24 cm.								
500	##	SaI ncludes	index.							
650	#0	Sa Television	n pilot programs\$z	United States\$vC	atalogs.					
650	#0	SaTelevision serialsSzUnited StatesSvCatalogs.								

FIGURE 33 | MARC 21 Book example

SPAR Ontologies

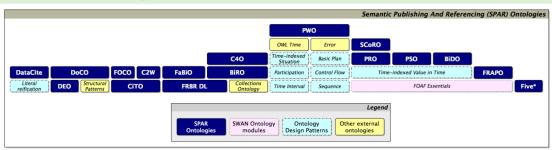


Semantic Publishing and Referencing Ontologies

In the past, several groups have proposed (Semantic Web) models, such as RDFS vocabularies and OWL ontologies, to describe particular aspects of the publishing domain. However, these models were mainly concerned with the description of the metadata of bibliographic resources (e.g., DC Terms, PRISM and BIBO). One of the first attempts to address the description of the whole publishing domain is the introduction of the Semantic Publishing and Referencing (SPAR) Ontologies. SPAR is a suite of orthogonal and complementary OWL 2 ontologies that enable all aspects of the publishing process to be described in machine-readable metadata statements, encoded using RDF.

Please use the following reference for citing the SPAR Ontologies suite in your works:

Peroni, S., Shotton, D. (2018). The SPAR Ontologies. In Proceedings of the 17th International Semantic Web Conference (ISWC 2018): 119-136. DOI: https://doi.org/10.1007/978-3-030-00668-6_8



While the current suite contains several ontologies (see the figure above), the original suite of SPAR ontologies comprises eight distinct modules:

- the FRBR-aligned Bibliographic Ontology (FaBiO) is an ontology for describing entities that are published or potentially publishable (e.g., journal articles, conference papers, books), and that contain or are referred to by bibliographic references;
- the Citation Typing Ontology (CiTO) is an ontology that enables characterization of the nature or type of citations, both factually and rhetorically;
- the Bibliographic Reference Ontology (BiRO) is an ontology meant to define bibliographic records, bibliographic references, and their compilation into bibliographic collections and bibliographic lists respectively.
- the Citation Counting and Context Characterisation Ontology (C4O) is an ontology that permits the number of in-text citations of a cited source to be recorded, together with their textual citation contexts, along with the number of citations a cited entity has received globally on a particular date;
- the Document Components Ontology (DoCO) in an ontology that provides a structured vocabulary written of document components, both structural (e.g., block, inline, paragraph, section, chapter) and rhetorical (e.g., introduction, discussion, acknowledgements, reference list, figure, appendix);
- the Publishing Status Ontology (PSO) is an ontology designed to characterise the publication status of documents at each stage of the publishing process (draft, submitted, under review, etc.);
- the Publishing Roles Ontology (PRO) is an ontology for the characterisation of the roles of agents people, corporate bodies and computational agents in the publication process. These agents can be, e.g. authors, editors, reviewers, publishers or librarians;
- the Publishing Workflow Ontology (PWO) is a simple ontology for describing the steps in the workflow associated with the publication of a document or other publication entity.

The main part of the aformentioned ontologies are based on other available and more general models. Two of those have been developed within the SPAR Ontologies:

- the Essential FRBR in OWL2 DL Ontology (FRBR) is an expression in OWL 2 DL of the basic concepts and relations described in the IFLA report on the Functional Requirements for Bibliographic Records (FRBR), also described in lan Davis's RDF vocabulary. It is imported by FaBiO and BiRO.
- the Discourse Elements Ontology (DEO) is an ontology that provides a structured vocabulary for rhetorical elements within documents (e.g., Introduction, Discussion, Acknowledgements, Reference List, Figures, Appendix). It is imported by DoCO.

