# Crowdsourcing Semantic Resources

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of Doctor in Philosophy by

**Anton Roscoe Minnion**

May 2015

# Contents

# Illustrations

## List of Figures

# List of Tables

# Preface

This dissertation is original, unpublished, independent work by the author, Anton R. Minnion.

# Abstract

Finding easier and less resource-intensive ways of building knowledge resources is necessary to help broaden the coverage and use of semantic web technologies. Crowdsourcing presents a means through which knowledge can be efficiently acquired to build semantic resources. Crowds can be identified that represent communities whose knowledge could be used to build domain ontologies. This work presents a knowledge acquisition approach aimed at incorporating ontology engineering tasks into community crowd activity. The success of this approach is evaluated by the degree to which a crowd consensus is reached regarding the description of the target domain. Two experiments are described which test the effectiveness of the approach. The first experiment tests the approach by using a crowd that is aware of the knowledge acquisition task. In the second experiment, the crowd is unaware of the knowledge acquisition task and is motivated to contribute through the use of an interactive map. The results of these two experiments show that a similar consensus is reached from both experiments, suggesting that the approach offered provides a valid mechanism for incorporating knowledge acquisition tasks into routine crowd activity.

# Acknowledgements

This work would not be possible without the support and encouragement of my family. In particular my mother Jean, my father John and my step-mother Sandra. Additionally I would like to acknowledge my fellow occupants of Room G22, Martyn Lloyd-Kelly, Matias Garcia, Adam Wyner, Muhammed S. Khan and Luke Riley. I would also like to thank my supervisors Peter McBurney, David Jackson and Valentina Tamma.

I would also like to acknowledge Nick Anning, academic, investigative journalist and colleague who sadly passed away in 2003. Without his advice and support I would never have started my academic career.

# Chapter 1

# Introduction

This research did not start as a semantic web project. It started by looking at ways to make maps smarter so that they could be used to perform more sophisticated tasks. The problem was that traditional digital maps such as *GoogleMaps* or *OpenStreetMap* are built on quantitative data, with all the entities on the map largely defined by their quantitative properties. Map objects were defined as being points, lines or polygons; each rooted to a coordinate system and the distance between objects measured in metres and miles. The problem with this quantitative foundation is that it fails to represent how people view the world. A new way of representing locations on maps was needed, one which could communicate the qualitative properties of that location. So instead of something being ten metres away, it could be described as being nearby. So that an architect can describe a building in the terminology that makes sense to him/her, while a restaurant critic could describe the same building according to the properties of the restaurant it houses. As a result of investigating the idea of building qualitative maps, a solution emerged — ontologies. Ontologies take many forms and can relate to knowledge domains at both higher and lower levels. For the purposes of this thesis, an ontology provides a description of the knowledge needed to describe a particular subject domain.

Ontologies usually consist of concepts, which determine what type of objects are being described, and object properties which establish how concepts can be linked. They may also contain the instances of the concepts along with data properties that can be used to link concepts to static data such as strings, integers or even coordinates. Additionally, ontologies embody a logical system that can be understood and used by computers to infer new knowledge and make new connections. If maps were based on ontologies, the kind of qualitative links that are needed to build smarter maps, and more adaptable for specialist groups of users, could be created.

The problem of using ontologies in this way is that they are difficult to make. There is a scarcity of people with the skills needed to build ontologies. In order to bring the power of ontologies to a broader range of applications, less resource-intensive ways of creating ontologies are needed. Some work has already been done which aims to automate the building of ontologies, typically these analyse a body of domain-specific data from which an ontology can be inferred. These techniques can broadly be referred to as 'ontology learning' and are described in more detail in Section 2.5 of this thesis. The problem with existing ontology learning techniques is that they still require human input to make them work. Moreover, the input that is needed has to come from people with specific skills and understanding of ontologies. The approach offered in this thesis aims to reduce the need for expert human input in the ontology learning process by obtaining consensus through crowd mediation.

In recent years, when human-resource issues presents itself in terms of data collection, it is not long before crowdsourcing is mooted as a potential solution. Crowdsourcing seeks to use the interconnectedness of the web to collect data and perform tasks by utilising human input. Traditionally crowdsourcing has been seen as a way of distributing large tasks so that they can be performed quickly. But another aspect of crowds is that the can be distinguished by the subject or the task that they address. *GalaxyZoo*[1], for example, which looks to harness crowd input to categorize observed galaxies according to their shapes, takes advantage of the potential scale of the crowd to perform a huge task. However, to contribute to *GalaxyZoo* a participant would most likely be motivated by an interest in astronomy and would likely have a level of expertise. Therefore, we can attribute a greater degree of astronomical expertise to the crowd contributing to *GalaxyZoo* than would normally be present. Along with the collective knowledge of the crowd, the fact that they are willing to spend time contributing towards such a project indicates a good level of motivation.

Using crowdsourcing to build an ontology may be a trivial task compared with mapping the shape of billions of galaxies, however it does offer major challenges. While the scale of the crowd is of much less importance, the sophistication of the tasks that need to be performed is much greater. However, if a crowd can be identified whose interests encompass a specific and identifiable knowledge domain, if a consensus over the knowledge within the crowd can be reached, and if that crowd can be motivated to perform the tasks required, then it may feasible that ontologies and other semantic resources could be crowdsourced. This is the hypothesis that is addressed in this research.

The research questions being addressed by this work are:

- Can a mechanism be found for eliciting knowledge from the crowd in such a way as to facilitate ontology building?
- Can conflicting crowdsourced knowledge be reconciled to provide a consensual model of a domain?
- Can knowledge that is unable to be reconciled to a consensus be effectively identified so that it can be dealt with manually?
- Can the mechanism be embedded into routine crowd activity so that the ontology building task is hidden?

[1]http://http://www.galaxyzoo.org

- To what degree can crowdsourcing produce a consensus over the description of a knowledge domain?

The rest of this work is organised as follows. In Chapter 2 a review of the relevant literature is made, providing the context in which this research can be placed. Chapter 3 presents a mechanism through which a consensual ontology model can be obtained from a crowd. The first two experiments are described in Chapter 4 and Chapter 5. These experiments tests this approach to see if it works with more traditional ontology building methods with crowds that are task-aware. The third experiment, described in Chapter 6 uses a digital map that allows users to create a map resource that is of interest to the crowd, while at the same time manipulating the crowd into building a consensual ontology. A discussion on the results of these two experiments will take place in Chapter 7 and a conclusion provided in Chapter 8.

# Chapter 2

# Literature Survey

## 2.1 Preface

This literature survey will summarise the most relevant and influential work relating to the employment of ontology in knowledge-based systems (KBS) with emphasis on the knowledge acquisition process that communicates human knowledge to electronic form.

### Notation

In this and subsequent chapters the term 'expert' is used to describe the domain expert, that is to say the person or people who hold the real-world knowledge regarding the domain which is being modelled. Unless otherwise specified, the 'expert' will have no or unknown competency in regard to knowledge engineering and ontology engineering processes. 'Knowledge engineer' is used as a generic term to describe the person or people who have the required technical skills to implement knowledge engineering and ontology engineering methodologies and to construct KBS and their component parts. The distinction between *knowledge acquisition* and *knowledge elicitation* should also be made. Knowledge elicitation is the process of obtaining knowledge from the expert through techniques such as interviews and brainstorming [99]. Knowledge acquisition is a broader set of activities, including knowledge elicitation, along with the explication and formalization of that knowledge into a form in which it can be used in KBS [18].

## 2.2 Knowledge Engineering

### 2.2.1 Development of Knowledge Engineering

Knowledge engineering has traditionally been defined as the process of developing information systems in which knowledge and reasoning play pivotal roles [99]. Over time these information systems (or KBS) have evolved and been adapted to meet the needs of ever more sophisticated applications which employ increasingly powerful reasoning techniques that rely on knowledge models with greater expressiveness.

Studer et.al. [105] provide an analysis of the general change in thinking among knowledge engineers that occurred towards the turn of the last millennium. The changes described are important as they have led to the development of new ideas and practices which have come to underpin modern knowledge engineering, paving the way for the incorporation of resources such as ontologies that enable sophisticated reasoning tasks. Studer et.al. characterise these changes as being part of a general paradigm shift from a 'transfer view' of knowledge engineering, in which KBS presents a direct reproduction of human knowledge, to a 'model-based' view where there is more emphasis on the problem-solving dynamic and where knowledge is organised into multiple models that contain only the required knowledge to address particular sets of problems [105]. According to Studer et.al.:

> "building a KBS means building a computer model with the aim of realizing problem-solving capabilities comparable to a domain expert" [105].

This statement means that the objective of modern knowledge engineers is to solve problems in a way *comparable* to the expert, meaning that there is no need to simulate the cognitive processes associated with human problem-solving, as might be the case in a knowledge transfer approach [105]. According to Studer et.al., one factor leading to this paradigm shift was the identification of problem-solving methods (PSM) that were independent from the human cognitive process but, nonetheless, could be adapted to support the development of more sophisticated knowledge representations that support a similar degree of inference to human-cognitive approaches.

Making problem-solving knowledge explicit and regarding it as an important part of the knowledge contained in KBS, is the rationale that underpinned the increased focus on PSM in the pre-web era [105]. Credit for the identification and early formalisation of PSM is given to Clancey [17] who identified common problem-solving behaviours in early attempts at building KBS, even when the representations produced from these systems differed considerably; and from these generic behaviours he developed a heuristic classification. Studer et.al. break down this heuristic classification into a set of roles and inferences whereby observable knowledge, such as a temperature or a name, can be associated with an abstract, and from that abstract a heuristic match can be made from which a solution can be inferred [105].

From this process it can be seen that not only the observable knowledge, but also the knowledge required to perform the various inferences, needs to be acquired from the expert. Studer et al. consider the identification of PMS to be a major stimulus behind the development of the two main approaches to knowledge engineering that emerged in the 1980s, Role-Limiting Method (RLM) and Generic Task (GT) approaches [105]. RLM employ a particular PSM and are therefore limited to addressing the tasks appropriate to that PSM [105]. One important aspect of RLM is that the particular PSM employed is not influenced by the knowledge acquisition process used [105]. Therefore the interaction with the expert that is required to formulate a knowledge representation is separate from the process of identifying the appropriate PSM. Bylander et.al. introduces GT approaches as a way of exploiting the “interaction problem” which states that the representation of knowledge is dictated by the problems which that knowledge is being acquired to solve [11]. GT approaches associate each problem-solving task with a particular knowledge acquisition method [11], meaning that unlike RLM, each task is influenced by the knowledge acquisition process employed. GT approaches require a pre-determined and fixed knowledge structure of the domain to be specified, that provides the knowledge needed to solve a task [105]. This means that multiple PSM can be employed to fit specific problem-solving tasks. While the work presented in this thesis is not directly related to KBS, it is important to provide as broad background as possible so that a greater understanding can be achieved.

### 2.2.2 Knowledge Engineering Methodologies

#### CommonKADS

The evolution of knowledge engineering described above has helped establish the theoretical principles that govern modern knowledge engineering practices, particularly in terms of making the distinction between types of knowledge, and in the need to specify problem-solving functionality. From this research, a focus on devising knowledge engineering methodologies emerged with the aim of providing a more controllable development cycle for building and maintaining KBS.

