The Use of

Web Ontology Language & Semantic Web

For the Digitalization

of the

Cooley Collection

(Semester 2)

Eoin Dalton

20070289 KITMA-4

Supervisor: Mary Clarke

Bachelor of Science (Hons) in Information Technology Management

Waterford Institute of Technology

Waterford Institute of Technology

The Use of
Web Ontology Language & Semantic Web
For the Digitalization
of the
Cooley Collection

Abstract

In 2017 the Luke Wadding Library received Professor Mike Cooley's personal artefacts and belongings which he established throughout his distinguished career. The main purpose of this project is to investigate the use of the semantic web and ontologies to model and Catalog all the concept's contained within the collection. The project will investigate the technologies that make up the semantic web and the benefits of using an ontology. There will also be research into some of the programming languages available that can manipulate ontologies to allows the creation of a graphical interface that enables librarians to add the concepts to the ontology. To demonstrate how this may be done an ontology editor called Protege will be used to build a prototype ontology of some of the concept's in the collection.

Eoin Dalton [20070289]

Supervisor: Mary Clarke

Bachelor of Science (Hons) in Information Technology Management

Waterford Institute of Technology

Acknowledgements

The first set of acknowledgements for the project will go to the supervisors Brenda O' Neill and Mary Clarke. Their experience, advice and know how the whole way through the project was in invaluable. The next acknowledgement must go to Dr Larry Stapleton who back in September spoke so enthusiastically about Professor Mike Cooley and what his collection will bring to the collage. He spoke about the previous work that had been undertaken and future projects that students may partake in. Without him this project might not have been possible.

Finally, to the Luke Wadding Library and in particular Peggy Mc Hale and Kieran Cronin. They always found time for meetings even with their busy schedule. Kieran was a great help in sending relevant information about the libraries processes. To the library porters who at all times made sure access was giving to the room containing the collection. The libraries openness and support made them a pleasure to work with.

Table of Contents

Lists of Figures & Diagrams

Lists of Terms

1 Introduction

1.1 Professor Michael Cooley

This project originated through Dr Larry Stapleton who is a senior lecturer at Waterford Institute of Technology, and the Luke Wadding Library and their tireless efforts in securing the personal collection of Professor Mike Cooley.

Professor Mike Cooley was born in Tuam, Co. Galway in 1934. He worked as an engineer and was a leader in the trade union movement. He is probably best known for his work on the Lucas plan. Mike Cooley is an innovator in the field of human-centred, socially useful production, (Award, 2018) where he wrote numerous important books such as Architect or Bee, (Cooley & Noble, 1999). Human-centred Systems (Cooley, 1987).and Delinquent Genius: The Strange Affair of Man and His Technology (Cooley, 2018). He was also a founding chairman of the academic journal AI & Society in 1987 and Greater London Enterprise Board (GLEB) in 1982.

In 1981 he won The Right Livelihood Award. This award honours and supports courageous people and organisations that have found practical solutions to the root causes of global problems (Award, 2018). Mike Cooley won this award because of his work in 'designing and promoting the theory and practice of human-centred, socially useful production'.

The Lucas Plan came about in the 1970's, at which time The Lucas Aerospace company was making aircraft systems and equipment and they let it be known to employees that redundancies were coming their way. At the time Lucas Aerospace employed nearly 18,000 employees and half of their work was related to production for the military. At the time a group of people under the leadership of Professor Mike Cooley came together and created a plan to try and save the jobs of the employees.

They decided to create their own corporate plan for the company referred to as the Lucas plan. The Lucas Plan took around two years to draw up. Their team looked at market analysis and gave an economic argument. They even came up with new staff training methods. They focused on the knowledge, skills and experience of all the employees and how they could be used in an alternative socially enhancing manner. The equipment used in the production of military arms could be used also to produce more socially useful items. They then came up with alternative solutions for winding down the company's military production and at the same time keeping all the employees in their jobs (Smith, 2014).

They proposed making medical equipment like dialysis machines and looked at alternative energy techniques and using turbines as their experience in aerodynamics was highly beneficial to this. Other ideas included heat pumps and solar and fuel cells. Some ideas the employees came up with were way ahead of their time and it just shows, when given the chance, people can be very creative in their thinking (Steven, 2006).

The AI & Society journal founded by Professor Mike Cooley has been running since 1987 and is one of the world's leading forums in its field. It focuses on issues like cognitive, social and cultural issues around design and artificial intelligence. As you might tell by now Professor Mike Cooley is a very well-respected man in Ireland and in the greater academic circle for achievements in his lifetime. This is why this research is so important as it focuses on getting his collection digitized for the Luke Wadding Library so that the collection can be made available to the wider academic population, historians, trade unionists and the general public alike.

1.2 The Cooley Collection

The collection was obtained by the library in 2016 through the hard work of Dr Larry Stapleton who personally knows the Cooley family. The collection arrived in boxes and is stored on the bottom floor in room LB 13 of the Luke Wadding Library as shown in figure 1.1. In total there was around sixteen boxes but according to Dr Larry Stapleton this is not the total collection. It looks like the content of Mike Cooley's office was just placed into boxes the shipped from Galway to Waterford.

The collection is vast and contains all types of items. Items such as reports as seen in figure 1.2. Pictures and picture reals as seen in figures 1.3 and 1.8.VHS tapes and cassette tapes seen in figures 1.4 and 1.5.

Figure 1.1: Boxes containing Mike Cooley's personal belongings

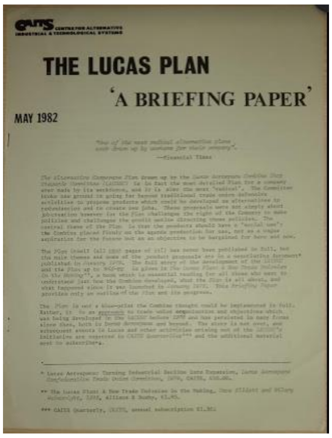


Figure 1.2: The Lucas Plan

Figure 1.3: Photographs for industrial relations marches.

Figure 1.5: VHS tapes and books.

Figure 1.4: Tapes recorder and tapes.

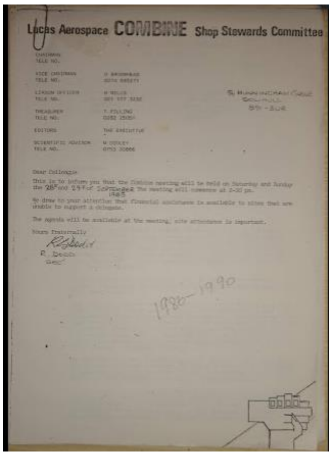
NEWS

THE GUARDIAN Tuesday March 2 1982

Shop stewards fight 'typically British' job cuts

After £21 million record profits last year, Lucas Aerospace is now sacking 1,050 workers. Peter Hildrew explains why

Figure 1.6: Newspaper articles.



Lucas Aerospace COMBINE Shop Stewards Committee

Figure 1.7: Documentation from Lucas Aerospace Meetings.

Figure 1.8: Picture real.

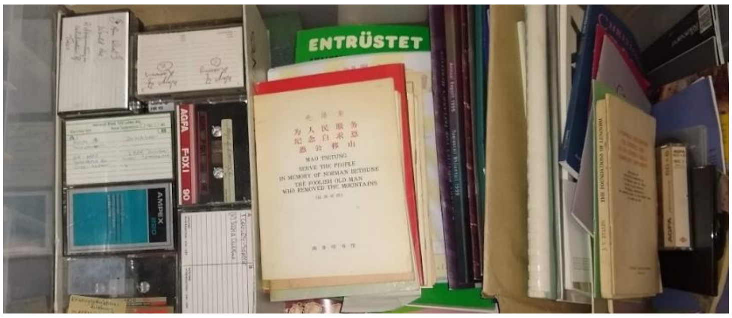


Figure 1.9: cassette tapes and letters.

Some other items that include newspaper articles, letters and correspondent letters seen in figures 1.6, 1.7 and 1.9. As seen from the figures above there is a varied number of items in the collection and that is only scratching the service. There will be more items that will be uncovered as the collection is been digitalized. In the first meeting with the Luke Wadding Library, Kieran Cronin spoke about the current metadata standard used by the library called MAchine-Readable Cataloguing (MARC). He also mentioned about the limitations of that standard for such a collection. The first research project that was undertaken about the Cooley collection looked at what is the best metadata standard for a of this type of collection and this project will speak about that a little later on.

One aspect of the collection that was noticed was the connection between the different items. For instance, in figure 1.2 you will see the Lucas Plan, A briefing paper in figure 1.6 you will see a newspaper article about job layoff and in figure 1.7 you will see a document containing notes about meetings held about the layoffs that were about to take place. There is a strong connection between all them items and that is because they are all related to the Lucas plan. As the collection unfolds you will see more connections between items because it's in relation Mike Cooley's life's works. He wrote many books and there will be correspondent letters and research about those books. It would be nice when the collection is digitalized that if a user is looking for information on the Lucas Plan that there is links to all the relevant documents for the concept.

1.3 The Luke Wadding Library

The Luke Wadding Library is part of Waterford Institute of Technology (WIT). The Library managed to obtain Mike Cooley's collection through the hard work of Dr Larry Stapleton. The introductory meeting with the Library took place on Wednesday 21th of November 2018. In attendance was Brenda O' Neill (supervisor) Kieran Cronin and head librarian, Peggy Mc Hale who had taken over from Kieran's position earlier in that year. The meeting took place in room LB13 which was the location of the collection. The conversation was about the previous research that had been undertaken. Kieran spoke about the current metadata standard that was used by the Library for cataloguing. Kieran was impressed by the previous research project that was undertaken where Metadata Encoding Transmission Standard (METS) was chosen as the most suitable standard for cataloguing the Cooley Collection.

2 Previous Research Undertaken

Last year (Academic year 2016/2017) a student from one of the computer science degrees named Matthew Hendrick undertook a research project. It involved researching the different metadata standards that are currently available and which standard would be the most suitable for cataloguing such a collection. His research looked at standards such as the Metadata Concept Description Schema (MODS), Dublin Core (DC), Machine Readable cataloguing (MARC) currently being used by the library and the Metadata Encoding and Transmission Standard (METS). Matthew's research looked at the advantages and disadvantages of all of the standards. If you refer to Figure 2.1 below you can see the comparisons of all four standards taken from Matthews research.

	MARC	MODS	Dublin Core	METS
Interoperability	✓✓	✓✓	✓✓	✓✓✓
Extensibility	✓	✓	✓✓	✓✓✓
Describe a wide range of items	✓	✓	✓✓	✓✓✓
Descriptive Metadata	✓✓	✓	✓✓	✓✓✓
Structural Metadata	✓	✓	✓	✓✓✓
Administrative Metadata	✓	✓	✓	✓✓✓

Figure 2.1: Comparisons of currents standards.