The aforementioned ontologies form the original set of SPAR ontologies. However, this set has more recently been extended with other complementary ontologies that extend the coverage of the possible description of the publishing domain:

- the Scholarly Contributions and Roles Ontology (SCoRO) is an ontology based on PRO for describing the contributions that may be made, and the roles that may be held by a person with respect to a journal article or other publication (e.g. the role of article guarantor or illustrator);
- the Funding, Research Administration and Projects Ontology (FRAPO) is an ontology for describing the administrative information of research projects, e.g., grant applications, funding bodies, project partners, etc.;
- the DataCite Ontology (DataCite) is an ontology that enables the metadata properties of the DataCite Metadata Schema Specification (i.e., a list of metadata properties for the accurate and consistent identification of a resource for citation and retrieval purposes) to be described in RDF;
- the Bibliometric Data Ontology (BiDO) is a modular ontology that allows the description of numerical and categorical bibliometric data (e.g., journal impact factor, author h-index, categories describing research careers) in RDF;
- the Five Stars of Online Research Articles Ontology (FiveStars) is an ontology written in OWL 2 DL to enable characterization of the five attributes of an online journal article peer review, open access, enriched content, available datasets and machine-readable metadata.
- the FAIR* Reviews Ontology (FR) enables the description of reviews of scientific articles and other scholarly resources.

This website and all its content are licensed under a Creative Commons Attribution 4.0 International License excepting where stated differently.

Back to top

Essential FRBR in OWL2 DL Ontology (FRBR)

URL http://purl.org/spar/frbr (alternative at w3id.org)

DOI 10.25504/FAIRsharing.b34b43

Documentation http://purl.org/spar/frbr.html

Source http://purl.org/spar/frbr.xml (RDF/XML)

http://purl.org/spar/frbr.ttl (Turtle) http://purl.org/spar/frbr.nt (N-triples) http://purl.org/spar/frbr.json (JSON-LD)

Repository https://github.com/sparontologies/frbr

The Functional Requirements for Bibliographic Record (FRBR) is a general model, proposed by the International Federation of Library Association (IFLA), for describing documents and their evolution. It works for both physical and digital resources and it has proved to be very flexible and powerful. One of the most important aspects of FRBR is the fact that it is not associated with a particular metadata schema or implementation.

FRBR describes all documents from four different and correlated points of view: Work, Expression, Manifestation and Item; each of which is a FRBR Endeavour. These can be illustrated by considering of the book Alice's Adventures in Wonderland by Lewis Carroll as an example:

- Work. A FRBR Work is a high-level abstract Platonic concept of the essence of a distinct intellectual or artistic creation, for example the ideas in Lewis Carroll's head concerning Alice's Adventures in Wonderland, independent of any representation of these ideas in a particular form. A Work is realised through one or more Expressions;
- Expression. A FRBR Expression is the realisation of the intellectual or artistic content of a Work. Thus the original text of Alice's Adventures in Wonderland and its Italian translation Le Avventure di Alice nel Paese delle Meraviglie refer to different Expressions of the same Work. An Expression is embodied in one or more Manifestations:
- Manifestation. A FRBR Manifestation of a work defines its particular physical or electronic embodiment, for example, the particular format in which Alice's
 Adventures in Wonderland is stored: as a printed object or in HTML, represent two quite different Manifestations. In publishing, different manifestations of a journal
 article will all bear the same Digital Object Identifier (DOI), which identifies the Expression of the work, not its various Manifestations. However, a paperback and a
 hardback version of a book will bear different International Standard Book Numbers (ISBNs), since these identifiers are assigned at the Manifestation level. A
 Manifestation is exemplified by one or more Items;
- Item. A FRBR Item is a particular physical or electronic copy of Alice's Adventures in Wonderland that a person can own, for example the printed version of the book you have in your bookcase, or the Mobipocket format copy you have downloaded to read on your e-book device. All Items that are identical to one another for example books from the same printing, are exemplars of the same Manifestation.

Starting from Ian Davis's RDF vocabulary that expresses the basic concepts and relations described in the IFLA report on the Functional Requirements for Bibliographic Records (FRBR), we have created an expression in OWL 2 DL of FRBR in order to be used in other SPAR Ontologies that require this kind of descriptions, e.g., FaBiO and BiRO.