*CommonKADS* is a knowledge engineering methodology that, through a long period of development, has emerged as one of the pre-eminent candidates for building KBS. Schreiber et.al. provide a comprehensive description of the CommonKADS methodology [99]. CommonKADS is a Knowledge Engineering methodology that boasts considerable academic and industrial support, and which is currently used to support numerous projects across a variety of disciplines [109, 118]. The development of CommonKADS arose from the need to build industry-quality knowledge systems on a large scale and in a structured, controllable and repeatable way [99].

CommonKADS accepts that knowledge engineering is not simply the transference of human conceptualisation to a machine, but the construction of ‘purposeful abstraction of some part of reality’ [99]. Here we can see an emphasis on the purpose of the produced model, and an acknowledgement that a complete model of human conceptualisation may be unnecessary. According to Schreiber et.al., knowledge has a ‘stable’ structure that can be analysed through distinguishing specific types of knowledge; however, the role played by that knowledge is essentially too fluid a concept to manage using a rigid development process. This fluidity is reflected in the level of adaptability accommodated by CommonKADS [99]. This implies that older GT approaches that require pre-defined and fixed knowledge models are flawed because they are too rigid. CommonKADS specifies a number of constituent model-types that, when combined, provide a knowledge engineering solution that meets the needs of any industrial or institutional requirement.

While the aim is to provide some flexibility, Schreiber et.al. note that certain knowledge models can be employed across a range of applications to identify a set of task templates to help facilitate the re-use of models [99].

To manage CommonKADS, it is suggested that specific 'process roles' be assigned to effectively manage human input and human oversight of the knowledge engineering process. Of specific relevance to this work is that Schreiber et.al. clearly see the knowledge provider (expert) as having a key role. The expert is independent from the knowledge engineering process, and whose primary role is to provide the requisite information to develop the necessary knowledge models [99].

Schreiber et.al. acknowledge 'bogus' or non-useful expertise as an issue and emphasise the need to consider the nature of the knowledge provider in subsequent knowledge engineering processes. The issues of bias and other human fallibilities are briefly discussed, followed by an analysis of elicitation techniques including interviews, protocol analysis and repertory grids [99]. In the majority of knowledge elicitation the starting point will be an interview. Unstructured interviews can be used, that allow the expert to express themselves without any formal constraints [99]. Essentially unstructured interviews are an informal scoping exercise that allow the expert to express the salient features of a given domain in a way that is familiar. A structured interview may also take place at a later stage using leading questions and constrained dialogue techniques. Structured interviews allow for more useful expert input in terms of explication into usable knowledge models, however, this inevitably curtails the freedom of the expert to communicate their knowledge in a way that is natural to them [99].

According to Schreiber et.al., a common problem in interviews is that they fail to capture aspects of the domain which cannot be verbalised. Usually this is because instinctive knowledge is built from a rationalisation that is natural to the expert but difficult to 'decompile' in an expressible form [99]. Informal interviews are used in the initial stages of development, while a structured interview is often used to fill in gaps in the knowledge model after a degree of model-building has already been completed. Beyond interviewing, a further method of knowledge elicitation is protocol, analysis which requires the knowledge engineer to observe the expert at work and establish common decision-making processes from which rules can be derived [99]. While the knowledge elicitation process provides ways to obtain knowledge from a single expert, it fails to address the possibility that knowledge elicitation may be achieved through collaboration between groups of experts.

### MIKE

*MIKE* (Model-based and Incremental Knowledge Engineering) is another knowledge engineering methodology that is described in Neubert [76] and later in Angele et.al. [3]. Like CommonKADS, MIKE was developed to address the need to produce large, reliable and maintainable KBS [3]. MIKE specifies that, alongside the models that contain the static knowledge of the domain, a model of generic problem solving-methods should also be developed along with the heuristic rules that govern their application to the static knowledge. The division into two types of knowledge model, constitutes an implementation of the distinction between 'symbol level' and 'knowledge level' as described by Newell [77]. The knowledge level exists above the symbol layer (the static knowledge) and is linked with the idea of rationality, i.e. some form of rationality can be imposed on the symbol level from the knowledge level [77].

In MIKE it is acknowledged that developing a formal specification directly from interviewing the expert is difficult, and therefore it suggests building 'mediating' representations [76] or semi-formal representations that can be understood by the expert. This allows them to be part of the modelling process, particularly in terms of evaluation [3]. A formal representation is then derived that establishes a model for *static*, *functional* and *dynamic* knowledge representation, these contain, respectively, the static, problem-solving and heuristic knowledge aspects [3]. MIKE uses Knowledge Acquisition and Representation Language (KARL) [33] to build the formal model.

#### Towards Ontologies

A recurring feature of modern Knowledge Engineering Methodologies is that they rely on multiple models that provide conceptualisations of different knowledge areas. Understanding what these models are, how they work and how they are devised, provides the flexibility needed for these methodologies to adapt to differing applications. However, the creation of these model-sets (knowledge bases) for each KBS has associated costs in terms of the time, effort and expertise required. These costs have led to an increasing focus on the development of ontologies to describe knowledge domains that can be reused across multiple applications, or extended to meet new requirements. Indeed, while Schreiber et.al. do not discuss ontology in any great detail, there is an acknowledgement that ontology points the way forward to providing the knowledge base that will underpin the next generation of KBS:

> "Ontologies should be seen as a 'natural next step' on the road to more expressive information modelling" [99]

## 2.3 Ontology, Ontology Engineering and Ontology Building

Ontology is a borrowed term from philosophy, describing a discipline that seeks to understand the nature of things — or more specifically, what is the essence of a thing that remains constant even when that thing changes [36]. Cimiano considers computers to be "essentially symbol-manipulating machines [that need] clear instructions about how to manipulate these symbols in a meaningful way"; and ontologies as models that provide these instructions [15]. In more precise terms, Gruber [40] describes how ontology, in its modern sense, is used to facilitate knowledge representation and knowledge sharing. Gruber describes ontologies as an explicit specification of a conceptualisation; and a conceptualisation as an abstract, simplified view of the world that we wish to represent for some purpose [40]. Here, ontologies are explicit documents, used to specify the things that exist in a given domain.

According to Gruber, another important aspect of ontology is *ontological commitment*, whereby agents within an ontology are said to *commit* if all their observable behaviour is consistent with definitions provided by that ontology [40]. Gruber asserts that an ontology should provide a guarantee of behavioural consistency between entities within a given domain[40]. Gruber goes on to state that formal ontologies should be designed according to the following five criteria: clarity of communication, coherency, extendibility, minimal encoding bias and minimal ontological commitment [40]. Coding bias occurs when design choices are made only for the purpose of efficient implementation and should be avoided when possible. Avoiding ontological commitment where it is unnecessary is important because unnecessary commitments may reduce interoperability with other ontologies. Gruber aims to provide a a set of design principles with the objective of improving interoperability between ontologies, it does not provide any explicit methods to actually design and build ontologies.

To see how ontologies are actually developed for use in KBS, the development of ontology engineering needs to be discussed.

> "Ontology engineering refers to the set of activities that concern the ontology development process, the ontology life-cycle, the methods and methodologies for building ontologies and the tool suites and languages that support them." [36]

Gómez-Pérez et.al. provide a comprehensive description of ontology engineering processes and demonstrates how ontology can be incorporated into information systems through ontology engineering [36]. Ontology engineering is concerned with formalising the processes that are required to build a formal ontology. While Gómez-Pérez et.al. provide a comprehensive overview ontology engineering process, it is the chapters relating to ontology building that are of most relevance to this work. In Chapter 3 [36] ontology building methods and methodologies are discussed. This discussion uses case-studies to illustrate the 'classical methodologies and methods' which have emerged in order to build ontologies [36]. For this review a brief summary of some of the presented case-studies is now provided in order to give general context for the approach offered .

Cyc [63] is an early attempt to build a common-sense knowledge base which containing over a million manually entered assertions and for which a bespoke language, *CycL* was devised to allow numerous types of assertions to be made [36]. Cyc has a three stage procedure for acquiring the common-sense knowledge to be modelled; firstly a process for the manual extraction of knowlege; secondly a process for computer-aided extraction; and thirdly a fully computer-managed process for extraction [36]. Cyc can be considered a ontology because it has a communication layer in its design that allows it to be utilised by other intelligence systems [36] that include modules for interaction with database systems, thesauri, agents systems and natural language tools for interacting with WWW [36]. However the fact that an argument has to be made that Cyc provides an ontological resource implies that there may be difficulties in conceptually reconciling it with modern ontologies.

Gómez-Pérez et.al. [36] describe the ontology building method initially described in Uschold and King [112] and later extended in Uschold and Gruninger [111] (to include more detail regarding the scoping phase of ontology building and to specify the role of competency questions in the methodology) as the "first method for building ontologies" [36]. Uschold and King provide a 'skeletal methodology' for ontology building that has four stages [111], identifying scope and purpose of ontology to be built, the actual building of the ontology, evaluating the ontology and, finally, documenting the ontology. The informal part of the scoping phase consists of brainstorming, followed by grouping concepts into naturally forming categories, with the objective of forming viable definitions [111]. Interestingly, no mention is made of who might be performing the brainstorming and categorising, whether an expert is present and, if so, how they communicate with the ontology engineer. Uschold and Gruninger strongly advocate a middle-out (as opposed to top-down and bottom-up) term acquisition process that requires the user to define the fundamental terms used in the domain before defining any related terms which are either more abstract or more specific [111].

Uschold and Gruninger also discuss the broader need for agreement (or consensus) to be made regarding the concept definitions[111], yet while a solution for dealing with term ambiguity is specified, only a set of guidelines [111] are provided for dealing with general disagreement in definition.

METHONTOLOGY provides a more developed ontology engineering methodology that allows for the development of ontologies at a knowledge level [66, 67]. The transmission of knowledge from the expert to the knowledge engineer in this case occurs through the creation of intermediate representations which can be understood by both expert and knowledge engineer alike [67]. The basic idea of producing an initial model from the expert, processing it and then referring back to the expert, is seen in the teach-back technique put forward by Johnson and Johnson [52]. These intermediate representations are built by first constructing a glossary of terms that are used in the domain and obtained through initial consultation with the expert, then from that glossary a concept-classification tree is created that specifies basic class relationships such as inheritance and disjointness. From the concept-classification tree the intermediate representation is derived which consists of a set of attribute tables that describe the properties of domain-concepts, relationships between these concepts, any logical axioms that have been defined as well as any constants and formulas used [67]. An example of an intermediate representation is presented in López-Fernández. et.al. [67]. López-Fernández. et.al. make the claim that the intermediate representations created are understood and validated by the domain expert and the human end-user alike [67].

The knowledge acquisition process employed in the given example of METHONTOLOGY does not differ hugely from that generally used in knowledge engineering (see section 2.2.2) in employing formal and informal interviewing techniques. Here, the expert is consulted as and when the knowledge engineer feels that elicitation is needed [67]. In the description of the METHONTOLOGY implementation from López-Fernández. et.al. the expert is used, not as an integral part of the process, but as a point of reference that can be used to obtain 'clues as to what [the knowledge engineer] were to look for' [67]. While this may be appropriate for obtaining knowledge from experts who are reasonably well-versed in the concept of knowledge representation, it may be compromised if the expert fails to understand the intermediate representation, leading to a flawed validation process. There is also no procedure for using collaborations between groups of experts (unlike the methodology proposed by Uschold and King / Uschold and Gruninger).