By the end of the research paper it was clear that METS was the standard that would be best suited as it ticked all of the boxes for cataloguing this type of a collection. Matthew's research found that the advantages of METS are:

- It is built specifically for library use.
- It can fulfil the full requirements of an electronic collection.
- It can describe a wide range of items.
- It can incorporate other metadata to maximise interoperability.

Matthew found that METS had no disadvantages.

In Matthew's research he described all of the elements that made up the METS metadata standard. He then demonstrated how METS was used to describe a letter that he used from the Cooley collection. The letter in question was related to the famous Lucas plan in which Mike Cooley had a big influence.

3 Purpose & Goals for this Project

(Archive, 2009) states the METS standard is useful for two purposes that is "encoding digital concept for storage" and "encoding digital concept for transport". The METS standard nearly ticks all the boxes. But what about relationships between METS documents or certain areas of a METS documents. It may be possible to create relationship through the ID tag, but it doesn't describe what the relationship is.

In his research Matthew had used an ontology editor called Protégé to build an ontology of the METS tags. The ontology was great because it showed the connection between the different XML tags in the METS standard. Using Protégé it's possible to build an ontology of the collection by defining the different items contained within the collection. The library knows the properties of a book, the properties of a VHS tapes and the properties of pictures can be found online with a simple google search. If Protégé was used to define each concept and their properties in the collection, you only have to define each concept once, then it is possible to create individuals/instances of the concept as you need them. Using Protégé it is then relatively straight forward to define the relationship that may exist between those instances as will be demonstrated later in this report.

To be able to build such an ontology, some understanding of the technologies that makes up the semantic web are need. Web ontology language in bundled up among those technologies. What Matthew researched was based upon eXtensible Mark-up Language (XML) which is also part of the semantic web but sits near the bottom of the stack. Another important technology to understand is Resource Description Framework (RDF). This report will try to give readers an understanding of the most important technologies needed for creating of an ontology of the Cooley Collection.

The scope of this project is to investigate the Semantic Web and the Web Ontology Language (OWL), and how these technologies may digitalize the collection of Professor Mike Cooley's life's works that are stored in the Luke Wadding Library. Some of the research is going to focus on how companies used these technologies to implement navigational websites such as the Internet Movie Database (IMDB). It will also focus on institutions such as the University of Neuchâtel and the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences who used an ontology editor called Protégé to build an ontology that contains all the bell's in Bulgaria.

What if we applied this theory to the Cooley collection? (Oxford Dictionary, 2019) defines a collection as "a group of people or things". The Cooley collection is made up of things or what we could call concepts. For the bell ontology they needed to describe the concept of a bell and all the related metadata that is associated with a bell. That is just one concepts. The concept of a collection would contain many concepts. Within the collection you would have the concept of a book, report, picture, VHS tape, cassette tape, video real, letter just like what is shown in the images above. The good thing about an ontology is that it is easy to extend as new concepts appear when all the boxes are opened. Once the ontology is built it can be reused to represent other collections where people may add and remove the concepts they need and the concepts they don't need.

Once the ontology is built it is possible to create individuals of the concepts and this is where you can define the relationship between the various concept in the collection

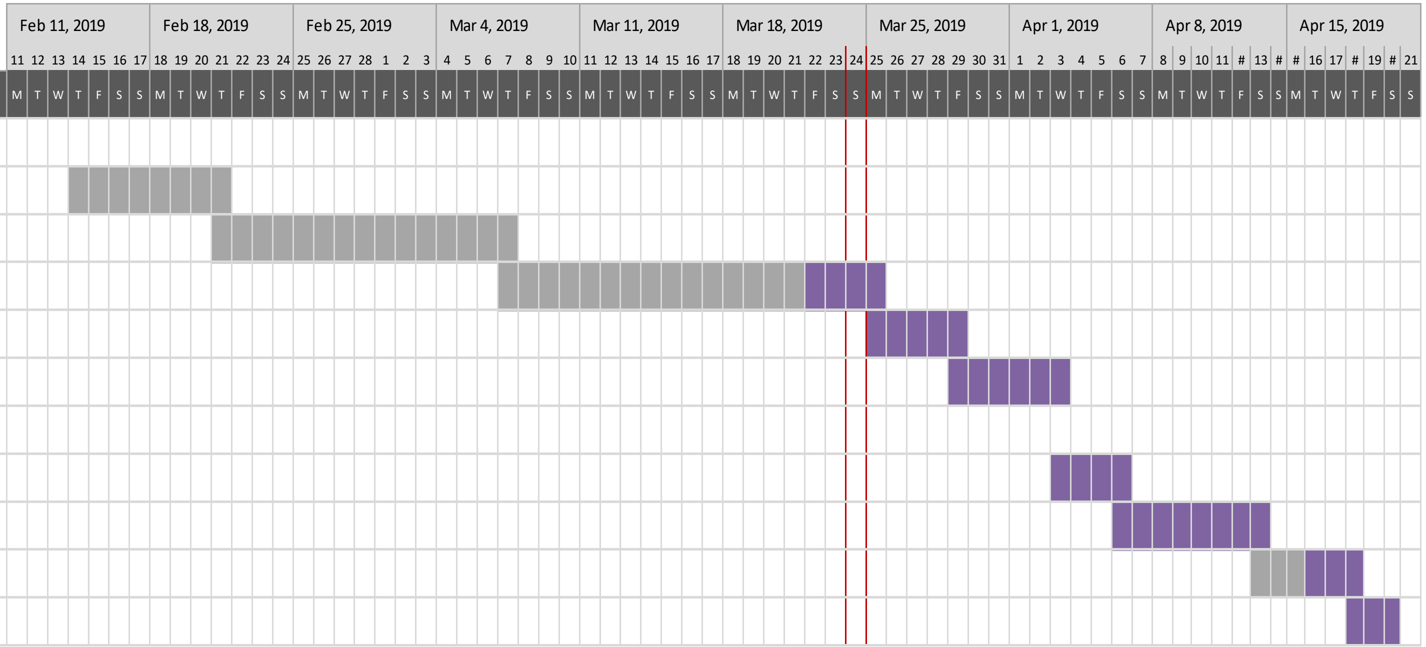
3.1 Project Plan

The Use of Web Ontology Language & Semantic Web For the Digitalizationof the Cooley Collection

Waterford Institute of Technology
Eoin Dalton

Project Start:	14/02/2019
Project Finish:	21/04/2019
Display Week:	1

TASK	ASSIGNED TO	PROGRESS	START	END
Phase 1 Title				
Task 1	Discuss XML & Ontologies	100%	14/02/2019	21/02/2019
Task 2	Discuss Owl, RDF, Protégé	100%	21/02/2019	07/03/2019
Task 3	Review of Ontology Implementation	80%	07/03/2019	25/03/2019
Task 4	Sub-Domain Prototype Analysis of the Cooley Collection	0%	25/03/2019	29/03/2019
Task 5	Sub-Domain Prototype Design of the Cooley Collection	0%	29/03/2019	03/04/2019
Phase 2 Title		0%		
Task 1	Sub-Domain Prototype Design of the Cooley Collection	0%	03/04/2019	06/04/2019
Task 2	Sub-Domain Prototype Implementation of the Cooley Collection	0%	06/04/2019	13/04/2019
Task 3	Poject Poster/Video	50%	13/04/2019	18/04/2019
Task 4	Report Review & Editing	0%	18/04/2019	21/04/2019



3.2 Project Risks

A meeting with the Luke Wadding Library in which the proposed idea was presented to Kieran Cronin took place in LB13 27/02/2019. Some of the benefits of semantic web-based ontologies were explained to Kieran and some web sites that have previously been built using the above technologies were demonstrated. Kieran stated that from the libraries point of view their main concern was the cataloguing of the concepts within the collection. Some of the benefits of their current cataloguing system is that it can verify ISBN numbers upon entry. Some of the concerns from Kieran's side was:

- Can a graphical interface be built on top of the ontology for cataloguing?
- Can it integrate with existing technologies currently used by the library?
- Will staff need training and expertise in this area in order for this to work?
- Is there some form of ISBN validation?

Trust is another big factor when building an ontology. The Cooley collection will contain may different definitions of concepts. Some concepts may already be defined on semanticweb.org and other websites, but there will be many new concepts that will need to be defined. Reuses is a big benefit of Ontologies so there has to be trust that the definition and vocabulary used in the concept is the same all around the world. One idea would be that a number of libraries or an institution such as the Library of Congress come together and agree on the terms that would be used. When we talk about the Cooley collection and in particular the Lucas Plan questions can be asked such as what kind document this may be. Is it defined as a report or a paper? This is questions that could be answered if many libraries come together and agree on the vocabulary.

All these points are both risks for this project and risks the library would be taking by adopting this approach. You just hope the benefits can out way the risks if these technologies will be fully adopted in the future.

4 State of the Art

4.1 What is an Ontology?

Previously ontologies where widely used in philosophy, for philosophers to study the existence of being. In recent years ontologies have be adapted for computer science. The artificial intelligence community refers to it as a “set of concepts or terms that can be used to describe some area of knowledge or build a representation of it” (Swartout & Tate, 1999). What this means is that it will describes a set of concept and terms within a domain. A way to share and reuse knowledge.

Standardized ontologies are being built by many professions so that domain experts can share information got to do with the current fields of study. (Noy, et al., 2000) gave some valid reasons as the why ontologies are built:

- To share common understanding of the structure of information among people or software agents
- To enable reuse of domain knowledge
- To make domain assumptions explicit
- To separate domain knowledge from the operational knowledge
- To analyse domain knowledge

4.2 What is the Semantic Web?

(Oxford Dictionaries, 2018) defines the semantic web as the “development of the World Wide Web in which data in web pages is structured and tagged in such a way that it can be read directly by computers”. (Berners-Lee et al. 2001) defines the, "The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.".

The semantic web is a vision of worldwide web consortium (W3C) to build a web of linked data. Most of the World Wide Web is based upon Hyper Text Mark Up Language (HTML) sometimes referred to as the “web of documents” (W3C, 2019). HTML is built upon defined tags so that the content can be display on web browser. Web pages are built for the purpose of displaying information to people. Humans can understand the web page but “their inherent meaning is not shown in a way that allows their interpretation by computers” (Cardoso & Sheth, 2005). The semantic web was developed to allow computers to understand information on the web referred to the “web of linked data” and that (W3C, 2019) say is the data that may be found in databases. The semantic web is based upon a stack or some people may call it a tower of technologies to help computer to interpret information better.