FIGURE 35 | Functional Requirements for Bibliographic Record (FRBR)

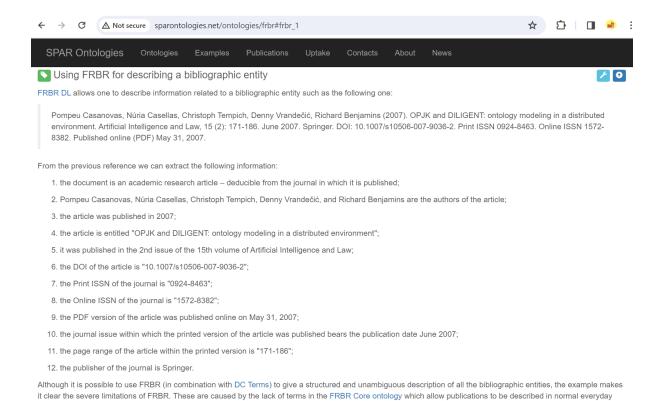
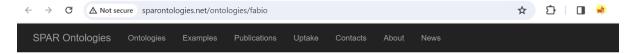


FIGURE 36 | FRBR - Limitations of FRBR - no everyday language

language. This issue is addressed by using FaBiO entities, has shown in one of its available examples.



FRBR-aligned Bibliographic Ontology (FaBiO)

URL http://purl.org/spar/fabio (alternative at w3id.org)

DOI 10.25504/FAIRsharing.2f3180

Documentation http://purl.org/spar/fabio.html

Source http://purl.org/spar/fabio.xml (RDF/XML)

http://purl.org/spar/fabio.ttl (Turtle) http://purl.org/spar/fabio.nt (N-triples) http://purl.org/spar/fabio.json (JSON-LD)

Repository https://github.com/sparontologies/fabio

Reference Peroni, S., Shotton, D. (2012). FaBiO and CiTO: ontologies for describing bibliographic resources and

citations. In Journal of Web Semantics: Science, Services and Agents on the World Wide Web, 17

(December 2012): 33-43. Amsterdam, The Netherlands: Elsevier.

https://doi.org/10.1016/j.websem.2012.08.001

 $Open \ Access \ at: \ http://speroni.web.cs.unibo.it/publications/peroni-2012-fabio-cito-ontologies.pdf$

FaBiO, the FRBR-aligned Bibliographic Ontology, is an ontology for recording and publishing on the Semantic Web descriptions of entities that are published or potentially publishable, and that contain or are referred to by bibliographic references, or entities used to define such bibliographic references. FaBiO entities are primarily textual publications such as books, magazines, newspapers and journals, and items of their content such as poems, conference papers and editorials. However, they also include blogs, web pages, datasets, computer algorithms, experimental protocols, formal specifications and vocabularies, legal records, governmental papers, technical and commercial reports and similar publications, and also anthologies, catalogues and similar collections.

FaBiO already imports several entities from existing standards for bibliographic entity descriptions, i.e., FRBR, DC Terms, PRISM and SKOS. In addition, FaBiO has been developed so to limit any restriction to its classes as well as the domains and ranges of its properties. This flexibility has the great advantage of allowing FaBiO to be used together with other models.

In particular, FaBiO classes are structured according to the FRBR schema of *Works, Expressions, Manifestations* and *Items*. The following Graffoo diagram shows additional properties that have been added to extends the FRBR data model by linking Works and Manifestations (fabio:hasManifestation and fabio:isManifestation0f), Works and Items (fabio:hasPortrayal and fabio:isPortrayedBy), and Expressions and Items (fabio:hasRepresentation and fabio:isRepresentedBy).

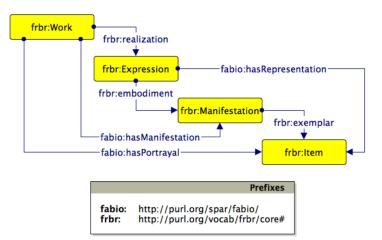


FIGURE 37 | FRBR-aligned Bibliographic Ontology (FaBiO) - improves on FRBR with everyday speech but not fully OWL 2 Description Logic (DL) complaint and does not address libraries.