The NeOn methodology for ontology engineering, described in Suarez et.al. [106], is an example of a modern approach which puts emphasis on how ontologies are used and reused within the context of an increasingly networked world. So far, there are various instances of successful real-world implementation that have used NeOn [13, 37, 65, 114]. NeOn places much emphasis on defining processes which allow for efficient re-use and interoperability with other ontologies and semantic resources [106]. Perhaps the biggest difference from previous methodologies discussed is that NeOn concentrates on how to build ontologies within the context of the broad networks of interconnected semantic resources in which they will exist. NeOn, like previous ontology engineering approaches, does not define a strict work-flow; instead it defines procedures to follow given particular scenarios. In total there are nine of these scenarios [106] that deal with a range of important and emerging challenges such as re-engineering, ontology merging and the incorporation of ontology design patterns. Only the first of the given scenarios covers the initial specification and modelling that is performed. While it is acknowledged that the ontology specification, which includes the conceptualisation through knowledge acquisition, is an essential aspect of any ontology building endeavour, there are no defined processes in this regard and there is no indication as to how knowledge elicitation should be performed.

From reviewing the evolution of knowledge engineering and ontology engineering methodologies, we can see that the relationship between the domain expert and knowledge engineer has not hugely evolved, and in some of the most recent methodologies is not addressed at all. There has been massive progress in improving processes that enhance the way in which ontologies can be used; however, from reading the literature on ontology engineering methodologies, there is little evidence that the way knowledge is acquired from the expert has changed since the establishment of ontology engineering in the 1990s.

## 2.4 Bottlenecks

For the purpose of this work the two distinct 'bottlenecks' are defined, the *knowledge re-engineering bottleneck* and the *knowledge acquisition bottleneck*. A bottleneck describes, in broad terms, the major challenges that hinder the adoption, refinement and effectiveness of knowledge systems. It should be noted that in some publications these two bottlenecks are considered as a single bottleneck encompassing the challenges present in both [7]. While the re-engineering bottleneck is not directly relevant to this work, in order to illustrate the distinction between the two bottlenecks a brief description is provided.

### 2.4.1 The Knowledge Re-Engineering Bottleneck

Hoekstra describes the knowledge re-engineering bottleneck as "the general difficulty of [facilitating] the correct and continuous reuse of pre-existing knowledge for a new task" [47]. Looking back at the NeOn methodology we can see that of the nine scenarios, for which solutions are defined, five specify situations in which resources need to be either re-engineered or restructured in order to facilitate reuse [106]. This general difficulty occurs for two reasons: firstly, because the proliferation of semantic resources built in an uncontrolled manner that need to be reconciled to some standard [47]; and secondly, because of the scarcity of skilled knowledge engineers. A broader analysis of the knowledge re-engineering bottleneck and how it is being overcome can be found in Hoekstra [47].

In addition to devising more sophisticated methodologies to facilitate re-engineering processes, there has been a move towards collaborative ontology engineering in order to distribute the process thus reducing development time and improving oversight. This has manifested itself in various efforts to develop software tools to facilitate collaborative ontology development. These include *OntoWiki* [5], *WebProtégé* [110] and *Moki* [25] along with various other projects too numerous to list.

### 2.4.2 The Knowledge Acquisition Bottleneck

Of greater significance to this work is the knowledge acquisition bottleneck (KAB). Unlike the knowledge re-engineering bottleneck, the knowledge acquisition bottleneck is a long-standing issue associated with AI and expert systems. Feigenbaum describes the KAB in 1984 in his paper *Knowledge Engineering: The Applied Side of Artificial Intelligence* [31]. Here Feigenbaum divides all knowledge that needs to be acquired by expert systems as being either *facts* consisting of the tangible knowledge that can be easily expressed by being written down or *heuristic* meaning that part of knowledge that ‘constitutes the rules of expertise, the rules of good practice, the judgemental rules of the field [and the] rules of plausible thinking’. The problem, as Feigenbaum saw it, was that the acquisition of both forms of knowledge was a “tedious, time-consuming, and expensive procedure” [31]. In more succinct terms KAB can be characterised as “the difficulty to actually model the domain in question”[16].

One solution proposed to deal with KAB is to find ways of automating the construction of ontologies from data-sets. These ontology learning methods are discussed in the next section.

## 2.5 Addressing the Knowledge Acquisition Bottleneck through Ontology Learning

### 2.5.1 Ontology Learning using Text Resources

Ontology learning is the idea that we can develop ontologies from data-sets that are known to encompass the scope of the target knowledge domain. Described another way, ontology learning is a way to automate knowledge acquisition using pre-existing data. Many ontology learning approaches use natural language processing (NLP) techniques to identify patterns which indicate relationships between terms. While most textual resources on the web do not have any explicit semantic structure that can be directly used with traditional knowledge acquisition, it is possible to glean some of this knowledge by processing the structure of the language itself. Several approaches exist to convert lexico-syntactic patterns into ontology-friendly structures. Lexico-syntactic patterns aim to reduce term ambiguity within text by specifying more restricted contexts in which the term can be defined. Furthermore, it can facilitate the identification of semantic relationships [69].

Some of the earliest work relating to the identification of patterns for enriching semantic content was carried out by Hearst [44] who identified certain textual cues within text (such as A *is a* B or A[,] *and other* B) that indicate a hypernym/hyponym relationship. While these rules were a starting point, more sophisticated approaches have been developed to negate the inevitable false-positives and lack of recall produced through the application of Hearst’s rules.

Snow et al. [103] describe a pattern-based approach that aims to replace the reliance on lexical databases, such as *WordNet* [32, 73], when determining hypernym/hyponym relationships. The advantage of the approach offered by Snow et.al. is that it avoids the computational burden of accessing these lexical resources and would be able to use patterns to define new and unknown hypernym/hyponym relationships emerging from the text. With the use of training data consisting of newswire text, the approach can be used to identify as yet unknown hypernym/hyponym relationships between noun pairs in new texts. In this approach lexico-syntactic patterns that indicate hypernym/hyponym relationships are identified through the learning process and then applied to new text resources [103]. This approach outperformed WordNet's classification when compared to a human-generated list [103] that was used as a gold standard. The approach offered by Snow et.al. has since proved useful when applied to much larger data-sets [72].

Ritter et.al. [95] presented a pattern-based method to find hypernyms within arbitrary noun phrases. The authors ascribe to the goal of "machine reading", which aims to extract information from text to support a range of inferencing capabilities [29]. The focus is specifically centred on a process termed 'ontologizing', in which arbitrary noun phrases are analysed to discover hypernyms [95] that correspond with the class/subclass relationship present in ontology. Ritter et.al. use a support vector machine classifier to find the correct hypernyms with use of an adapted version of Hearsts patterns which excludes matches made where there is a known to be high frequency of returned error [95].

Maedche and Staab introduce an approach that incorporates NLP-based ontology learning into a broader knowledge acquisition process [68]. Acknowledging that fully automated knowledge acquisition from ontology learning is a distant goal, the authors propose ontology learning as a component part of a broader ontology engineering solution where the knowledge engineer plays a supervisory role [68]. With this in mind, a knowledge acquisition architecture is proposed that utilises machine learning to help facilitate the partial extraction of domain ontologies from web resources [68]. The architecture uses components of the *OntoEdit Ontology Engineering Workbench* [107, 108] alongside a resource processing and algorithm library [68]. OntoEdit provides the graphical interface for the knowledge engineer while the resource processing library provides various tools for manipulating the incoming web-resources, including the indexing of HTML documents, identifying explicit and implicit data-structures, and NLP processes that can be applied to the text such as tokenizing [68]. The algorithm library provides a set of resources that facilitate a 'multistrategy' approach in which specific algorithms can be used to acquire the type of knowledge that is required [68]. Using this approach has enabled users to identify likely concepts, cluster concepts hierarchically and determine a set of potential association rules [68].

Rios et.al. [94] present another approach that uses the analysis of domain-related texts to capture explicit knowledge in the form of definitions. This approach works by identifying a given sequence of words that follow any concept which has previously been identified from the text. This work builds on previous work that exploited lexical patterns analysis [12, 44, 85] and augments them using a clustering algorithm and additional contextual information extracted from the Web that improves the discovery of hypernym/hyponym relationships. Additional approaches to obtaining semantic properties from text analysis are provided by Ortega-Mendoza et al. [83] and Sang [97].

### 2.5.2 Ontology Learning using Web Resources

There are also examples of ontology learning that do not rely on natural language processing [21, 34]. Sabou et.al. [96] developed a framework for ontology learning from Web Services. Web services are considered as self-contained, self-describing, modular applications that can be published, located, and accessed via the Web [92]. The framework for knowledge acquisition offered by Sabou et.al. exploits the fact that these sources are expressed in a specific and logically restrained sub-language, making them amenable to automatic analysis [96]. Inevitably within the structure of a web service, some explicit or implicit reasoning knowledge, beyond that which is expressed in natural language, will exist that can be used for ontology learning.

The web also contains more explicit knowledge structures that can be leveraged to build ontologies. Folksonomies are explicit taxonomic structures generated in a bottom-up manner by encouraging system users to categorise their input through tagging activity. Tags allow users to categorise documents and other entities by associating a symbolic label which can then be reused by other users. Websites such as *flickr* and *Delicious* use and generate folksonomies through the tags created by their many users. Folksonomies can be an effective way of mapping users' cognitive collaborative understanding of a domain[28, 56].

While folksonomies do not necessarily have any heirarchical structure or restrictions, they can be enhanced through the use of rule-sets that allow for the basic modelling of concepts [27]. Folksonomies are dynamic because knowledge representations will change through usage, and the emergence of new knowledge will be captured. Moreover, because folksonomies are created automatically there is less need for specialised skills and they can, therefore, be employed in minority subject-domains which would not necessarily appeal to ontology engineers. Folksonomies can either be broad, where tags are created by a large number of interconnected users, or narrow, where tags are created and intended for idiosyncratic use by a particular group or community [27, 28].

The problem with folksonomies is that they are prone to three basic issues: polysemy, where a tag can be interpreted in more than one way, leading to misclassification; synonymy, where two or more tags impart identical or very similar meaning; and granularity, where there is great variety in the level of abstraction of tags [28, 35]. Given these problems and the general lack of semantic coherence [35], ways are being sought to incorporate ontological features with folksonomies with the aim of producing adaptable knowledge management systems. These would provide the flexibility, collaboration and information aggregation of folksonomies, with the standardization, automated validation and interoperability of ontologies [27, 28].

The hybrid systems that have emerged from this research have variously been described as *folksontologies* [113], *semantically enriched folksonomies* [4] or *flexonomies* [53].

In order to build such models, a number of methods have been developed aimed at combining folksonomies with ontologies. These range from using ontologies to augment and replace existing tags within to help improve interoperability and reasoning capabilty [1]; to approaches that use algorithmic methods to construct entirely new ontologies from a set of tags so that identical items could be tagged at different levels of specificity [28].

Abbasi et.al. [1] devised a system called *T-ORG* that automatically organises and categorises tags. This categorisation is done by manually selecting concepts from single or multiple ontologies related to the domain [1]. Concepts in the ontology that are not required are then removed (pruned), redundant and conflicting concepts are refined and the ontology or ontologies are augmented with missing concepts to fill in the gaps, again, this is a manual process. Then using, lexico-syntactic patterns (discussed earlier) concatenated with the tag names to be organised, a call is made to the *Google* search API [1]. By finding terms that the patterns frequently match within the text of the returned Google search results, a good guess can be made as to which concept they belong to in the ontology [1].