4.3 Semantic Web Stack

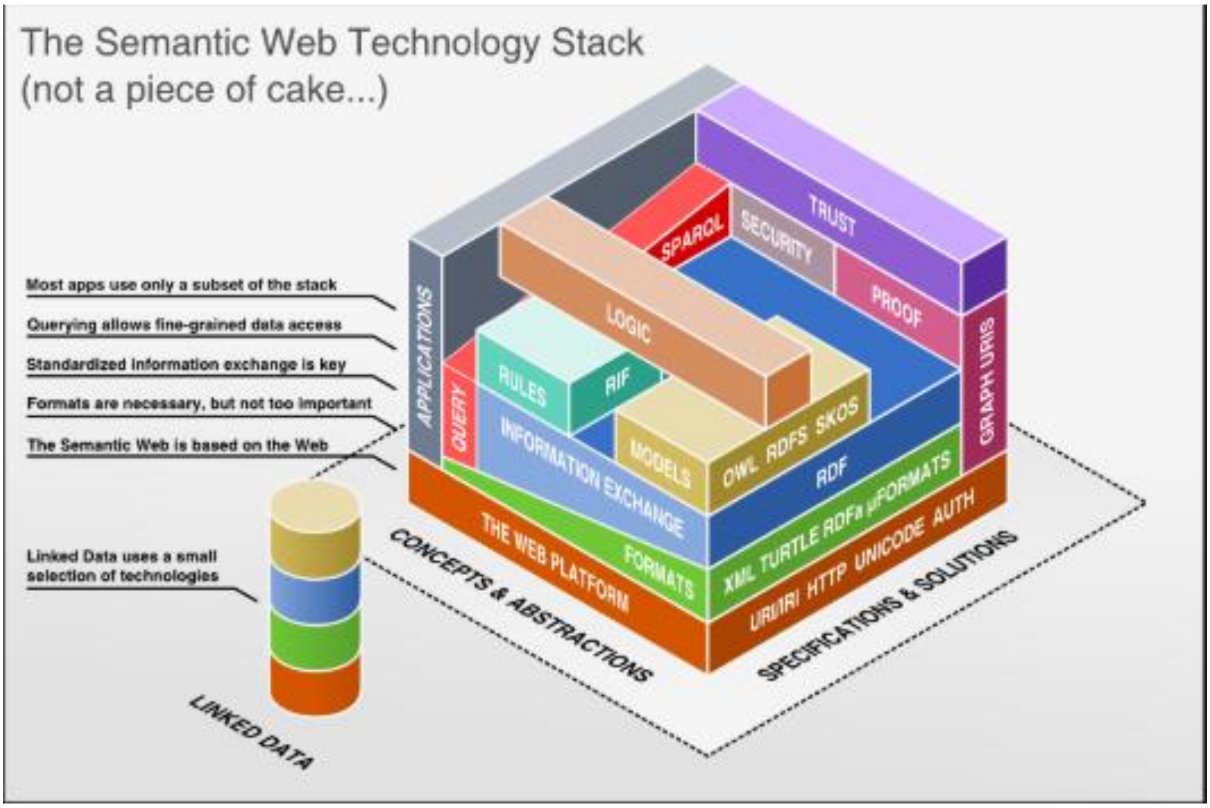


Figure 4.1: Semantic Web Stack (candy, 2019).

As referred to above the worldwide web consortium (W3C) have been assisting on the construction of a set of standard the will bring together the web of documents and the web of data so to computers will be able to do more useful work and “developing systems that can supports trusted interaction over a network” (W3C, 2019). The goal of the technologies is to enable users build places to store data on the web, create new vocabularies to write rule for manging the data. This may be achieved by using the technologies referenced is figure 4.1, RDF, OWL, SPARQL and SKOS (W3C, 2019). First, we must identify the standardized technologies that the stack is built upon like URI/IRI, URL and XML.

4.4 Uniform Resource Location (URL) & Internationalized Resource Identifier (IRI)

A URL is what we all use to locate a resource on the web. It will also contain the protocol we are using to access the resource such as HTTP, HTTPS. The URL will have a certain format if you look at the URL below it will start with the protocol followed by the domain name in the case below the semantic web which is where the webpage can be found in the standards section of www.w3.org’s website.

Protocol Domain Webpage

https://www.w3.org/standards/semanticweb/

The IRI is an extension of Uniform Resource Identifier (URI) and are used to “identify abstract or physical resources on the web” (FUSION, 2019). (W3C, 2019) defines the URI as a sequence of character chosen from a small number of US-ASCII characters. The IRI extends the set of characters and adds characters from the universal character set or UNICODE as referred to in figure 4.1. IRI where designed to replace URI because of the limitation of the URI. The limitation being that they only use a subset of ASCII characters. The IRI's use of Unicode which incorporates Latin characters and more, gives a wider range to identify more resources on the web.

4.5 eXtensible Mark-up Language (XML)

XML is similar to HTML in that is used tags. HTML uses tags like <p> which tells the web browser that a paragraph is going to follow. All the tags in HTML are predefined whereas with XML enables users to define their own tags. If you look at the following example referred to in figure 4.2 it is describing details about a contact. This example is giving information about a contact name Jorge and gives some other information about him like surname his organization and email and phone number. Businesses can use XML to standardize how business documents such as a sales orders should look. This makes it easier for computers to process the information when it is exchanged as they already the common field contained within the sales order. It is this interoperability is what makes XML standout.

```
<Contact contact_id="1234">
     <first_name> Jorge </first_name>
     <last_name> Cardoso </last_name>
     <organization> University of Co-
imbra </organization>
     <email> jcardoso@dei.uc.pt </
email>
     <phone> +351 239 790 051 </
phone>
</Contact>
```

Figure 4.2: XML example describing a Contact (Cardoso & Pinto , 2015).

4.6 Resource Description Framework (RDF)

RDF was built by W3C " and is "a language for encoding knowledge on Web pages to make it understandable to electronic agents searching for information." (Noy, et al., 2000). It was build using XML as the underlying technology and also incorporated URI to identify resources. RDF is represented by a subject, predicate and concept model as referred to in figure 4.3.

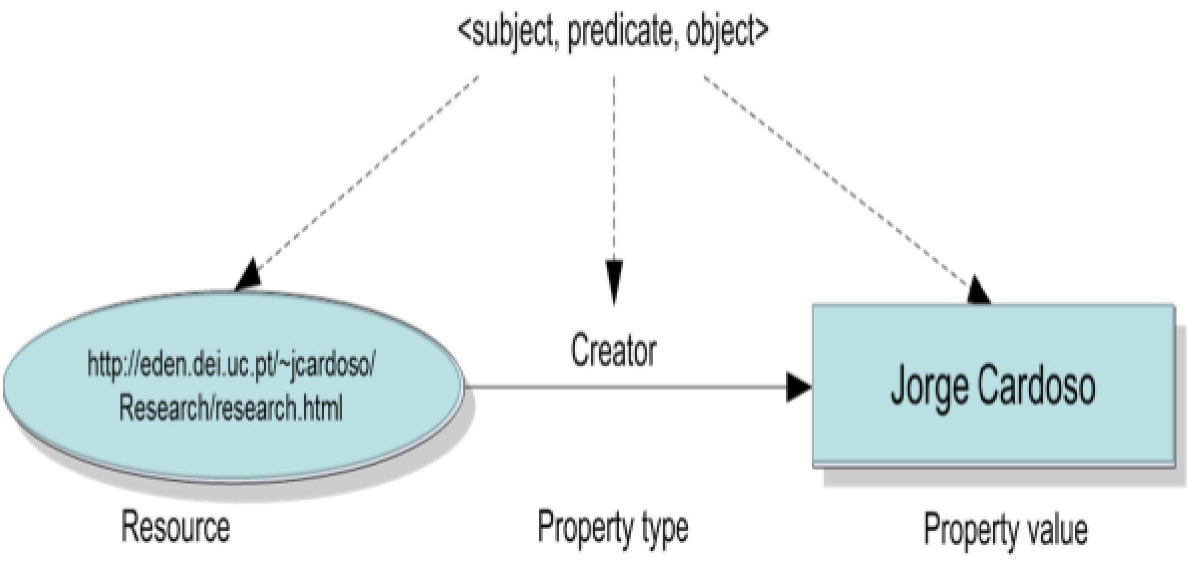
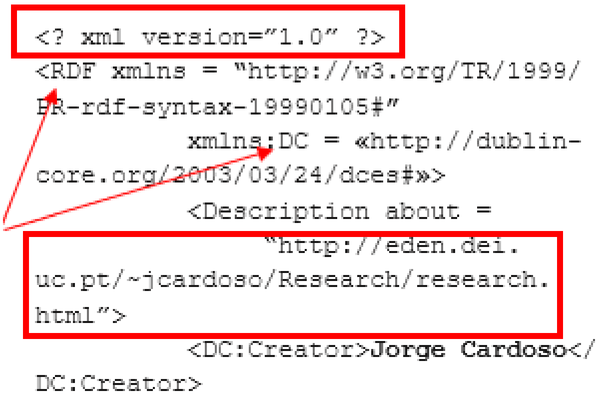


Figure 4.3: RDF triple (Cardoso & Pinto , 2015).

If you refer to figure 4.3 the subject or resource can be a something that is found in the world or universe . This means it can be a physical concept, documents, concepts, numbers (W3C, 2019). The predicate or property type is the relationship that exists between the subject and the concept. The concept or property value refers a value the relates to the resource in question. For example, the resource may be a car that is made by Ford. In this case Ford is the property value the property type is made by and the resource is the car.

Shown to figure 4.4 is the syntax that is associated to figure 4.3. Notice the boxes highlighted in red it is point to use of XML tags and the IRI identifying the resource. Also shown is the use of XML names-spaces to refer to external XML schema that have be built. The first is referencing the RDF xmlns schema and the second is referencing Dublin DC which is a metadata schema. This is telling the editor to verify this triple to those two predefined XML schemas.



```
<? xml version="1.0" ?>
<RDF xmlns = "http://w3.org/TR/1999/
PR-rdf-syntax-19990105#"
        xmlns:DC = «http://dublin-
core.org/2003/03/24/dces#»>
        <Description about =
            "http://eden.dei.
uc.pt/~jcardoso/Research/research.
html">
        <DC:Creator>Jorge Cardoso</
DC:Creator>
```

Figure 4.4: Syntax for RDF triple example (Cardoso & Miguel Pinto, n.d.).

4.7 Web Ontology Language (OWL)

OWL is a language the was designed for the semantic web. It was "designed so that information can be integrated and processed over the web" and It is used to "describe rich and complex knowledge about things, groups of things and relations between things" (W3.Org, 2018). OWL builds on RDFS and lets us define ontologies. (DuCharme, 2013) states that "Ontologies are formal definitions of vocabularies that allow you to define complex structures as well as new relationships between your vocabulary terms and between members of the classes that you defined". OWL was originally developed in 2004 by the W3C OWL Working Group. Owl 2 is an updated version of OWL the was developed in 2009 by the same group and the current version was developed in 2012. OWL is part of the semantic web technology stack (W3Schools.com, 2019) which also includes RDF (Resource Description Framework), RDFS (Resource Description Framework Schema) and SPARQL (SPARQL Protocol and Query Language) (W3.Org, 2018).