Dewey Decimal Classification

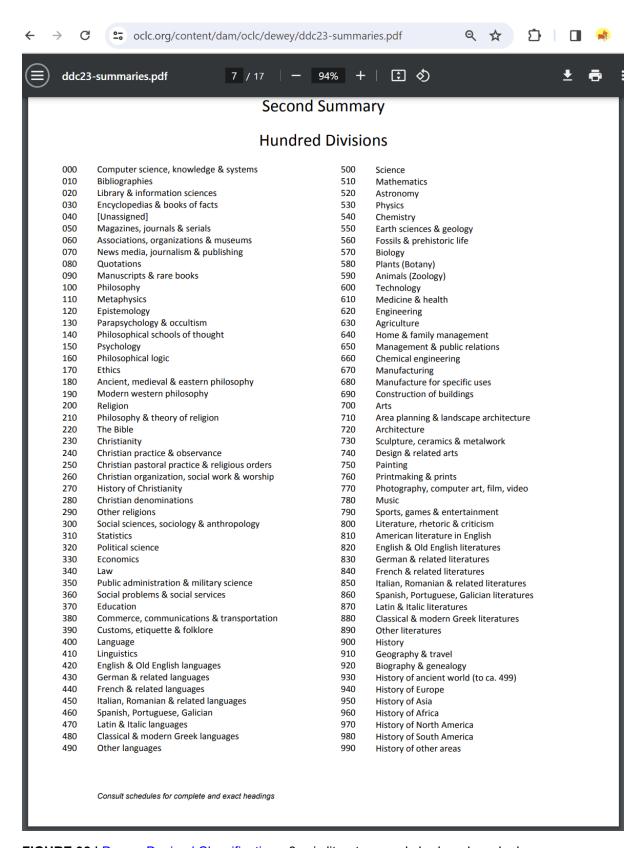


FIGURE 38 | Dewey Decimal Classification - 8xx is literature - only broken down by language

The Library of Congress Genre/Form Terms (LCGFT)

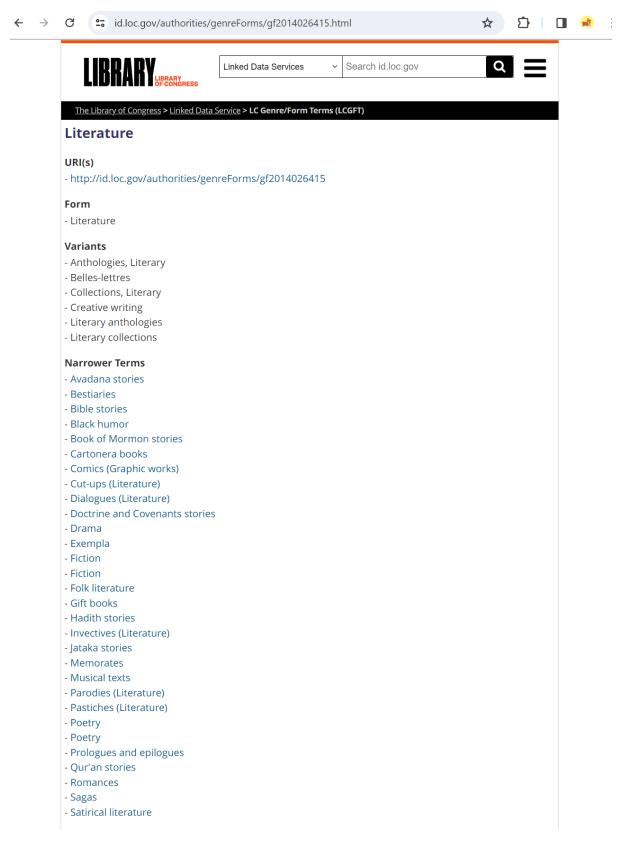
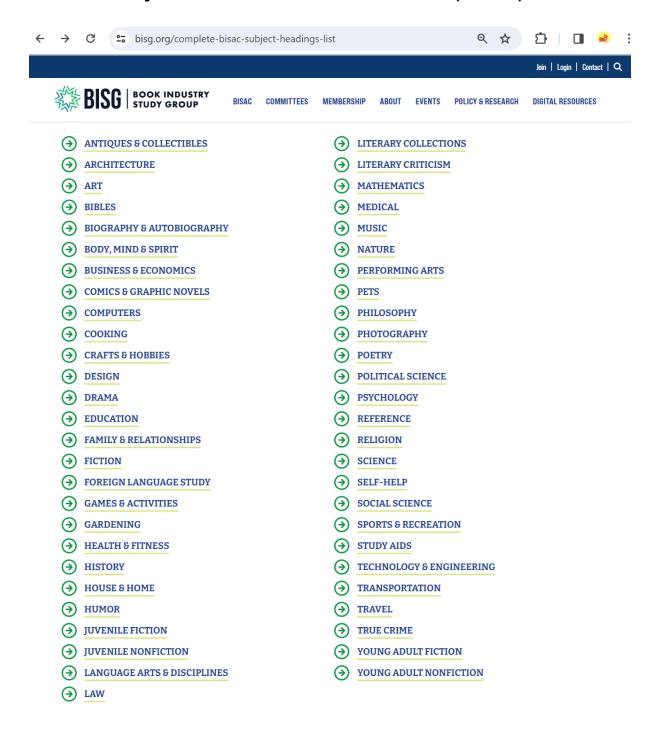