Allen and Schneider provide a case-study on how a folksontology can be created [2]. *Cisco Systems, inc.* wanted to create a folksontology to support their internal technical support team. The objective of building such a system was to streamline the response of technicians to incoming support requests by ensuring that these requests were placed correctly in the support infrastructure [2]. An additional goal was to enhance the ability to capture the knowledge from customers and map this knowledge to a pre-existing taxonomic structure so that all support requests could be consolidated into a restricted and standardised vocabulary [2].

Cisco then formed groups of experts for each major technical area, each group being given the ability to manage the tag framework, allowing them to perform tasks including tag removal and augmenting tags with meta-data [2]. Each group would be charged with specifying the set of tags that characterise the content of their area with the assistance of a set of rules to identify good tags to adopt [2]. The tags could then be managed with additional semantic information using a tag management application which used Resource Description Framework(RDF) by specially trained staff [2]. The ontology or ontologies were then pruned and refined so that they could efficiently interoperate.

Acquiring and enriching ontology structures from the analysis of Linked Data (discussed in section 2.9.1) has also been the subject of various research efforts [19, 59, 87, 120].

### 2.5.3 Shortcomings of Ontology Learning

Ontology learning has undoubtedly proved useful in providing genuinely bottom-up approaches to knowledge acquisition [21], especially when used with dynamic data-sets that are constantly changing. In this way ontology learning techniques can be effective in capturing emerging and alternative conceptualisations within a domain [34], especially when used with existing tag-based web resources.

However, ontology learning techniques are limited in what they can do. We see that approaches utilising NLP can be effective in discovering hypernym/hyponym relationships; however, ways to identify more nuanced semantic properties are underdeveloped or lacking completely. Those methods which exploit existing web structures clearly require the input of highly trained individuals to act as knowledge engineers. In the case-study offered by Allen and Schneider a 'taxonomist' is employed to facilitate the construction of the folksontology, while Abbasi et.al. is reliant on significant manual pre-processing, even if the final stages of classification are automated. Ontology learning, in general, offers much promise, particularly in terms of acquiring new and emerging knowledge, however, it is broadly acknowledged that large quantities of data by itself may not be adequate to build complete ontologies without significant human input [88].

## 2.6 Addressing the Knowledge Acquisition Bottleneck through the use of Controlled Natural Language

Denaux et. al. [23] propose an approach to ontology development that seeks to "seamlessly" integrate the development processes that are used to build the conceptual and the logical aspect of the ontology. To do this, Denaux et. al. advocate a holistic approach involving the use of three components; Kanga a methodology for building domain ontologies [60]; Rabbit [42, 43] a controlled natural language (CNL) that enables knowledge encoding in human-readable form; and ROO [24], a user-friendly ontology development tool that uses Rabbit. The approach offered by Denaux et. al. puts the domain expert in control of ontology development by making them use a CNL. This allows them to communicate all the information needed to build a domain ontology in a way that can easily be converted into a logical form. This approach is interesting as, by placing the domain expert at the centre of the process, it requires the consideration of wider issues such as interface design, and how language can be used to bridge the gap between knowledge engineers and domain experts.

This work represents a significant improvement on existing approaches in terms of making ontology development tools accessible, as presenting via a more human-readable language helps domain experts to comprehend the ontology logic. However, it still relies on complex software (ROO is an adaptation of *Protégé*) and there is still a comprehension hurdle that needs to be overcome by the domain expert. Whether a logical problem is presented in symbolic form, or whether it is written in natural language, there is still a required mathematical competency that you cannot expect all domain experts to have.

## 2.7 Addressing the Knowledge Acquisition Bottleneck through the use of Templates

Parreiras et. al. [86] suggest that the use of templates could potentially make ontology development easier. Templates are made up of reusable axioms or statements about a domain, produced by a domain expert in an appropriate form (such as a table, a set of rules or some other structured format), that can then be used by an ontology engineer as a building-block for building an ontology. Parreiras et. al. present a standard way to develop such templates. The approach extends existing metamodels, such as OMG OWL [82], using templates allows for greater level of abstraction which is of benefit to both the domain expert and the ontology engineer.

A limitation of using such an approach is that it still relies on domain experts to have some technical expertise. In the case of the approach offered by Parreiras et. al., for example, a requirement is the use of Unified Modelling Language to build the templates needed. This means that some technical expertise is needed on the part of the domain expert.

## 2.8 Addressing the Knowledge Acquisition Bottleneck through Collaborative Ontology Development

As mentioned earlier, there are various ways to conduct collaborative ontology engineering. There are many tools, including *WebProtégé* [110] and *Moki* [25], to facilitate this activity. These can, of course, also be used to address KAB as they facilitate the input of knowledge in a distributed way, potentially speeding up the acquisition process. Holsapple and Joshi provide a description of the benefits and shortcomings of collaborative ontology building [48]. Unlike standard ontology development environments, in a collaborative approach the final model of the acquired knowledge can be obtained from a diverse range of sources which are iteratively reconciled until consensus is reached regarding the ontological commitments [48]. Of course this supposes that the collaborative dynamic is not undermined by human-behaviours that could affect the oversight of such endeavours. According to Holsapple and Joshi, coordination of the design process may suffer if too many persons are directly involved and to overcome such situations a consensus-building mechanism needs to be employed. [48].

There is still a lot to learn about how collaborative ontology development works, and in particular what evaluation methods can be used. Strohmaier et.al. [104] provides an interesting study of the social dynamics that are present in various collaborative ontology engineering projects, with the objective of preparing the groundwork for a formal evaluation method [104]. Without much of a precedent to work with, Strohmaier et.al. look to provide a quantitative analysis of behavioural trends within collaborative ontology building endeavours, in view to providing a greater understanding and enabling better evaluation techniques.

In general, the tools used to facilitate collaborative ontology development still present a high expertise threshold in terms of their use. If the expert or experts holding the domain knowledge are not well-versed in the fundamental principles of ontology design then they will be unable to use such systems directly, meaning that there would still be a heavy reliance on the knowledge engineer.

## 2.9 Broadening the Semantic Web

We can see from the discussion on ontology learning that much work has been done to perform automated knowledge acquisition over web resources through either NLP or through the analysis of existing web-resources. However, because of the expert input required, in neither of these cases can the knowledge acquisition bottleneck be said to be fully addressed. Furthermore, all the ontology learning approaches outlined thus far are useless if there is no data-set that encompasses the domain. Because ontology engineering is an expensive process, knowledge engineering efforts have been concentrated in a few commercially viable areas. However, if we are to move significantly towards the 'universality' goals of the semantic web, where all information on the web can be understood by machines [8], a way of semantically enhancing a greater range of knowledge domains needs to be found. But do we really need to create resource intensive ontologies to do this? One approach to enhancing semantics on the web has been to develop Linked Data.

### 2.9.1 Linked Data

The term Linked Data refers to a set of best practices for publishing and connecting structured data on the Web [9]. Heath and Bizer provide a description of Linked Data, how it works and how it has been adopted [45]. The basic idea of Linked Data is to use the general architecture of the web to facilitate the sharing of structured data on a global scale [45]. By using Universal Resource Identifiers (URIs) to reference not just web pages, but also representations of real-world objects and abstract concepts, Linked Data provides basic semantic resources that can be easily accessed through the HyperText Transfer Protocol (HTTP) [45]. The Resource Description Framework (RDF) [57], a simple graph-based data model, is used to describe these concepts as well as specifying the links between these concepts and annotated web resources. RDF's core data model is the *subject*, *predicate*, *object* relationship, commonly referred to as triples [45] (See figure 2.1 for an illustrated example of triples in use).

In broad terms, Linked Data and the intended result of its adoption is described by the following quote:


FIGURE 2.1: Basic RDF structure describing 'Eric Miller' from the RDF Primer [70]

> "Just as hyperlinks in the classic Web connect documents into a single global information space, Linked Data uses hyperlinks to connect disparate data into a single global data space" [45].

RDF provides Linked Data with a simple and unifying data model, while the use of established technological architecture allows for easy access — all of which means that linked data establishes a better platform for data discovery and for self-description than traditional web [45]. Often, but not always, the conceptualisation of Linked Data is provided by an ontology that determines the abstract concepts and properties.

The success of Linked Data can be seen in its widespread adoption both in business and governmental domains. There are numerous examples of Linked Data being used in various contexts [14] and there are now some well-established Linked Data repositories and websites that use Linked Data including *DBpedia*[1] [62], a semantically annotated version of Wikipedia, *Geonames*[2] a geographical database and the *BBC* website [3] [58]. Many Linked Data resources can be found within a broader network called the Linked Open Data Cloud[4]. There has also been a concerted effort by government agencies to provide Linked Data. Hendler et.al. are generally positive about governmental uptake of Linked Data claiming that it has been embraced by the US Government and, as such, plays an increasingly important role in government information sharing [46]. Shadbolt and O'Hara provide a more mixed review of the state of Linked Open Data efforts by the UK government [100]. While praising much of the work that provides access to Linked Data, they acknowledge that uptake in some areas is slow due to the 'complex combination' of administrative and technical tasks needed to be performed, meaning

[1]http://dbpedia.org/
[2]http://www.geonames.org
[3]http://www.bbc.co.uk
[4]http://linkeddata.org/

that Linked Data supports only around 5% of available data on *data.gov.uk* [100].

Linked Data, no matter how plentiful, or how high the quality, is no substitute for the development of ontologies. Bechofer et.al. looked at how Linked Data is used in research environments in general and concluded that while Linked Data help to publicise the results of publication data, it cannot enable reusing, sharing or reproducing of research without the use of additional resources [6]. Jain et.al. goes further by claiming that Linked Data is only of limited value in building the semantic web [51]. The shortcomings of Linked Data according to Jain et.al. are that there is a lack of resources that provide a conceptual description, a lack of schema mappings and a general lack of expressivity [51]. Jain et.al. also suggest that the over-reliance of Linked Data libraries on upper-level ontologies like DOLCE [71] and SUMO [78](ontologies that provide the most-abstract conceptualisation), means that the level of conceptualisation may be inappropriate for making meaningful linkages between data-entities [51]. Moreover, Linked Data emphasises the need to enrich data rather than acquire knowledge. It can't be used as a means of determining domain knowledge. In fact, to make best use of it, domain knowledge needs to be pre-specified in the form of vocabularies, schemas or ontologies. While Linked Data has been effective in broadening the availability of semantic resources, the semantic qualities that it communicates are relatively weak.

## 2.10 Knowledge Acquisition through Crowdsourcing

From the discussion so far, it can be seen that the knowledge acquisition required to build domain ontologies is difficult to achieve without knowledge engineering skills; that ontology learning techniques cannot yet build the fully-functional ontologies required of the semantic web; and that the high-quantity, basic knowledge conceptualisation offered by Linked Data is insufficient in key areas. Therefore, if the KAB is going to be addressed we need a solution that avoids an over-reliance on qualified knowledge engineers, whilst being powerful enough to capture the knowledge required to allow a good level of automated reasoning.

## 2.11 Crowdsourcing

One approach advocated, is to use crowdsourcing as a means to acquire the knowledge needed to build ontololgies. Crowdsourcing is the idea that the web can facilitate the selection and aggregation of useful information from large numbers of internet users [20]. The qualification to contribute to such systems is set low in order to encourage as broad a contribution as possible. The central premise of crowdsourcing is that the sheer quantity of data that can be collected compensates for the lack of quality assurance that would be provided by an expert-led system. However, this is not to say that crowdsourcing cannot produce reliable and accurate results, but in order to do this, some form of user-mediation is needed so that erroneous (and even malicious) data is not incorporated.