OWL 2 ontologies are like other programming languages in that is provides classes and data values it also provides individuals and properties; these are then stored as semantic web documents. "OWL 2 ontologies can be used along with information written in RDF, and OWL 2 ontologies themselves are primarily exchanged as RDF documents" (W3.Org, 2018).

4.8 SPARQL

SPARQL is the standard query language for querying RDF triplestores i.e. semantic graph databases. The standard or protocol was designed by W3C and it allows "users to query information from databases or any data source that can be mapped to RDF" (Yankova, 2016). SPARQL works the same as SQL in that it enables users to retrieve, edit data in a relation database. SPARQL can also perform queries on any type of relational database through a middleware. The middleware performs a mapping between the relational database to RDF.

Another benefit to SPARQL is the unlike SQL it can query many data stores for information. SQL can only query within one database. This can be done because SPARQL is also a HTTP-Based transport protocol, this means that users can gain access to a SPARQL endpoints through any of the transport protocols that are currently standardized (Yankova, 2016).

5 Ontology Use Cases

5.1 Internet Movie Database (IMDB)

IMDB or internet movie database found that there was the "scope for vastly improving the knowledge representation on the IMDb and ensuring that it described using a single ontology that everybody can use to view and update the information" (Avancha, et al., 2001). To achieve this IMDB implemented they're ontology using resource description framework (RDF). IMDB used tree main classes in their ontology, Movie, AwardsandNominations and SchedulesandLocation. These classes were all contained within the imdb concept classes.

Each of the classes were used to store important information. AwardsandNominations kept information about awards and nomination the people associated with the movie received. Schedules and Location describe information about where the movie is being screened and at what times its being screened. The Movie class contain metadata about the movie. This is information such as titles and movie length, actors, directors (Avancha, et al., 2001).

One of the most important classes in the ontology on the moviePerson class which contained information relating to people that produced and acted in the movie. IMDB stated that this was one of the most important classes in the ontology because every person will be related to a movie in some way. This creates linkage between the different class's in the ontology. If you refer to figure 5.1 you will see the imdb website. The arrow is pointing to their use of the moviePerson class, you will see a director, writer and star. By clicking on any them hyperlinks users will be brought to biographic information about each and the list of movies they made or started in (Avancha, et al., 2001).

Figure 5.1: IMDB website, moviePerson class.

5.2 Bulgarian Bell Ontology

The Cooley collection may be not so different. All collections have an owner. The collection its self contains books, books have authors. all of the letters contained within the collection will have been written by someone. By developing the collection as an ontology, you are enabling users to gain more knowledge about the domain they are investigating.

In a University of Neuchâtel and the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences (Bogdanova, et al., 2019) used an ontology editor called Protégé to build an ontology of all the bell's in Bulgaria. They were able to define a general bell in an ontology as referred in figure 5.2. The bell would have technical data, historic data, media data, audio data and location. Once this was defined all they needed to do was to create individual or instances for each bell. It looks very similar to building an entity relationship model for a database but without the restrictions of database implementation.

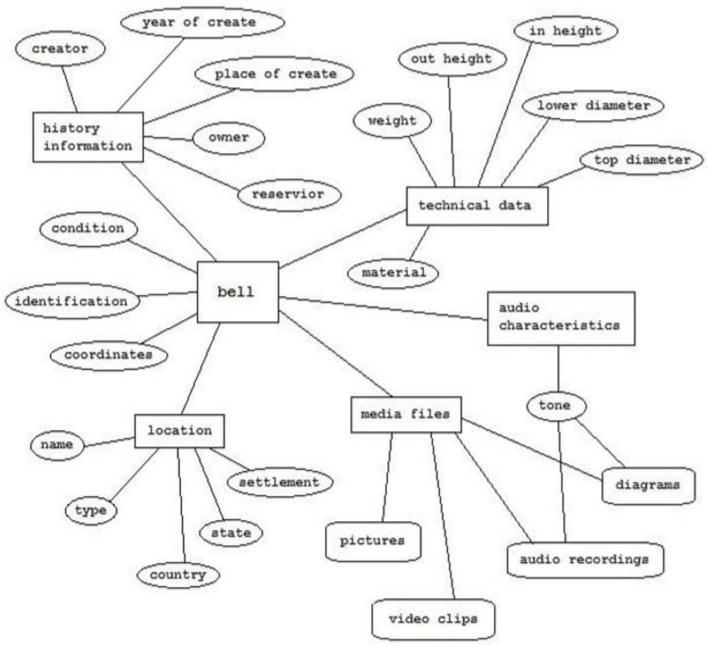


Figure 5.2: : Bell Ontology.

6 Ontology Development Process for the Cooley Collection

6.1 Development Software (Protégé)

Protégé is an opensource ontology editor the was developed be Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine (Stanford University , 2019). Protégé offer many features and addons to develop, reuse and edit ontologies. It supports has full support for OWL2, RDF and other description logic languages. One of the benefits of Protégé is the built-in reasoners. Reasoners are used to test that ontologies are consistent, and that the logic of the rules put in pace are correct. (Group, 2019)

Referred to in figure 6.1 below is the user interface developers are presented with when developing an ontology on Protégé. Giving is a brief description of the important tabs that are used in the development stage. All of the classes shown in figure 6.1 refer to concepts that are found in the collection. The data properties will refer to metadata that will be stored on each concept.

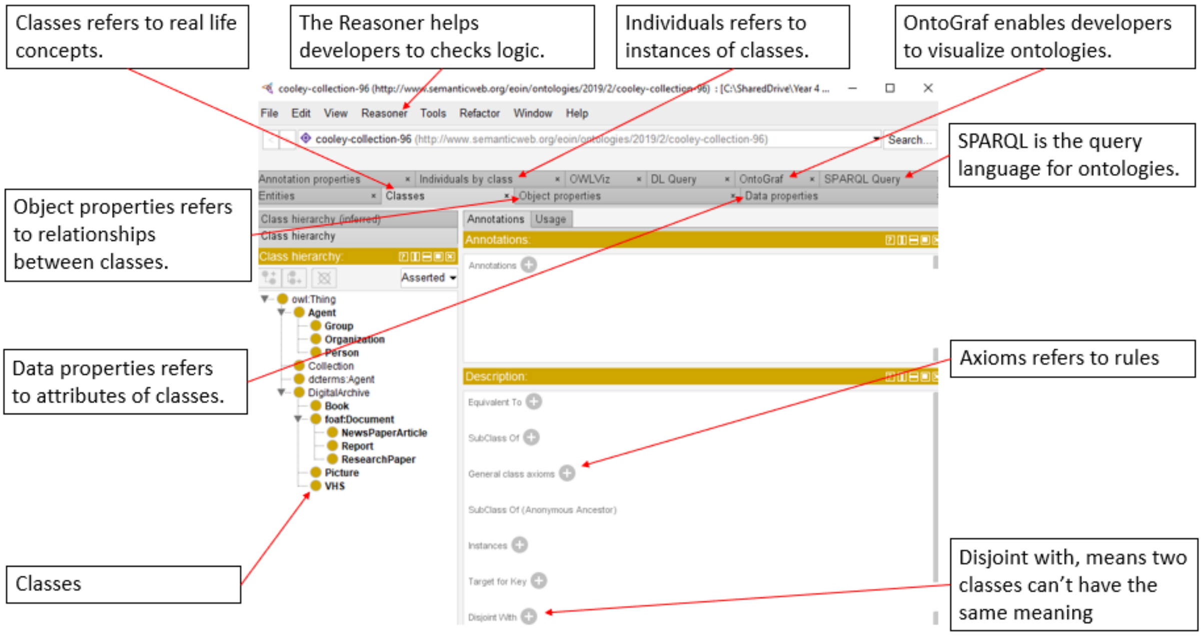


Figure 6.1: Protégé user interface.

6.2 Design & Development

(Noy, et al., 2000) refer to several guidelines that can be followed when developing an ontology and they are:

1. Determine the domain and scope of the ontology.
2. Consider reusing existing ontologies.
3. Enumerate important terms in the ontology.
4. Define the classes and the class hierarchy.
5. Define the properties of classes—slots.
6. Define the facets of the slots.
7. Create instances.

This section of the report will use these guidelines to develop a prototype ontology while giving a brief explanation of what is expected under each guideline. These guidelines follow an iterative approach in that these steps may be repeated multiple times until the ontology is complete.

6.2.1 Determine the domain and scope of the ontology.

The first step when building an ontology is to determine the domain and scope of the ontology. In order to do this (Noy, et al., 2000) refers to four questions that must be asked.

- What is the domain that the ontology will cover?
- What will the ontology be used for?
- What types of questions the information in the ontology should provide answers?
- Who will use and maintain the ontology?

The domain in which this ontology will cover is the Cooley collection and all the concepts contained within the collection. The ontology will be used like a schema that will map and relate to concepts within the collection. The idea will be to use a Java API like Apache Jena to manipulate the ontology in a way that enables librarians to catalogue the collection by creating instance of each of the concept and adding them to the ontology. When a user of the library queries the ontology with a query language like SPARQL they will be returned with the relevant information.

The questions that ontology will be answering will be about the Cooley collection. The could be returning information about the famous Lucas Plan and all the material that may be related to it. Other questions the ontology may answer would be books a particular Autor may have written. All the metadata that will be attached to the instances through data properties will enable users to gain knowledge about what they are searching for.

The last bullet point looks and the maintenance of the ontology. In order to develop an ontology that has interoperability and can be reused it is important for a group of librarians or an organisation like the Library of Congress to get involved so that many concepts that can be found in libraries can be defined such as reports, picture and VHS tapes. This is where the trust of the semantic web stack comes into effect. If an organisation such as the Library of Congress is involved it give greater trust to the ontology. Then concepts can be reused by libraries in whatever way is required such as creating an ontology of a collection such as the Cooley collection.

6.2.2 Consider reusing existing ontologies.

Reuse is one of the major benefits that ontologies offer. Ontologies for general medical science have been created and are found at (BioPortal, 2019) and are there to be reused by medical professions. At present there isn't to many ontologies that are built for digital libraries. One ontology was found that represented a book but at present the library uses MARC for cataloguing books. MARC uses IRI's to locate the resources so in the ontology for the library there only needs to be a class book with no data properties defined. When instances of books are created, they can have an object property that points to that resource using the IRI. So, book terms don't need to be defined.

For this project and to demonstrate reuses to the library two ontologies are being used. One is friend of a friend (FOAF) ontology and the other is Dublin Core (DC) which is a standard for maintaining metadata. These two ontologies are free to be used and are found at:

- Friend of a friend (FOAF) found at: http://xmlns.com/foaf/spec/20070524.html
- Dublin Core (DC) found at: http://dublincore.org/

These two ontologies can be located in Protégé by loading the ontology from URI. Referred to in figure 6.2 when a user clicks on the file, they can select to load ontology by URI, and they will be presented with a list of ontologies or a box where users can manually insert an URI of the ontology they wish to use. Notice the red arrow pointing to FOAF and DC ontologies used in this demonstration.