FIGURE 39 | The Library of Congress Genre/Form Terms (LCGFT) for Literature genre



FIGURE 40 | LCGFT Fiction Genre

Book Industry Standards And Communications (BISAC)



Please note: If your title doesn't have subject content, i.e. a blank book, please use the Non-Classifiable term below. Otherwise, use the Subject/General term in the closest descriptive area.

NON000000: NON-CLASSIFIABLE

FIGURE 41 | <u>BISAC</u> – Includes audiences: adult, young adult, and juvenile. Includes genres for each audience for both fiction and non-fiction.



FIGURE 42 | BISAC Adult Fiction (subset)



FIGURE 43 | BISAC Juvenile Fiction (subset)



FIGURE 44 | BISAC Young Adult Fiction (subset)

SPARQL

SPARQL does not work beyond Protégé 5.5.0. Two options are proposed:

- 1. Downgrade to Protégé 5.5.0
- 2. Use Snap SPARQL plugin

Option two did not work, so proceeded with option one. However, then the Protégé 5.5.0 small font bug appears. The fix for that is here. The fix on Windows 10 requires modifying javaw.exe properties. Other Java executables may have to be modified for other operating systems. The explanation for the small font fix is here.

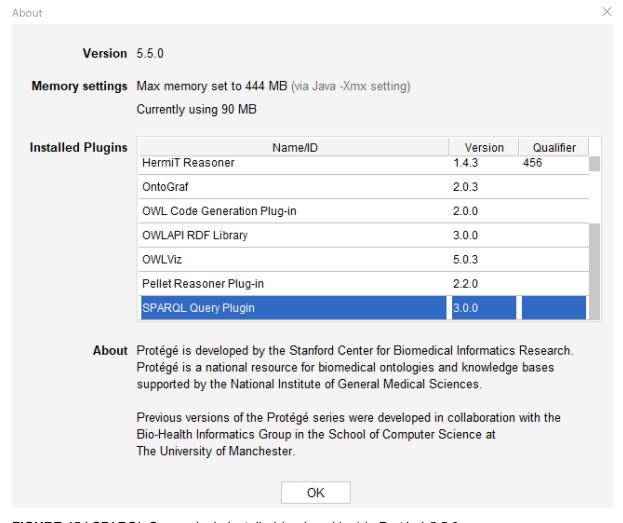


FIGURE 45 | SPARQL Query plugin installed (and working) in Protégé 5.5.0

SPARQL Query and Search

The intention was to use SPARQL to create a Search instance with the searchDate data property set with the current time when performing a search.