If we take the example of *Wikipedia* — by far the most used crowdsourcing platform — we can see that there is a body of research attesting to its accuracy and reliability [10, 26, 54, 64, 75, 117]. Given the scale of Wikipedia, this might seem remarkable; however, the key to the success relies not on the magnitude, but on the motivational dynamics of the crowd, and the mediation mechanisms available to them.

Panciera et.al. analyses the behaviour of contributors to wikipedia and notes that there is a core of users, or 'Wikipedians', who contribute, mediate and safeguard against malicious input in the majority of cases [84]. Preidhorsky et.al. notes that if Wikipedians are the contributors who make up the top 10% of users as judged by number of edits, then their contribution is 85% of the total edits[90]. These key contributors take on a group ethos and adopt informal rules and routines that are important in producing a consistent knowledge resource [84]. Further work on the motivational factors is provided by Nov [79], who noted that the leading motivations for these contributors were that they felt an ideological need to contribute, but most of all, because they saw it as fun.

### 2.11.1 Motivating Crowds to Perform Tasks

The traditional incentive of monetary gain can be used to motivate crowds to perform tasks, a process often referred to as 'microtasking'. This is seen in Amazon's *Mechanical Turk*[5] service, that allows crowds to perform tasks in return for a small payment. This kind of crowdsourcing, sometimes termed crowd work, is becoming more popular, and is seen by many as a trend that points the way towards how work will undertaken in the future [55].

Building ontologies may not be many people's idea of fun, it may not even be many knowledge engineers' idea of fun. However, as Wikipedia has proved, making something fun could be the key to motivating crowds to conduct knowledge acquisition tasks.

In 'Designing Games with Purpose', von Ahn and Dubbish introduce the idea that valuable tasks can be performed by crowds as a 'side-effect' of the primary activity they are engaged in [115]. The concept of Games With a Purpose (GWAP) is that computation can be a bi-product of gameplay, and it has proved very successful — as demonstrated by the success of reCAPTCHAG. This uses a security form that validates human users to identify a scrambled word along with a word taken from printed copy, thus providing a mechanism for digitising printed copy.

## 2.12 Using Crowdsourcing for Ontology Engineering

Noy et.al [80] look at the general feasibility of using crowds to perform ontology engineering tasks. They evaluate the performance of *MTurk* crowds, against that of a pool of domain experts in validating the conceptualisation. The evaluation is concerned with validating superclass/subclass relationships; the authors describe this activity as hierarchy-validation.

[5]https://www.mturk.com/

The authors perform this evaluation by repeating an ontology validation experiment from Evermann and Fang [30], where two ontologies were used (SUMO [78] and the Bung-Wand-Weber (BWW) ontology [39]) to extract natural language questions that could be presented to 32 paid student volunteers. In Evermann and Fang, the questions were manually constructed and could be answered with a simple true or false response.

The *MTurk* crowd user had to pass a basic test in order to qualify to participate, requiring correct responses to 8 out of 12 high school-level biology questions. There were also some safeguards against SPAM, a known issue with *MTurk*; whereby any result-set with 23 identical answers out of 28 was disqualified. Additionally, a bonus was paid to those users who contributed 75% correct answers when compared with a gold standard.

The results of the experiment presented by Noy et.al show that the *MTurk* crowd was less accurate than the domain expert (66.7% compared to 81.2% from the Evermann and Fang experiment). When given term definitions with each question, the crowd performed significantly better with an accuracy of 81.8%, as compared with the experts who returned 88.5% accuracy.

These results would suggest that crowdsourcing is a viable approach to validating ontology hierarchy structures, provided the crowd is given additional context with which to make decisions. Given the speed and availability of the crowd as compared to domain experts, this approach is attractive to ontology development efforts with limited resources.

### 2.12.1 Crowdsourcing for Ontology Alignment

Sarasua et.al. introduce *CrowdMAP*, a model for using 'microtasking' input to perform ontology alignment through the use of paid crowds [98]. Ontology Alignment, or Ontology Matching as it is sometimes called, is the process of making corresponding links between concepts across ontologies. The authors argue that there are four key benefits to employing crowds for ontology alignment. Firstly, humans can validate links without having to be provided with much context. Secondly, the task of verification can be broken down into atomic tasks that correspond to individual mappings, making them easier to deal with. Thirdly, even though ontologies can be large, crowdsourcing can cope with the scale of processing that is needed. Finally, ontology alignment is not a process that can be fully automated at present; incorporating crowdsourcing may help augment existing and future machine-driven approaches.

In Sarasua et.al. the research questions being addressed are:

1. Is ontology alignment amenable to microtask crowdsourcing?

2. How does such a human-driven approach compare with automatic (or semi-automatic) methods and techniques, and can it improve their results?

3. What correspondences between elements of different ontologies can be reliably identified via crowdsourcing?

While acknowledging that Games With a Purpose (GWAP) [116] might also be appropriate as a motivator, the authors opt to use a crowdtasking platform called *crowdflower*[6], which presents the mapping tasks to numerous fee-paying crowdtasking services, including *MTurk*.

The *CrowdMAP* workflow, in brief, has the following stages. Each concept mapping is presented to the user (member of the crowd), who can validate the mappings according to a selection of relationships, namely *equivalency* (where the concepts are the same), *subsumption* (where the concepts have a superclass/subclass relationship) and *merynomy* (where the concepts have constituent part-of relationship). Validation occurs for equivalence relationships if the task produces only equivalencies, i.e. if the first three results produced from the microtask are asserted to be equivalent. When other types of relationship are validated, or when there is a mixed response to the task, a greater degree of confidence is required.

Sarasua et.al. evaluate the performance of their approach by comparing the results of the crowdsourcing approach with mappings provided by the Ontology Alignment Evaluation Initiative. From this comparison, a precision and recall can be measured. Where the crowd was presented with a full set of mappings from a known ontology alignment, the approach achieved a 100% recall, indicating that the crowd was able to provide meaningful input for all mappings. This also proves that the crowd could easily deal with the scale of computation required. While the precision was less, the overall results show some promise and compare favourably with other alignment methods.

Ontology alignment is not directly part of the knowledge acquisition process and therefore this approach does not address the Knowledge Acquisition Bottleneck. Nevertheless, the value of CrowdMAP for the purpose of this work is that it demonstrates that crowds, with unknown expertise, can effectively make judgements on concept and relationship validity.

### 2.12.2 Crowdsourcing for Knowledge Acquisition

So far, the examples given show that crowdsourcing can be used for ontology engineering tasks outside of knowledge acquisition. The following examples show approaches that address aspects of the knowledge acquisition process.

Siorpaes and Hepp provide an early attempt to use games to facilitate ontology engineering tasks [102]. The authors present *OntoGame*, a platform for performing various ontology engineering tasks. *OntoGame* is intended to address what the authors term "the Incentive Bottleneck" by providing an entertaining and competitive means of validating ontology relationships. With reference to the Ushold and King/Gruninger methodology [111, 112], the tasks where game motivation could be used are identified. These include:

- Collecting named entities to describe a domain

[6]http://crowdflower.com/

- Associating entities with specific types (instances, classes, properties etc.)
- Building taxonomic structures
- Ontology alignment
- Dividing ontology structures into modules
- Lexical enrichment through meta-data
- Ontology population[7]

The prototype game presents pairs of users with a *Wikipedia* article generated from the random article function. Then the users are asked to make a judgement as to whether it is an instance or a class (concept). If both agree on the definition, then they are rewarded with points. If they are in agreement, then they can obtain a bonus point if they can both suggest the same super-class.

Of the nine participants for this prototype, seven were educated to degree level or higher, five in Computer Science and two being ontologists. It would be fair to say that this was a fairly qualified crowd. This means the results should not be seen as truly indicative of typical crowd input.

The authors acknowledge that this is early work; however, the results showed that the consensus is generally reached (see table 2.1).

| **Description** | | |
|---|---|---|
| Consensus reached on class/instance choices | 103 of 116 | 88.79% |
| Consensus reached on correct class/instance choices | 102 of 103 | 99.03% |
| Consensus reached on super-class/instance-of-relations | 62 of 67 | 92.54% |

Table 2.1: Results from the OntoGame prototype experiment

*OntoGame* is more significant for testing the user experience of participants, and although it is a qualified crowd, this proved generally positive. *OntoGame* is a promising start to incorporating gamification into the ontology development process and demonstrates a method which could be adapted to knowledge acquisition. However, at present it appears not to have been developed much beyond this basic study.

Good and Wilkinson propose a method for carrying out ontology engineering tasks using volunteers [38]. Good and Wilkinson use seed ontologies which they derive from ontology learning techniques. The data-set used for the ontology learning in the example given was the Medical Subject Headings (MeSH) thesauri. The authors estimated that there was considerable error ($\tilde{4}0\%$) in the ontology learning process due to the difficultly in the mapping of ontology *owl:subClassOf* relationships. The high error rate was attributed to the design of the thesauri which includes some relationships that can easily

[7]Not strictly an ontology engineering task, but included anyway as it is an important task that could use game dynamics.

be misinterpreted. MeSH includes some reasonably sophisticated relationships, such as meronymy, where a concept is a constituent part of another, along with subsumption and degrees of similarity. All of which may be misinterpreted by the learning algorithm employed.

Good and Wilkinson use volunteers to assert (*true*, *false* or *don't know*) the validity of the relationships produced through the ontology learning process. They initially use majority voting to validate, whereby a simple majority decision was used to determine the valid option. They also tried a system where the time taken between answers is used as a crude measurement of trust, giving greater weight to fast answers. Three additional machine learning algorithms were also used to modify the weight attributed to each volunteer's answers, these were; 1R, Support Vector Machines and Naive Bayes.

In all, 25 volunteers were recruited. As there was no lower limit in the number of questions that could be answered, the distribution exhibited a characteristic long tail pattern, with most answers came from a few volunteers while most contributed a small number of answers. In fact, only 5 volunteers responded to more than 25% of the questions.

The results were compared against a manually created gold standard. The evaluation for each method deployed produced a simple accuracy measure, along with two F-Measures based on the average precision and recall of correctly asserted 'true', and correctly asserted 'false answers'. The results are not particularly conclusive; however, they do appear to show superior results from machine learning approaches as compared to simple majority or weighted majority approaches (See table 2.2).

| **Aggregation Method** | **% correct** | **F-false** | **F-true** |
|---|---|---|---|
| A Single Volunteer | .62 | .17 | .75 |
| Majority Vote | .64 | .23 | .77 |
| Time Weighted Vote | .63 | .47 | .71 |
| 1R | .71 | .56 | .78 |
| SVM | .75 | .64 | .78 |
| Naive Bayes | .75 | .64 | .81 |

Table 2.2: Results table from Good and Wilkinson

Good and Wilkinson acknowledge that this is a preliminary study, stating that more work needs to be done on improving the incentive strategy to encourage greater volunteer contributions.

## 2.13 The Potential of Online Communities

The advent of the web, and the subsequent technological advancements that have led to Web 2.0, has provided a platform for the development of online communities. Online communities have removed the need for members of a community to be based in the same locality. As a result these communities have become increasingly based on personal relationships, shared interests and a sense of empathy [89, 93]. It has been suggested that a precise definition of an online community is of less importance than who makes up the membership and what its function is [89].