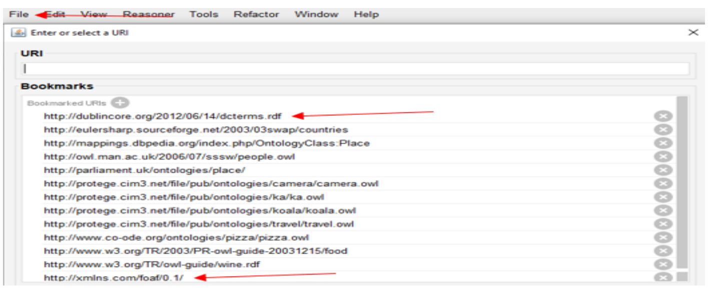


Figure 6.2: Loading FOAF & DC ontologies.

When developers are reusing ontologies the may merge the whole ontology with an existing ontology or they may only want parts of the ontology. The two ontologies that are reused in this project are quite extensive so there are only parts of the ontologies that are required. Let's look at the steps involved in this process. Referred to in figure 6.3 the first step is to navigate to refactor and find copy/move/delete axioms.

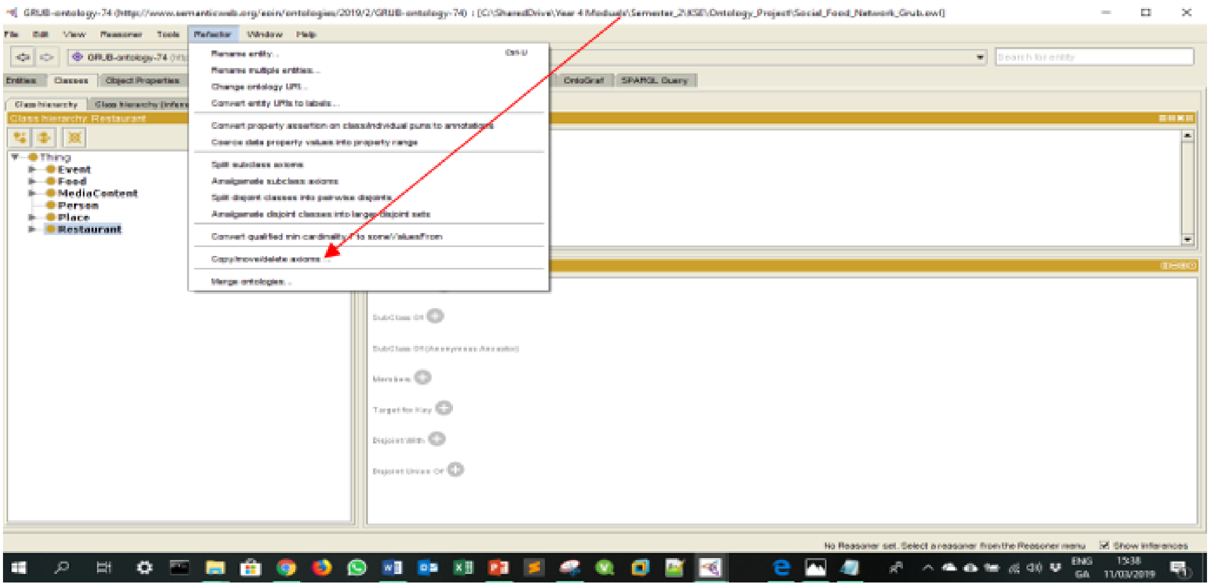


Figure 6.3: Copy/Move/Delete axioms.

After that users are presented with the screen referred to in figure 6.4. For the purpose of this example the axioms are copied by definition.

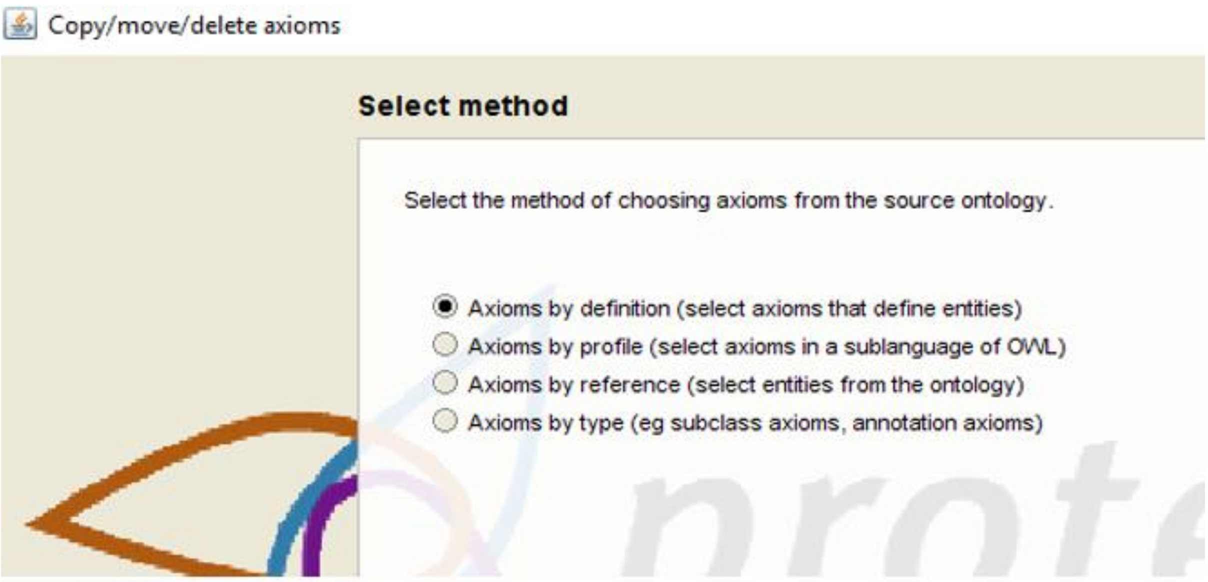


Figure 6.4: Selection Method.

Referred to in figure 6.5 below, users have the option to choose from classes, object and data properties. Users may also choose individuals, datatypes and annotation properties. The three classes that were chosen were agent, person, and organisation and the rest are data properties such as first name and surname. Later using OntoGraf we can see how these classes and data properties are used.

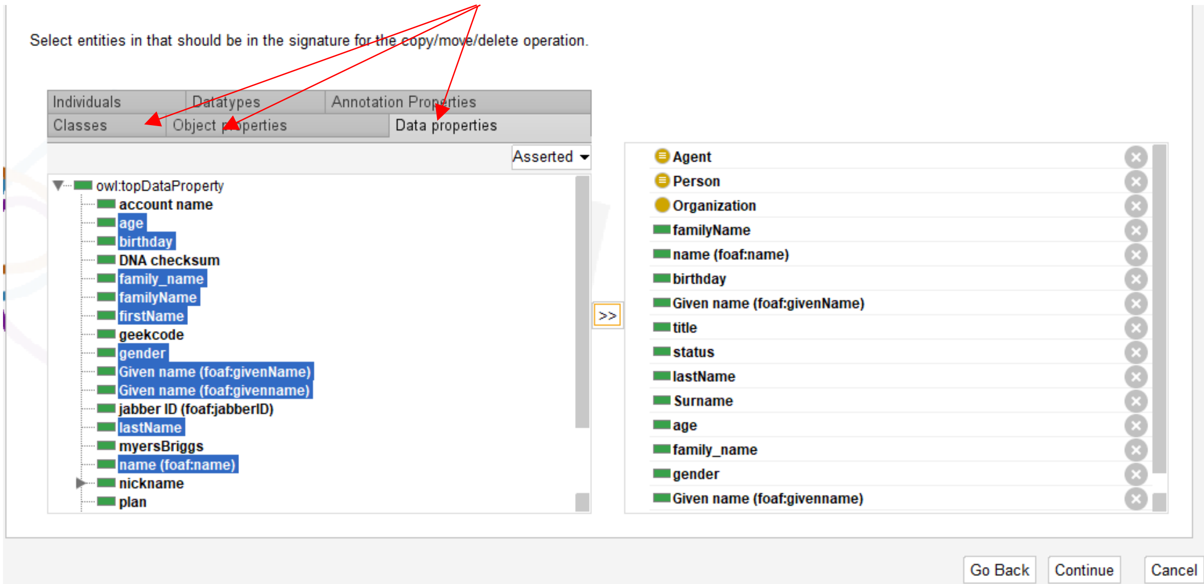


Figure 6.5: Selecting the axioms to be used.

After choosing all the parts of the ontology the user clicks continue and are presented with the screen referred to in figure 6.5 where they may copy, move or delete axioms.

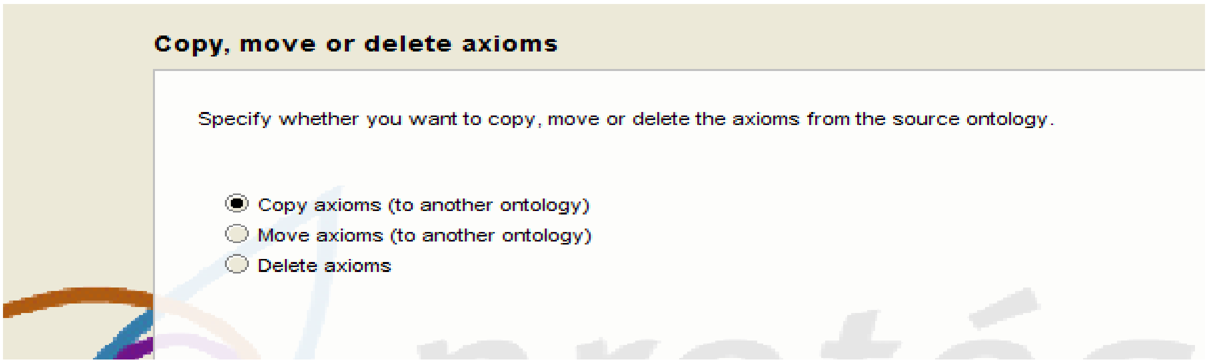


Figure 6.6: Copy, move or delete axioms.

For the demonstration it was a matter of copying the axioms. Users are then asked if the wish to copy into an existing ontology or create a new ontology referred to in figure 6.7. In this case an existing ontology was chosen. To do this all ontologies must be open in the same work space and not on separate screens. The same steps are followed for the DC ontology.