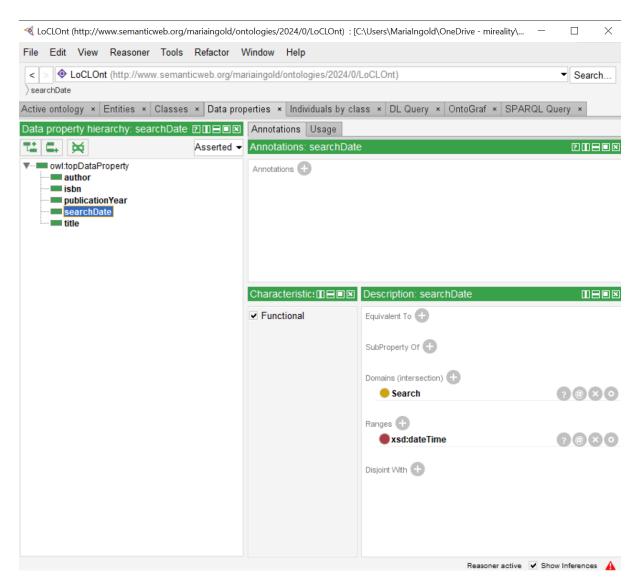


FIGURE 46 | Search class: searchDate data property of dateTime type

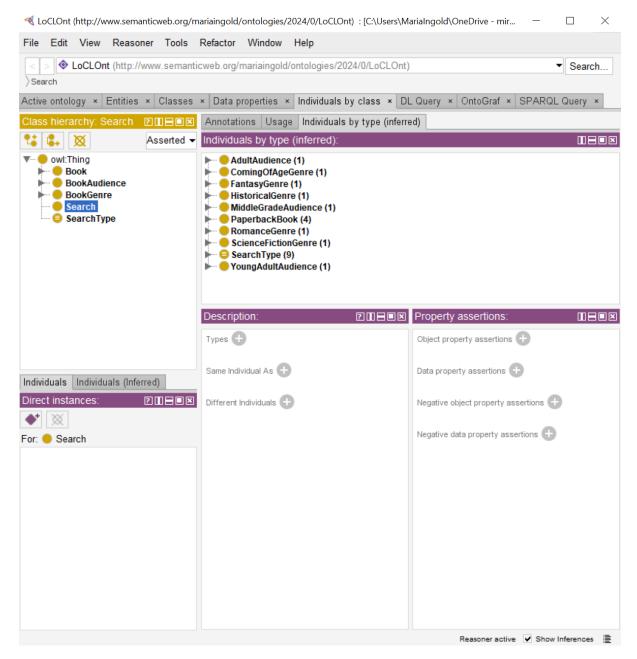


FIGURE 47 | The intention was to get SPARQL to create a Search instance with searchDate set.

However, because SPARQL Query for Protégé is only a *query* language, no updates like insertion of instances are allowed. Furthermore, NOW(), to get the current datetime, is not supported in SPARQL Query 3.0.0. As this is the only SPARQL seemingly able to work Protégé, creating an instance and updating the time of the search using SPARQL has proved impossible.

This SELECT query works:

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX LoCLOnt:
<http://www.semanticweb.org/mariaingold/ontologies/2024/0/LoCLOnt#>
SELECT ?title WHERE {
    ?book rdf:type/rdfs:subClassOf* LoCLOnt:Book;
    LoCLOnt:title ?title.
}
```

However, adding now(), even though it passes the SPARQL Query Validator, does not work. Stack overflow confirms this.

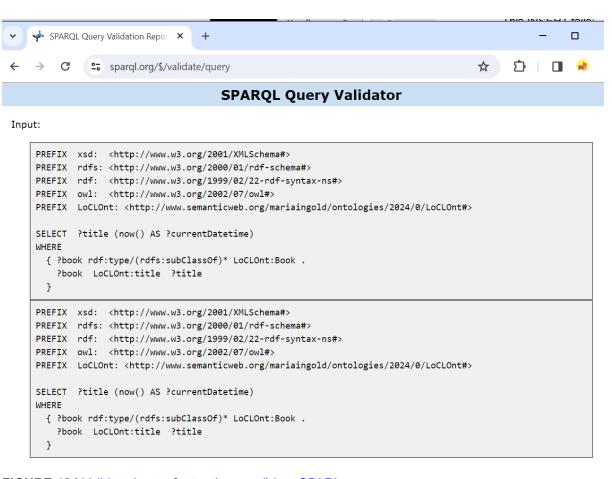


FIGURE 48 | Validated use of returning now() in a SPARL query

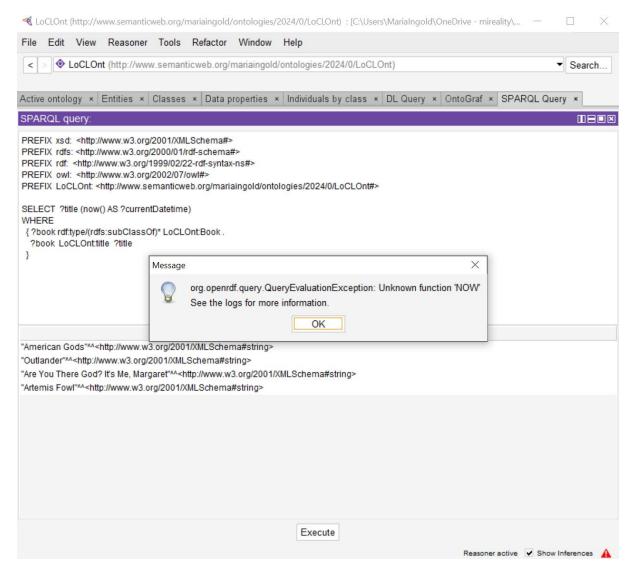


FIGURE 49 | The now() fails on a query apparently because now(), while part of SPARQL is not supported in this implementation.

While passing the <u>update validator</u>, the INSERT also fails. While it could be because of using now(), it fails on INSERT, because this is SPARQL Query, not SPARQL Update and INSERT is unsupported. Others on Stack Overflow also found that they cannot do updates on Protégé: <u>here</u>, <u>here</u>, and <u>here</u>.

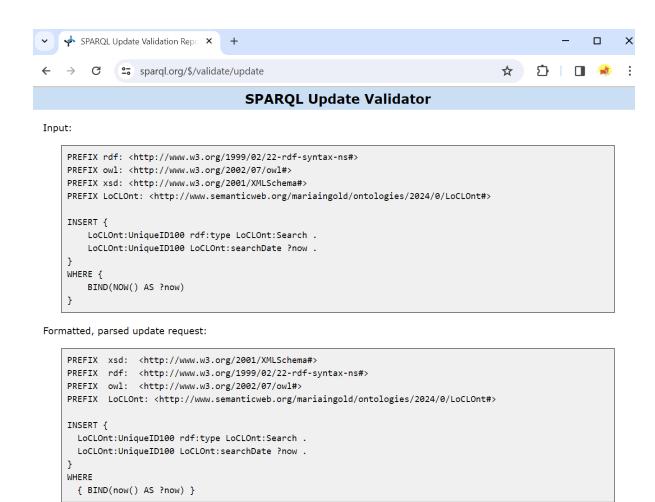


FIGURE 50 | Validating use of INSERT and now() in SPARQL

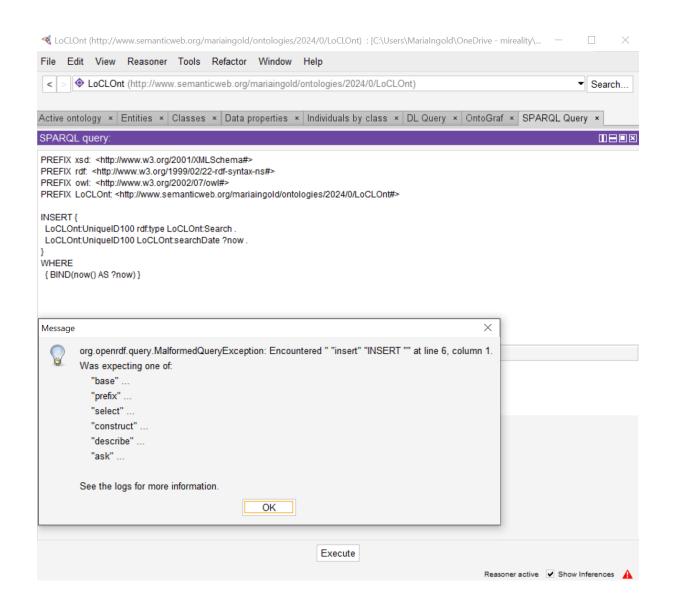


FIGURE 51 | SPARQL Query INSERT error in Protégé. Notes that it is not an expected query.