An online community can be seen as being a qualified crowd, in that while the level of expertise of each member is unknown, the formal and informal membership criteria will provide a minimum expertise threshold for participation. For this study, the term 'community crowds' will be used to describe a crowd in which the precise expertise of each user is unknown but where membership of a particular online community provides qualification. Community crowds undoubtedly exist: they take various forms, have different functions and are often placed within the context of broader networks of communities such as *Facebook* or *Wikipedia*. For example, the contributors, moderators and guardians of a Wikipedia article could be seen as an online community that exists within a broader online community of *Wikipedia* contributors. What links the members in a community such as this is a set of interests, abilities and motivations. A contributor will contribute because they feel they understand the topic, because they are enthusiastic towards communicating this knowledge [119], because they want to enhance their own reputation within the community [61] or because they want to protect the integrity of the article [10, 54, 64].

To bring this back to the context of knowledge acquisition for domain ontology development, we can see that online communities offer six key potential benefits:

1. The collective knowledge of the membership will likely encompass the scope of the domain

2. There is a pre-existing set of motivations that, if harnessed, can overcome the incentive bottleneck

3. Online communities are ubiquitous, covering a breadth of knowledge that goes way beyond that which has already been semantically described

4. Large-scale online communities can offer the scalability benefits typical of crowdsourcing

5. As with folksonomies, emerging and changing conceptualisation can be captured

6. A greater degree of expertise can be assumed than in unqualified crowds

While online communities take many shapes and forms, it is possible to identify those communities that encompass a knowledge domain. For example, the contributors to the movie database *IMDB*[8] represent an online community whose knowledge encompasses the movie domain. Smaller online communities also exist that encompass knowledge from smaller domains such as local history projects, or niche interests such as arts and crafts or outdoor pursuits. The big question is, can we harness the dynamics within

[8] http://imdb.com

these communities to acquire knowledge for ontology building?

We could employ game dynamics; however, this would be a contrived approach and would rely on the memberships of those communities being interested in playing such games.

A better approach might be to incorporate the process of eliciting knowledge into activity that the membership are going to perform anyway. For example, if the community is already performing data collection activities, then incorporating some of the techniques used to verify semantic concepts and properties, as described earlier in this chapter, could be an effective way to build semantic resources. The rest of this thesis describes an approach to acquiring concept hierarchies and object properties that can be incorporated into routine online community activity.

## 2.14 Summary

This chapter has outlined the process of building KBS. The importance of developing ontologies as component parts of KBS is stated, and ontology engineering (the process by which ontologies are designed and built) is discussed. The major challenges, or bottlenecks, are defined along with the case for developing new knowledge acquisition methods to address these challenges. Automatic generation of ontologies through the use of learning algorithms is investigated and the shortcomings of these approaches is stated. The need to spread the use of ontologies to broaden the semantic coverage of the web along with an evaluation of Linked Data — so far the most successful approach to enhancing the coverage of semantic capability across the Web. Finally, crowdsourcing is advocated as a possible way forward to acquire the knowledge needed to support the expansion of ontology coverage. Current approaches that incorporate crowdsourcing into ontology engineering tasks are then looked at. While progress has been made, much of this effort aims at validation and alignment tasks that are unrelated to knowledge acquisition. Approaches are identified that could be adapted to perform knowledge acquisition processes. *OntoGame* demonstrates that a gaming mechanism could be used to incentivise crowds to perform ontology engineering tasks, but that approach is underdeveloped. Good and Wilkinson offer a more developed approach; however, it is aimed at augmenting and validating ontology learning processes and, as such, is still reliant on domain knowledge resources being available. The use of online communities in a crowd context is then discussed, and potential benefits are identified.

While crowdsourcing has increasingly been used in knowledge engineering processes, it has not been used for directly eliciting domain knowledge. There is a tendency to view the scale of crowds as the primary benefit of crowdsourcing, however in online community crowds there are additional benefits that used to assist the knowledge acquisition process. Online communities can be identified that encompass specific domains, provide *de facto* pre-qualification in terms of expertise and, if the knowledge acquisition

process can be incorporated into routine online community activity, then the incentive bottleneck can also be addressed.
In the next chapter an approach to eliciting knowledge from crowds that can be incorporated into online community activity will be presented.

# Chapter 3

# A Crowdsourcing Approach to Knowledge Acquisition from Online Communities

## 3.1 Obtaining Consensual Models from Community Crowds

### Notation

This chapter presents protocols for performing mediation on crowdsourced ontologies. The protocol is described using a series of diagrams, each defining a specific process. All references to the protocol diagrams are placed in brackets '( )', with the relevant diagram referenced at the beginning of the paragraph. These references may include state transitions (e.g X→Y) as well as references to the symbols used (e.g. $\alpha$). The term 'concept model' refers to the taxonomic positioning of concepts as defined by each participant. The term 'knowledge model' is used to refer to general knowledge representations, which may include ontologies.

### 3.1.1 Novel Approach

As has been noted previously, attempts to harness the full potential of crowdsourcing as a way of building ontologies are few and far between. This work proposes an approach to ontology building that exploits crowds to provide the required expertise. This expertise is based on a consensus view of the domain obtained by mediating the crowd input. This approach may also be used to support new knowledge acquisition processes, this is explored in Experiment 3 (See Chapter 6) where a map interface is used to obtain crowd input. The approach described in this chapter aims to reconcile ontologies containing disparate concept hierarchies so that agreed-upon concepts will emerge that, when combined through the process of mediation, will provide a consensual knowledge model of the target domain. This approach was devised by considering what anonymous crowds (where expertise is not known) are capable of and then adapting these abilities to the process of knowledge acquisition and ontology engineering; therefore it is different from existing distributed approaches which aim to connect knowledge engineers and other experts. While some development is needed to realise the full potential of this approach, the experiments conducted show that it can be used successfully, providing the possibility that it can be used as a potent tool in overcoming the knowledge acquisition bottleneck. While elements of crowdsourcing and other collaborative technology have been used to address the knowledge acquisition bottleneck, this approach is the first to use crowdsourcing as a primary device for addressing this problem.

### 3.1.2 Objective

This approach consists of a set of simple and adaptable protocols for eliciting concept hierarchies that can potentially be incorporated into routine online community activity. As focusing knowledge acquisition onto online communities is a novel approach, initially these protocols are going to be used for the acquisition of concepts. Ultimately this approach could be used to obtain more sophisticated aspects of knowledge models, such as object properties and semantic restrictions; however, more work is needed in to achieve this, particularly in terms of finding ways to efficiently exploit crowds in order to acquire usable input. Therefore, the objective is to see if we can gain a consensual knowledge model consisting of generally agreed-upon concept models that have been acquired from a community crowd through the employment of the approach offered. Gaining a consensus is the first stage of identifying a community crowd that can be used to obtain a useful and consensual domain representation. Obtaining a community crowd consensus will provide a knowledge model in which a degree of confidence can be attributed, making it suitable for further development by knowledge engineers.

It should be noted that even a lightweight ontology can have a sophisticated structure, and at this stage only the most basic elements can realistically be acquired from the crowd. Therefore no complete model is likely to emerge, and any acquired consensual model must be further adapted to meet the needs of any application that wants to make best use of this knowledge model. Deriving crowd knowledge using the approach outlined in this chapter would still be of great benefit, as alternative conceptual models may emerge that could enhance our understanding of the domain. Moreover, by incorporating this approach into online community tasks, there is the potential for automatically developing useful semantic resources that can be adapted and developed at a later stage.

### 3.1.3 General Design

The approach starts by allowing users to generate a set of basic ontology structures for a domain. These are then mediated, with the objective of adopting agreed-upon concept models and removing the marginally supported concept models that are likely to be invalid. Figure 3.1 provides a high-level overview of the crowdsourcing mechanism.

If a seed ontology is provided by a knowledge engineer it is extended by participants who contribute concept models. If no seed ontology is specified, then the participants have free reign over how they specify their concept hierarchy within their ontology. A more detailed discussion on seed ontologies can be found in Chapter 6 where an experiment is carried out to determine the most appropriate seed ontology to use. The mediation process is triggered once an adequate number of entities is acquired (A). The 'adequate' number will be specified either by the knowledge engineer who is overseeing the process, or simply by defining a time threshold that, when exceeded, triggers the mediation process. The volunteered ontologies are then processed to discover concept model matches (at this stage this is based on the labels given to them by the participants, however, more sophisticated ways of determining matches are possible; this is explored in Section 5.3.3). These matches are a good indicator of conceptual equivalence; however, to be more confident of this equivalence we must look at the concept model that has been specified. If the concept models of these matches differ (meaning that a concept conflict exists) then a consensus should be sought on the correct concept model. To this end, the set of concept models which are likely to represent identical concepts are mediated. The first stage of mediation (B) is fully automated and resolves conflicts where a broad agreement threshold (majority adoption threshold) is met. The secondary stage (C) is where the conflict requires further human mediation (referred to in this thesis as *semi-automated mediation* as it uses manual input in conjunction with automatic adoption rules); if this fails then a domain expert or knowledge engineer will act as arbitrator (D).

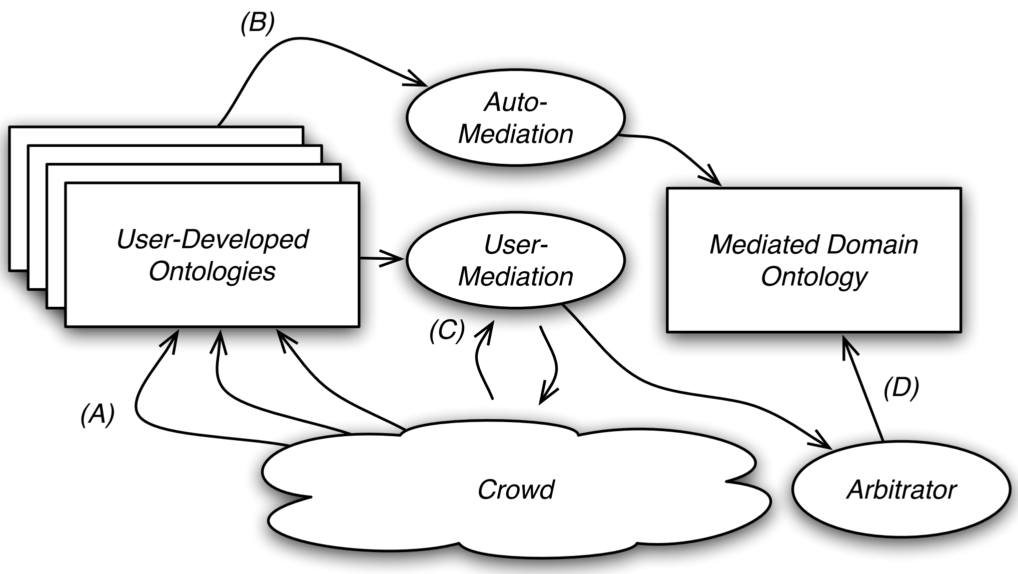


FIGURE 3.1: General Overview of Crowsourcing Mechanism

For each concept processed there are four possible outcomes:

1. Majority adopted: fully adopted across the ontology set and overriding any matching concept that has a conflicting concept model

2. Minority Adopted: consensus reached across some ontologies and is therefore retained, but not adopted across the ontology set

3. Conflicting: existence of concept supported (at least to minority adoption), but the concept models are in conflict.

4. Pruned: entity did not acquire enough support, but will be manually verified to see if it should be retained.

In minority adoption cases (which includes most conflicting cases), adjudication from the knowledge engineer will be required. Adoption thresholds should be set according to the nature of the crowd as these will largely determine the quantity of interaction that is required.