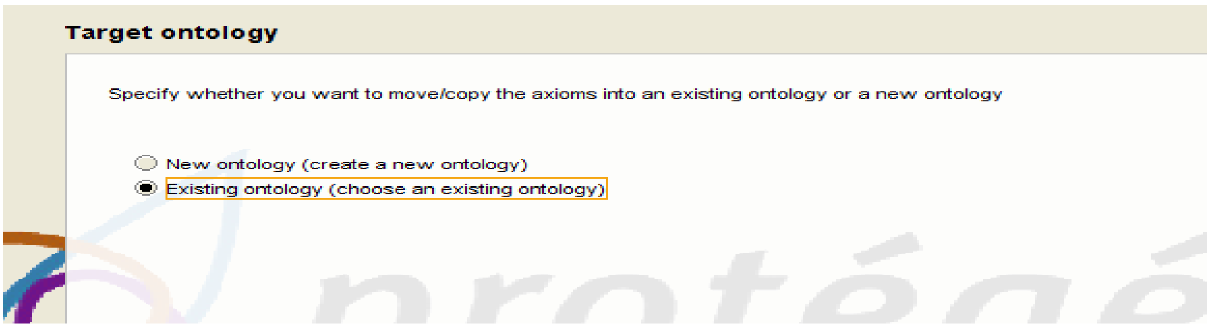


Figure 6.7: Choosing new or existing ontology.

6.2.3 Enumerate important terms in the ontology.

This involves writing down all terms and the properties of the terms that are contained within the ontology. (Noy, et al., 2000) states that it's important to get a full list of the terms and at this point overlap between concept and properties that the terms are representing is not a problem. In the next two stage is where these terms are defined. So, let's look at some of the terms that will be used in this ontology. Referred to in table 6.1 is the list of terms representing data properties and description of each term. The data properties represent the attributes of the classes.

Terms	Description
age	The age in years of some agent.
Birthday	The birthday of this Agent, represented in mm-dd string form, e.g. '12-31'.
Date	A point or period of time associated with an event in the lifecycle of the resource.
Date Accepted	Date of acceptance of the resource.
Date Copyrighted	Date of copyright.
Date Created	Date of creation of the resource.
Date Issued	Date of formal issuance (e.g., publication) of the resource.
Date Modified	Date on which the resource was changed.
Date Submitted	Date of submission of the resource.
Date Valid	Date (often a range) of validity of a resource.
Identifier	An unambiguous reference to the resource within a given context.
Gender	The gender of this Agent (typically but not necessarily 'male' or 'female').
Status	A string expressing what the user is happy for the general public (normally) to know about their current activity.
Date Available	Date (often a range) that the resource became or will become available.
Alternative Title	An alternative name for the resource.
Bibliographic Citation	A bibliographic reference for the resource.
Located	Location of a resource
FamilyName	The family name of some person.
FirstName	The first name of a person.
LastName	The last name of a person.
Surname	The surname of some person.
Gender	The gender of this Agent (typically but not necessarily 'male' or 'female').
Status	A string expressing what the user is happy for the general public (normally) to know about their current activity.
Name	The name for some thing.

Table 6.1: Data property terms.

Referred to in table 6.2 is a list of terms representing object properties. Object properties is how relationships are created between classes.

Terms	Description
Contributor	An entity responsible for making contributions to the resource.
Creator	An entity primarily responsible for making the resource.
Owns	Agent that owns an object or objects
PartOf	Refers to an object being part of a collection
Publisher	An entity responsible for making the resource available.
Author	Person that wrote some material.

Table 6.2: Object property terms.

Referred to in table 6.3 is a list of terms representing classes in the ontology. Classes represent real life concepts.

Terms	Description
Agent	An agent (e.g. person, group, software or physical artefact).
Group	A type of Agent.
Organisation	An organisation refers to a business.
Person	A human being (alive, dead, undead, real or imaginary).
Collection	Refers to a group of object or things.
DigitalArchive	Refers to an object the has been digitalized.
Book	Book is an object made up of pages usually rectangular in shape.
Document	Object containing written word.
NewsPaperAritcle	Written word about everyday life.
Report	Written word detailing some topic.
ResearchPaper	written word explaining some topic.
Picture	Still capture of an object.
VHS	Video tape containing some film.

Table 6.3: Terms about classes.

6.2.4 Define the classes and the class hierarchy.

When defining the classes and the class hierarchy (Noy, et al., 2000) refers to two approaches. The top down approach will start at the top down and start with the most general concept. The bottom up approach starts with the most specific term and work into more general terms. Referred to in figure 6.8 is the top down and bottom up approaches that may be taken when defining the classes and class hierarchy.

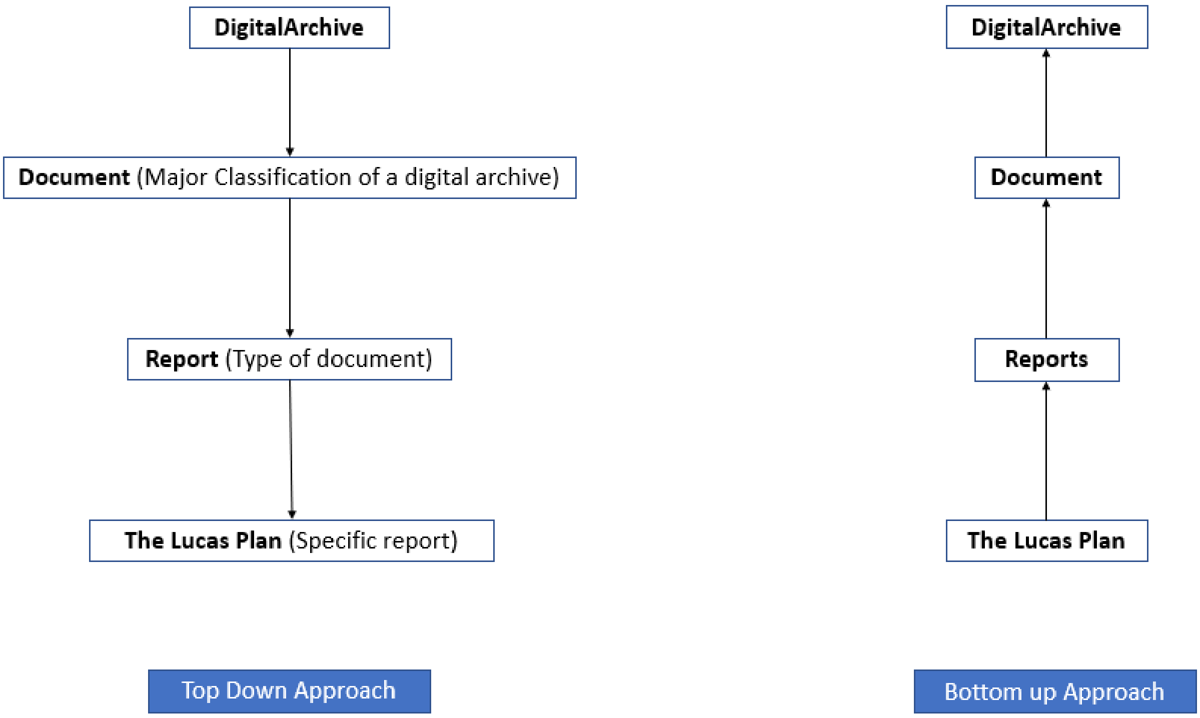


Figure 6.8: Top down & bottom up approaches.

Shown below in figure 6.9 is the class hierarchy seen in Protégé when developing the ontology. For this prototype the bottom up approach was used. After looking through the collection four object were picked to use for this project. The object included a VHS tape, Lucas Plan, newspaper article related to the Lucas Plan. These objects would be seen as the most specific term and the idea is to work up.

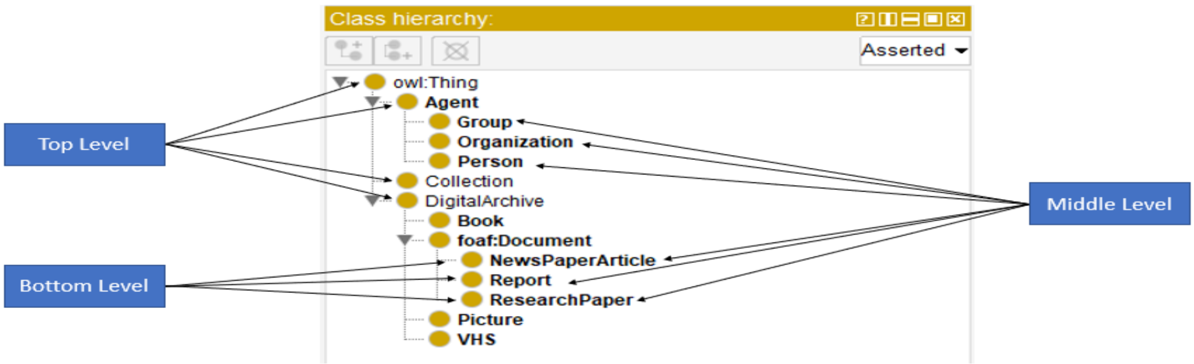


Figure 6.9: Class hierarchy in Protege.

```
<Declaration>
    <Class IRI="#VHS"/>
</Declaration>
<Declaration>
    <Class IRI="http://xmlns.com/foaf/0.1/Agent"/>
</Declaration>
```

Figure 6.10: Class declaration.

Shown is figure 6.10 is the class declaration of two classes called VHS and Agent. Notice how one is referenced with an IRI and one is not. That is because of the reuses of foaf ontology so it’s locating the class with the IRI. The other class has been defined for the purpose of this project.

6.2.5 Define the properties of classes—slots.

The properties of classes refer to the relations that may exist between classes and the attributes of each class. Because this is an iterative process the properties may change many times throughout the development process, the idea for this prototype was to use as many predefined properties as possible in order to demonstrate ontology reuses. If the library was to use the ontology approach, they would be the ones to decide what ontologies could be reused and what properties would need to be defining for the concepts.

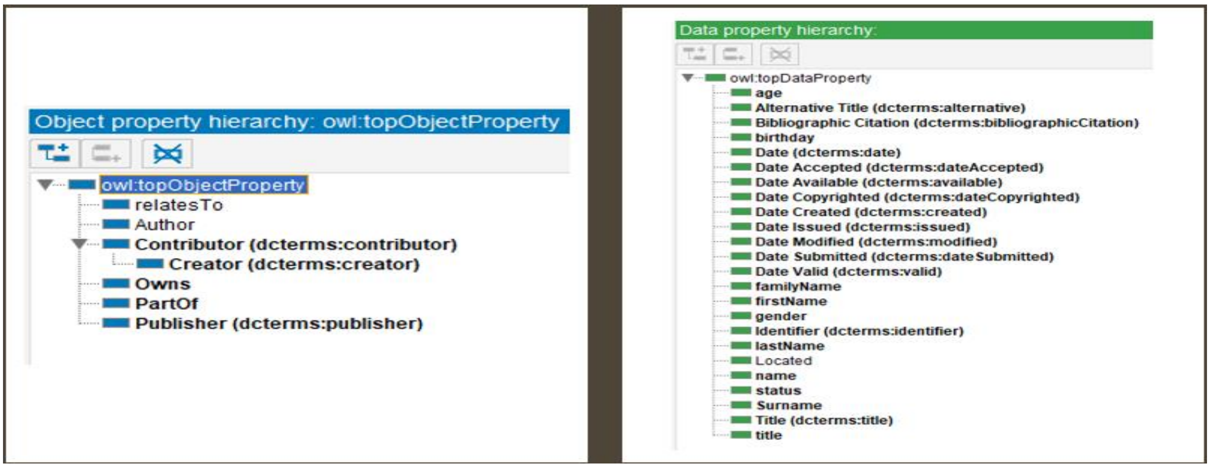


Figure 6.11: Object & Data Properties.

Referred to in figure 6.11 is all the object and data properties that have been reused and defined. Notice the dcterms, this is referring to the Dublin Core all the other object properties have been defined for the purposes of this project. The XML code generated by this is referred to in 6.12. Again, notice the use of IRI's to identify the property.

```
<Declaration>
    <DataProperty IRI="#Located"/>
</Declaration>
<Declaration>
    <DataProperty IRI="http://xmlns.com/foaf/0.1/age"/>
</Declaration>
```

Figure 6.12: Declaration of data properties.

6.2.6 Define the facets of the slots.

To define the facets of the slots refers to the rules or axioms that are put in place when defining classes, object and data properties. The rules could be:

- Value type.
- Cardinality.
- Allowed values.

The value type could refer to types such as string, integer, float, enumerated and instance. Cardinality refers to how many values a property may hold. These could be functional, inverse functional, symmetric, asymmetric, transitive, reflective and irreflective. Allowed values that are a property can incorporate. Allowed values are domain and range.

```
<DataPropertyRange>
    <DataProperty IRI="http://xmlns.com/foaf/0.1/name"/>
    <Datatype abbreviatedIRI="xsd:string"/>
</DataPropertyRange>
```

Figure 6.13: Value type, string.

Referred to in figure 6.13 is a value type. The data property name and is from the foaf ontology and it states that the value of name is of type string.

```
<FunctionalDataProperty>
    <DataProperty IRI="http://xmlns.com/foaf/0.1/age"/>
</FunctionalDataProperty>
```

Figure 6.14: Cardinality, functional data property.

Referred to in figure 6.14 is showing cardinality. This is a functional property and can only contain one value. The data property age is from the foaf ontology and this code is saying it can only have one value which make sense because a person can't have more then one age.

```
<ObjectPropertyDomain>
    <ObjectProperty IRI="#PartOf"/>
    <Class IRI="#Collection"/>
</ObjectPropertyDomain>
```

Figure 6.15: Allowed values, object property domain.

Referred to in figure 6.15 is an allowed value. This code is stating that the object property PartOf has a domain of Collection the domain refers to the object in question. Shown is figure 6.16 is the range of the data property PartOf which is book. In English this is saying that a book can be part of a collection.

```
<ObjectPropertyRange>
    <ObjectProperty IRI="#PartOf"/>
    <Class IRI="#Book"/>
</ObjectPropertyRange>
```

Figure 6.16: Allowed values, object property range.

6.2.7 Create instances.

The last step involves creating the instances/individuals. Referred to in figure 6.17 is the process of creating individuals in Protégé. At the point all classes object and data properties are created so when a user creates an individual, they just assign the appropriate properties.

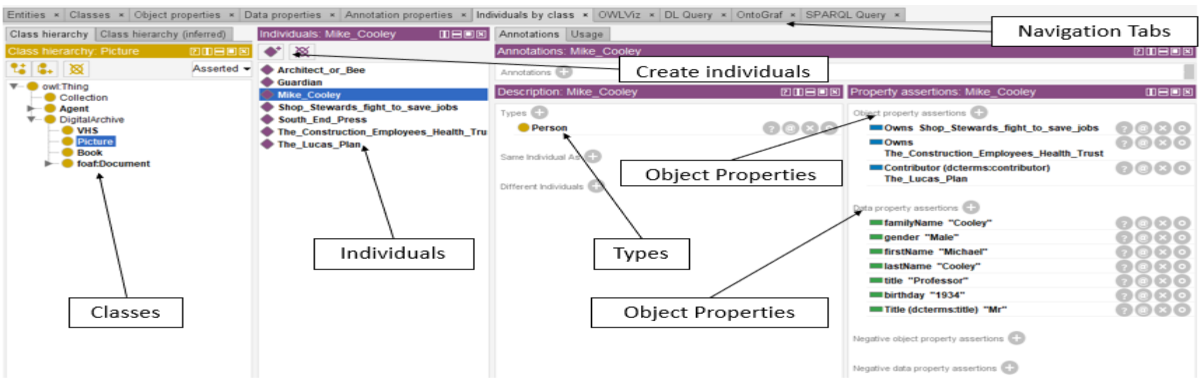


Figure 6.17: Creating individuals.

Shown is figure 6.18 is the ontology represented through OntoGraf. This view shows all the classes and individuals. There is an individual called Mike Cooley owns four objects. A book called Architect or Bee which is published by the South End Press which is an organisation. A new paper article that was published by the Guardian called Shop Stewards fight to save jobs which relates to the Lucas Plan. The Lucas Plan which is a report and which Mike Cooley contributed to. A VHS tape called The Construction Employees Health Trust. This image demonstrates the connection it is possible to get while using an ontology language like OWL.

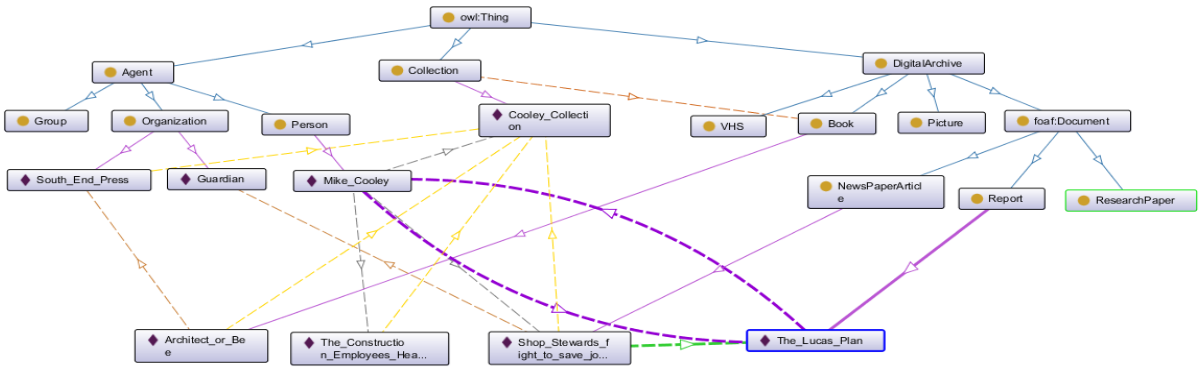


Figure 6.18: OntoGraf representation of ontology.

7 Conclusion and Evaluation

7.1 Research

The research undertaken for this project involved gaining an understanding of the semantic web and the technologies that the semantic web is build upon. The other important aspect of this project was the understanding of ontologies and the role they could play with digitalizing the Cooley collection. The point of the research is to demonstrate how an ontology of the collection may be developed and some of the advantages in doing so. By developing an ontology of the collection, the Luke Wadding library would be defining many new concepts that don't already exist on the world wide web.

Once the ontology is built professions from many different backgrounds could take advantage of it and reuse the different concepts contained within the collection. Students and researcher from within Waterford Institute of Technology would benefit from gaining more knowledge about what is contained within the collection. This could be achieved with object properties that creates relationships between the individuals. If a user of the library searches for the Lucas Plan, they main gain more insight because any correspondents related to the Lucas Plan can be easily accessed through hyperlink on that page similar to the IMDB website. The use of ontologies and the semantic web is relatively new, and it is only now organisations are starting to understand the benefits and possibilities of the technologies. As stated above the idea of this research was to gain understanding of these technologies and to demonstrate these advantages to the Luke Wadding library and to the readers of the project.

One other aspect of this project which was very satisfactory was being involved in the Computing and Maths Department Research Day and Cooley Launch where the background of this project was presented to the Cooley family and distinguished guests of the collage and can be found in Appendix 9.3.

7.2 Ontology Design

The design of the ontology was a success in that it demonstrated how ontologies may be reused. How classes are implemented, and object and data properties are attached to the classes. It explained some of the axioms that can be used to set value type, cardinality and allowed values.

However, there was some disappointment in not being able to demonstrate how a user interface may be built on top of the ontology in order to manipulate and add to it. By using a programming language like Java this is possible however there just wasn't enough time to implement something like this.

7.3 Future Work

There is a lot of future works that could come from this project. As mentioned before many concepts need to be defined to build this ontology. Each concept could be seen as its own ontology. Future work could be defining these concepts. While researching for this project a Java API called Apache Jena was identified. This API can be imported into most integrated development environments and can work with OWL, RDF as well as being used as a triple store for RDF. One video showed how an OWL file generated from Protégé where a class called person was created and three individuals of this class were created. Then Apache Jena was used to display these three individuals in a combo box using Java Swing.

8 Bibliography

Antoniou, G. & van Harmelen, F., 2004. Semantic Web Primer. *The MIT Press Cambridge,* Volume 10, p. 14.

Antoniou, G. & van Harmelen, F., 2004. Semantic Web Primer. *MIT Press Cambridge,* Volume 10, p. 2.

Apache Jena, 2000. *jena.apache.or.* [Online]
Available at: https://jena.apache.org/about_jena/about.html
[Accessed 03 04 2019].

Archive, I. V. R. L. &., 2009. *Univercity Collage Dublin.* [Online]
Available at: http://www.ucd.ie/ivrla/workbook2/wrepmets.html
[Accessed 22 02 2019].

Avancha, S., Kallurkar, S. & Kamdar , T., 2001. *Design of Ontology for The Internet Movie Database (IMDb),* s.l.: s.n.

Award, T., 2018. *Right to Livelyhood Award.* [Online]
Available at: https://www.rightlivelihoodaward.org/laureates/mike-cooley/
[Accessed 01 12 2018].

Berners-Lee, T., Hendler, J. & Lassila, O., 2001. The Semantic Web. *The Scientific American,* 248(5).

BioPortal, 2019. *Ontology for General Medical Science.* [Online]
Available at: https://bioportal.bioontology.org/ontologies/OGMS
[Accessed 05 04 2019].

Bogdanova, G., Stoffel, K., Todorov, T. & Noev, N., 2019. *Building OWL Ontology of Unique Bulgarian Bells Using Protégé Platform,* Sofia: s.n.

candy, f., 2019. *Discover ideas about Web Technology.* [Online]
Available at: https://www.pinterest.ie/fangs2/
[Accessed 23 02 2019].

Cardoso , J. & Miguel Pinto, A., n.d. *The Web Ontology Laguage and its Applications.* s.l.:s.n.