### 3.1.4 Concept Model Conflicts

Figure 3.2 illustrates an example of a concept model conflict. Because the concept of 'campus' has different parent concepts in each knowledge model, a conflict exists. Note that this does not indicate that one concept model is more valid than the other, neither does it exclude the possibility that both could be valid. In this example, a likely outcome is that neither concept model achieves majority adoption through manual mediation; this gives the knowledge engineer scope to incorporate both, perhaps by making region a subclass of area. Conflict detection in this approach is as much about identifying areas that need expert input as it is about removing invalid concepts.

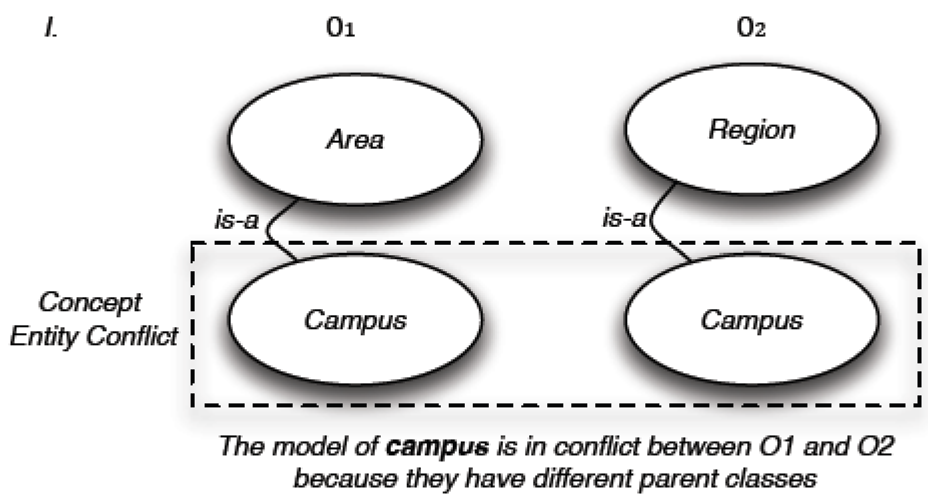


FIGURE 3.2: Example of concept model conflict

### 3.1.5 Mediation of Concepts

Automated mediation is performed with the objective of converging the ontologies towards a consensual model. This section outlines the specific rules that govern exactly where and when ontologies are modified. The main goals are to identify potential concept matches; to override minority concept models where a set number of the acquired ontologies agree upon an alternative concept model; and finally, to identify conflicts which human input (either through semi-automated mediation or through expert adjudication) will be required to resolve. Table 3.1 specifies the particular rules used to perform automated mediation.

| Stage | Rule |
|---|---|
| Formatting | **De-capitalise:** Entity names that are not predefined have their first character converted to lower-case. **Composite Names:** Entity names that contain multiple words are converted to lower-case strings with spaces being replaced with a underscore character (_). Where the user has used camel-case the initial capital is made lower-case and an underscore is inserted. **Edit Distance:** Where entities have names that are one edit-distance apart they are assumed to be equivalent. All subject entities are made uniform with the label identified first being adopted. |
| Automated Mediation (I) | **Majority Adoption:** If a concept model meets the majority adoption threshold across the ontology set, it is universally adopted. **Minority Adoption:** Where an identical concept model meets the minority adoption threshold across the ontology structures but does not meet the majority threshold, the concept model is retained, but not adopted across the set. Minority adoption cases are sent to manual mediation to see if they can be majority adopted, however if this fails, then the concept model is earmarked for expert adjudication |
| Automated Mediation (I) | **Pruning:** All concept models that fail to meet the minority adoption threshold are temporarily removed so that they can be verified or rejected by the participants during manual mediation |

Table 3.1: Rules for automated mediation of concepts

#### Stage 1: Formatting

The first task in handling the data is to identify where concepts are identical across the ontology set. Here, an attempt is made to match the labels of each entity, creating a set of entity matches [81]. While matching labels is simplistic, it still provides a good indicator of equivalence [50]. Going forward, simple string-matching would have to be replaced with a smarter, context-aware solution for determining identical concepts. In order to match strings, and because we are dealing with the crowd, we must assume that there will be differences in formatting along with the inevitable spelling inconsistencies.

A list of typical errors that may be present in any acquired ontology structure set are summarised in table 3.2.

| **Labelling Error Type** | **Example** |
|---|---|
| Simple Label Mismatch Error (case) | *label — Label* |
| Simple Label Mismatch Error (spelling) | *colour — color* |
| Simple Label Mismatch Error (composite words) | *taxiCab — taxi_Cab* |
| Complex Label Mismatch Error | *TV — television* |
| Synonyms | *record — album* |

TABLE 3.2: Label Mismatch Error Examples

There are various ways in which complex label mismatch errors or synonyms can be identified and automatically dealt with (see Miller and Hristea [74] for example), however in two of the three experiments described (See Chapter 4 and 6) there is no deliberate attempt to do this. Simple label mismatch errors, however, can be addressed using the *de-capitalise*, *composite* word and *edit distance* rules listed in table 3.2. In Experiment 2, *WordNet* is used to detect some complex label mismatch errors.

**Stage 2: Automated Mediation I (Adoption)**

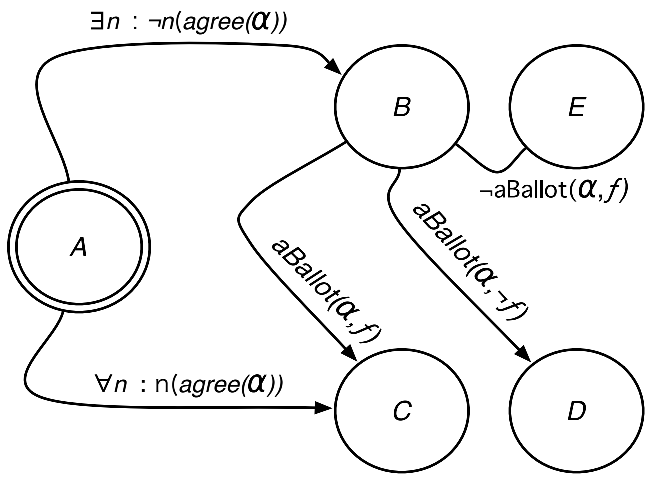


FIGURE 3.3: Automatic Concept Mediation Protocol

Automated concept mediation is the first step taken after formatting and is outlined in figure 3.3. The concept mediation protocol is used to mediate over individual concept models ($\alpha$) that have been acquired from our crowd. If all of the participants ($n$) who have specified identical concepts are in agreement over the concept model, then there is no need to mediate (A$\rightarrow$C). If any conflict exists in the concept model, then the concept model is auto-balloted (A$\rightarrow$B). Auto-balloting is a term used here to relate to the semi-automated mediation process where a human ballot takes place. In effect, the auto ballot is a simple calculation of the proportion of models that are identical. If, after the automatic ballot, a threshold ($f$) for majority adoption is met, meaning that a set proportion of the constituent models are consistent with each other, then the concept model that meets the threshold is approved (B$\rightarrow$C) and the contradictory entities are removed and replaced (B$\rightarrow$E). If the threshold is not met then the conflict is earmarked for further mediation (B$\rightarrow$D) where the concept model validity is voted on by all the participants (minority adopted).

## Adoption Rules

At this point any concepts that have identical labels but have different concept models (those concepts in conflict) have either been majority adopted (with the minority concept model removed) or earmarked for manual mediation. The proportion of ontologies that are required to agree in order for majority adoption to occur is referred to as the *majority adoption threshold*. The majority adoption threshold can be adjusted according to the characteristics of the participatory crowd; typically a larger crowd with more disparate knowledge would have a lower majority adoption threshold than a smaller expert crowd. If a concept model is supported above the *minority adoption threshold* but does not meet the majority adoption threshold, then it is retained within the ontologies that it was originally specified in and is earmarked for further mediation to try and gain a broader consensus. As with the majority adoption threshold, the minority adoption threshold should be lower in less controlled and broader participation crowd scenarios.

### Stage 3: Automated Mediation II (Pruning)

Pruning is an additional measure aimed at removing individual concepts that are only supported by a proportionately low number (determined by the pruning threshold) of acquired ontology structures. These concepts do not enter into the first stage of automated mediation as there is no known conflict to mediate (as no conflicting concept with the same name exists) yet the support for these concepts has not met the adoption thresholds. Pruning acknowledges that some useful models could be lost in this stage. Any concept removed at this point is presented to all the users during semi-automated mediation. If agreement is reached, the concept could be reinstated or even majority adopted.

**Example of Majority Adoption and Pruning**

The following is a brief example of the how adoption thresholds might be applied to a set of crowdsourced ontologies.

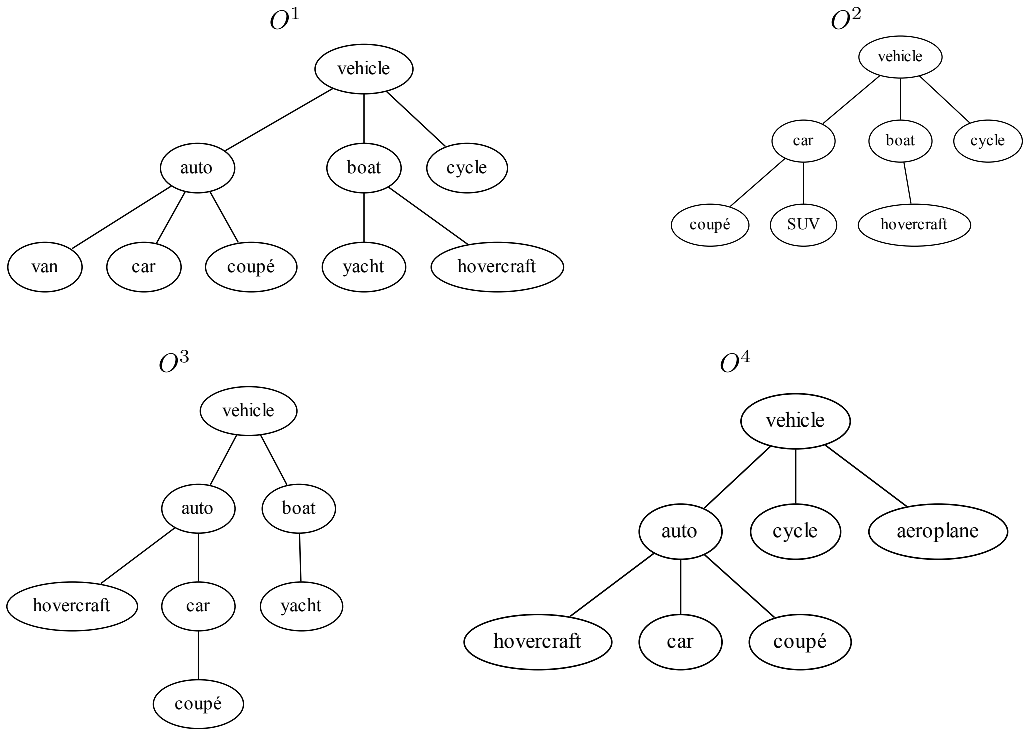


FIGURE 3.4: Mediation example: concept hierarchies pre-mediation

Figure 3.4 shows four 'ontology structures' (in reality simple concept hierarchies or taxonomies, as we are not using object properties or other elements typically present in an ontology). Each of these ontology structures represents an individual contribution from the crowd. The ontologies are labeled from $O^1$ to $O^4$ and contain four differing knowledge representations of the vehicle domain. For this example, the minority adoption threshold is set at 33% and the majority adoption threshold at 66%. After applying the automated mediation protocol and using the specified adoption thresholds, the following outcome is achieved:

| | |
|---:|---|
| Majority Adopted with no conflict | {auto, boat, cycle, vehicle} |
| Majority Adoption after resolved conflict | {car} |
| Minority Adopted | {yacht} |
| Conflicting | {hovercraft, coupé} |
| Pruned | {van, SUV, aeroplane} |

The following is a breakdown of why each concept is classified in such a way during mediation process. The statements in square brackets refer to Figure 3.3 in Section 3.1.5. *Auto*, *boat*, *cycle* and *vehicle* are majority adopted because they are supported by at least 66% of the ontologies and there is no conflict in the concept model [A → C]. *Car* is also majority adopted, the minority concept model in $O^2$, which sees *car* as a direct subclass of *vehicle*, is overwritten by the majority assertion (which meets the majority adoption threshold) that *car* is a direct subclass of *auto* [A → B → C for *car* $\sqsubseteq$ *auto* and A → B → E for *car* $\sqsubseteq$ *vehicle* ]. *Yacht* is minority adopted as it meets the minority adoption threshold but does not meet the majority adoption threshold [A → B → D]. *Hovercraft* and *coupé* remain in conflict as two competing concept models exist in the ontology set, neither of which meet the majority adoption threshold [A → B → D]. Finally, *van*, *SUV* and *aeroplane* are all pruned (temporarily removed) as they do not exceed the minority adoption threshold [A → B → D].