Cardoso , J. & Sheth, A., 2005. Semantic Web Process: powering next generation of processes with semantics and Web services. *Heidelberg, Springer-Verlag ,* Volume 3387.

Cooley, M., 1987. Human centred systems: An urgent problem for systems designers. *AI & SOCIETY,* Issue 1.

Cooley, M., 2018. *Delinquent Genius: The Strange Affair of Man and His Technology.* s.l.:s.l: Spokesman Books.

Cooley, M. & Noble, D., 1999. *Architect or Bee?: The Human/Technology Relationship.* s.l.:South End Press.

Devopida , 2019. *devopedia.org.* [Online]
Available at: https://devopedia.org/semantic-web
[Accessed 23 02 2019].

Dictionary, O., 2018. *oxforddictionaries.com.* [Online]
Available at: https://en.oxforddictionaries.com/definition/semantic_web
[Accessed 20 10 2018].

DuCharme, B., 2013. Learning SPARQL. *Beijing: O' Reilly,* p. 39.

FUSION, 2019. *IRI, URI, URL, URN and thier differances.* [Online]
Available at: http://fusion.cs.uni-jena.de/fusion/blog/2016/11/18/iri-uri-url-urn-and-their-differences/
[Accessed 25 02 2019].

Group, I. S., 2019. *HermiT OWL Reasoner.* [Online]
Available at: http://www.hermit-reasoner.com/
[Accessed 03 04 2019].

Noy, N., McGuinness, . F. & Deborah , L., 2000. *Ontology Development 101: A Guide to Creating Your First Ontology.* [Online]
Available at:
https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html
[Accessed 25 02 2019].

Oxford Dictionary, 2019. *oxforddictionaries.com.* [Online]
Available at: https://en.oxforddictionaries.com/definition/collection
[Accessed 23 02 2019].

Smith, A., 2014. *The Lucas Plan: What can it tell us about demorcratising technology today?.* [Online]
Available at: https://www.theguardian.com/science/political-science/2014/jan/22/remembering-the-lucas-plan-what-can-it-tell-us-about-democratising-technology-today
[Accessed 01 12 2018].

Stanford University , 2019. *protege.stanford.edu.* [Online]
Available at: https://protege.stanford.edu/about.php
[Accessed 03- 04 2019].

Steven, 2006. *ligcom.org.* [Online]
Available at: http://libcom.org/history/1976-the-fight-for-useful-work-at-lucas-aerospace
[Accessed 02 12 2018].

Swartout, W. & Tate, A., 1999. Ontologies. *IEEE INTELLIGENT SYSTEMS,* Volume 2, p. 18.

T.R.L, A., 2018. *Right to Livelyhood Award.* [Online]
Available at: https://www.rightlivelihoodaward.org/
[Accessed 01 12 2018].

W3.Org, 2018. *OWL-Semantic Web Standards.* [Online]
Available at: https://www.w3.org/OWL/
[Accessed 16 10 2018].

W3C, 2019. *Internationalization Resource Indentifier.* [Online]
Available at: https://www.w3.org/International/O-URL-and-ident.html
[Accessed 25 02 2019].

W3C, 2019. *RDF 1.1 Concepts and Abstract Syntax.* [Online]
Available at: https://www.w3.org/TR/rdf11-concepts/
[Accessed 25 02 2019].

W3C, 2019. *Semantic Web.* [Online]
Available at: https://www.w3.org/standards/semanticweb/
[Accessed 25 02 2019].

W3Schools.com, 2019. *Introduction to XML.* [Online]
Available at: https://www.w3schools.com/xml/xml_whatis.asp
[Accessed 14 02 2019].

Yankova, M., 2016. *What is SPARQL.* [Online]
Available at: https://www.ontotext.com/knowledgehub/fundamentals/what-is-sparql/
[Accessed 25 02 2019].

9 Appendix's

9.1 Supervisor Meetings

Attendance: Eoin Dalton, Mark Clarke

Week: 2 **Date:** 24/01/2019 **Time:** 9.30

Topics:

Introduction as Mary is a new supervisor.

Eoin explained what project was to Mary. Explained who Professor Mike Cooley is, what the METS standard is and why is was chosen.

Eoin explained his plan going forward for semester two.

Plan for upcoming weeks work, and subsequent weeks, if appropriate:

Eoin has first review of the Cooley collection in room LB.13 in the bottom floor of the library.

Eoin is going to look through the boxes containing the collection and chosen 3-4 item in which METS standard will be used to describe the items.

Additional Notes:

Mary is going to review Eoin's project from semester one and work done so far in semester two.

Attendance: Eoin Dalton, Mark Clarke

Week: 3 **Date:** 31/01/2019 **Time:** 9.30

Topics:

Mary suggested provides a list of suggested heading for Eoin to follow. Mary went through each heading and explained the heading in detail.

Eoin told Mary about a meeting with the Kieran Cronin from the Luke Wadding Library scheduled for Monday the 4th of February.

Eoin told Mary about the research he was planning on doing and some question the may be appropriate for this research.

Plan for upcoming weeks work, and subsequent weeks, if appropriate:

Review table of content suggested by Mary.

Meeting with Kieran in the Library with a view to collecting appropriate metadata for chosen items.

Work on report.

Additional Notes:

Mary going to send email to get clarification on what is expected in the report this semester.

9.2 Meeting with the Luke Wadding Library

Attendees: Eoin Dalton, Kieran Cronin

Date: 04/02/2019 **Time:** 15:20

Location: Kieran Cronin's Office

Topics discussed:

In relation to the chosen items, what type of metadata would the library want to keep?

How are object a currently digitalized?

Showed Kieran Archivists Toolkit a free open source software for cataloguing objects.

Outcomes:

Kieran gave Eoin some useful link in relation to metadata for different types of object. He also demonstrated how to search the website for the different types of objects.

Kieran explained how some books are sourced online from eBooks. When the object is not available online the library will digitalize the object their self's.

Kieran showed an object that was digitalized, there was a picture taken and it was stored as a JPEG image and it was retrieve through a URL.

Briefly showed Kieran the Archivists Toolkit.

Kieran suggested of making a template of four or five object that were chosen so they could be used later. He also suggested showing how each tag in the METS standard is used in the cataloguing process.

Links passed to me by Kieran for help in gaining metadata.

http://www.loc.gov/standards/mets/METSOverview.v2.html
https://www.loc.gov/standards/mods/v3/mods-userguide-elements.html
https://lccn.loc.gov/89027428/mods
https://witcat.wit.ie/cgi-bin/koha/opac-detail.pl?biblionumber=277855

https://web.library.yale.edu/cataloging

https://catalog.loc.gov

https://lccn.loc.gov/2010274282/mods

https://www.worldcat.org

Attendees: Eoin Dalton, Kieran Cronin

Date: 27/02/2019 **Time:** 15:00 – 15:40

Location: LB13 Room containing the collection.

Topics discussed:

Introduced Kieran to the idea of the Semantic Web and using an OWL with a view to build an ontology of the collection.

Showed Kieran the template that was in working process and demonstrated how you would create an individual of a book.

Showed Kieran the IMDB website the is built using a web ontology language and introduced him to the set of standards that the semantic web is make up of also showed him the Bulgarian bell collection that was implemented with the use of an ontology.

Outcomes:

While Kieran liked the idea of using a modern technology that would create great linkage within the collection, he had some worry that will need investigation.

There was a few item Kieran was worried about:

Could a user interface be implemented from the ontology that would allow to the continuation of the cataloguing process?

Would this be transparent to Librarians performs the cataloguing process?

Would it require staff training and expertise?

Could ISBN validation be built-in like then current system?

9.3 Computing and Maths Department Research Day and Cooley Launch

Computing and Mathematics International Research Day
Announcing the
Professor Mike Cooley Collection

PRELIMINARY PROGRAMME
Thursday March 7th 2019

1.30 Welcome Addresses:
Dr. Peter McLoughlin, Head of Science and Computing
Dr. Mark White, Vice-President for Research, Innovation & Graduate Studies (TBC)

Opening Address: "Professor Mike Cooley: Systems Engineer and Visionary", President Michael D. Higgins (TBC)

2.00 "Architect or Bee? A Tribute to Mike Cooley"
Distinguished Professor Karamjit Gill,
Editor-in-Chief, Artificial Intelligence and Society

2.30 "Deep Learning versus Wise Learning"
Distinguished Professor Petros Groumpos, University of Patras, Greece.

3.00 Tea/Coffee Break and Poster Session

Computing Research at WIT: The Cooley Collection
3.30 "The Cooley-WIT Link: Cross-Faculty Human Centred Systems Research and Teaching at WIT"
Dr. Larry Stapleton, INSYTE

3.50 "The Cooley Collection: The Challenge of Curating Complex Objects in a University Library"
Kieran Cronin, Librarian, Luke Wadding Library (TBC)

4.10 Extending Digital Document MetaData Standards - Modelling Machine-Readable Knowledge in the Cooley Collection"
Dr. Larry Stapleton, Brenda O'Neill, Matthew Kendrick and Eoin Dalton

4.30 Curating Objects in a Digitised Library: Challenges of 3-D Objects in the Cooley Collection: Patrick McKinerny, WIT

Mathematics Research at WIT
4.45 Presentation 1
5.00 Presentation 2

5.15 Close: Dr. Padraig Kirwan, Head of Department of Computing and Mathematics.

INSYTE
Centre for Information Systems and Techno culture
Waterford Institute of Technology

9.3.1 Computing and Maths Department Research Day and Cooley Launch Presentation

Hello ladies and gentlemen.

As Larry said my name is Eoin Dalton and I am currently in year 4 for studying Information Technology Management.

For my 4th year project I am researching the use of an ontology to model the Cooley Collection.

Ontologies have been used to build some famous websites and knowledge bases that you may have used or heard of, such as Amazon and IMDB movie website

If you look at the image in the presentation you will notice on ontology that was built so that people could represent knowledge about bell's in Bulgaria.

The people who built the ontology defined the concept of a bell and all so all the metadata that needed to be defined to describe a bell. They only needed to define the model once and when this was done the model could be used to create individual instances for each bell that they wanted to store knowledge on.

Some of the benefits of modelling the collection like this is that allows any knowledge discovered about the collection be shared easily. Ontologies all so allow for the representation of relationships between concepts and individuals.

Where this may be useful in the collection is relationships between books that Professor Mike Cooley has written, and the correspondents or research contained within the collection that is related to those books.

Once the ontology is built is can be reused by other developer who need to implement similar ontologies. Developers can extend and remove the parts of the ontology they need and don't need to suit the ontology they are building.

What I am investigating is could something similar be done to represent Professor Mike Cooley's Collection. If this model can be built, can a user interface be built on top of this to enable the librarians to catalogue each object that is contained within the collection and can this knowledge be shared with reader in the library and to the wider public.