### Stage 4: Semi-Automated Mediation

Semi-automated mediation uses human input in conjunction with threshold-based automated decision-making processes, to validate or eliminate concepts emanating from the crowd. The objective of semi-automated mediation is to reduce, as much as possible, the number of concepts which will need to be addressed by the knowledge engineer.

### Stage 4a: Reinstatement

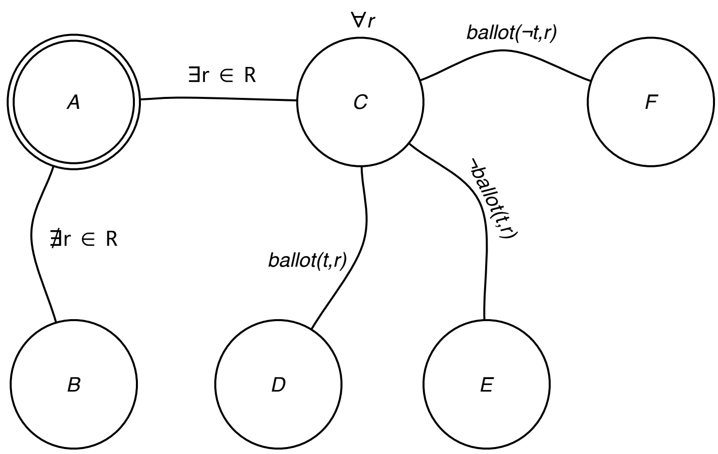


Figure 3.5: Reinstate Protocol

Semi-automated mediation is now performed, in which participants are presented with an automatically generated list of questions relating to the adoption of concept models which failed to be resolved through automated mediation. The first task is to reinstate any useful concepts removed during the pruning stage of the automated mediation. The protocol for the reinstating process is outlined in Figure 3.5. If no set of pruned concepts exist, then there is no need to reinstate any concepts in the ontology set, and therefore no ballot needs to take place (A→B). If the pruning produces a set of removed concepts ($R$), then a reinstate ballot takes place on each removed concepts

($r$) (A→C). All the participants are presented with the removed concept, along with its parent concept and asked the following question or a variant:

> "Is *concept_name*, which is a type of *superclass_name*, a valid concept?".

Participants should be warned to be careful with questions that ask if a concept is a type of *Thing*, and should only approve if they feel the candidate concept is abstract enough not to have any possible superclass other than *Thing*.

The questions are presented to all participants and the result recorded. If the majority adoption threshold ($t$) is met after the ballot has taken place, then the concept is reinstated and incorporated across all the ontology structures (C→D). If the majority adoption threshold is not met, but the minority threshold is met, then it is adopted by those models who supported it (C→F). Otherwise the concept should be discarded (C→F). Here the majority adoption and minority adoption thresholds should be set higher than those used for the automatic mediation; this is because people are more likely to agree with a suggested concept model than they are to come up with the same concept model independently. A high majority adoption threshold should be set here — although this could be changed according to the needs of the required model. To completely eradicate any chance of non-consensual entities being adopted, the majority adoption threshold would be set at 100%, but this would likely create more work for the knowledge engineer who could have to evaluate a greater number of minority adoption cases. If the ballot fails to gain a majority, then the concept is removed (C→E).

If a reinstate candidate concept is a superclass of other concepts, then the removal outcome will also remove the child concept (overriding any previous reinstate decision made on the child concept). For this reason the reinstate protocol needs to be applied in a bottom-up manner moving from lower level concepts to high level concepts. If a lower level concept is majority adopted, then the parent concept will also be majority adopted.

It is likely that in a crowd scenario, a large number of unique concept models may emerge resulting in the generation of a large number reinstate questions that are need to be referred the crowd. Indeed, it is anticipated that reinstate questions will constitute the largest proportion of the generated mediation questions. This means that additional methods need to be found to help reduce the number of reinstate questions by automatically assessing the validity of the pruned concepts. In Experiment 2, an example of how this might be achieved is given in (see Section 5.3.2). While this work does not incorporate a generalised method for reducing the number of reinstate cases to mediate over, some of the text-based methods outlined in section 2.5.1 might also be applicable. Any method used here will largely depend on the nature of the ontology domain and the reference resources available.

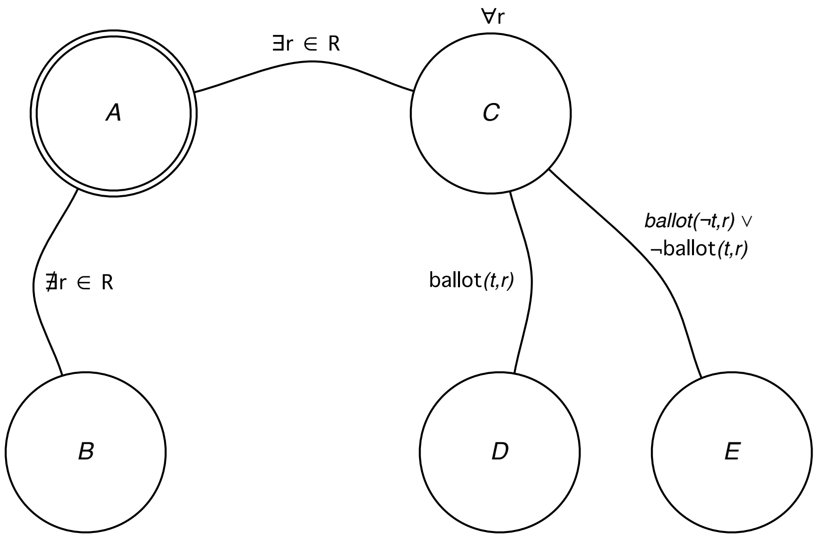


FIGURE 3.6: Minority Adoption Cases

#### Stage 4b: Minority Adoption and Conflict Cases

The next stage of semi-automated mediation is to identify those concepts which are above the pruning threshold, but which fail to meet the majority adoption threshold (i.e. those entities that met the criteria for *Minority Adoption* and/or where there are concept model conflicts and the total number of concepts matched meets the minority adoption threshold). In reference to figure 3.6, if a set (R) of minority mediation cases does not exist, then no ballot takes place (A→B). If there exists a set of minority mediation cases, then each concept (r) is presented to the participants in the same way as shown in stage 4a (using the same question phrasing) (A→C). If the ballot is successful and the majority adoption threshold is met, then it is adopted across all the ontologies (C→D). If the ballot fails, or the majority adoption threshold is not met, then the concept is retained but not adopted (C→E).

### Example of Semi-Automated Mediation

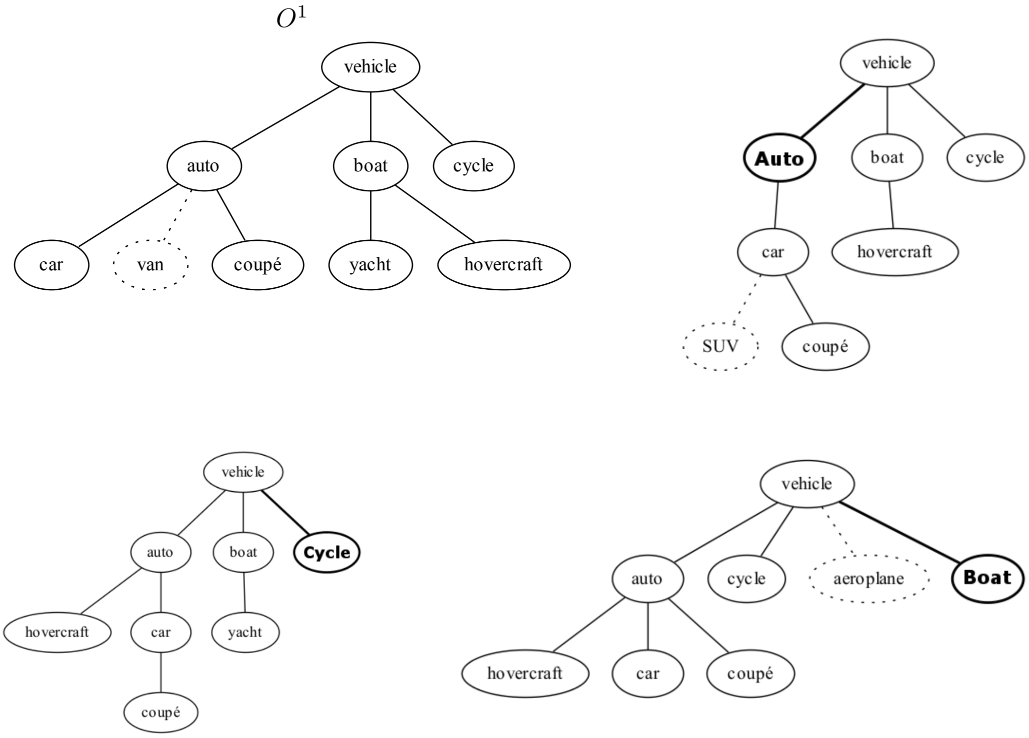


FIGURE 3.7: Mediation example: concept hierarchies pre-mediation

| **Mediated Concept** | $O^1$ | $O^2$ | $O^3$ | $O^4$ | **Outcome** |
|---:|---|---|---|---|---|
| aeroplane | | | ✓ | ✓ | *mA* |
| SUV | | ✓ | | | × |
| van | ✓ | ✓ | | ✓ | *MA* |
| yacht | ✓ | ✓ | ✓ | ✓ | *MA* |
| A {coupé $\sqsubseteq$ car} $\vee$ B {coupé $\sqsubseteq$ auto} | A | A | A | B | *MA (A)* |
| A {hovercraft $\sqsubseteq$ boat} $\vee$ B {hovercraft $\sqsubseteq$ auto} | A | A | B | B | *c* |

TABLE 3.3: Example manual mediation ballot

Using the example crowd input from earlier, Figure 3.7 and Table 3.3 show how semi-automated mediation would work. The statements in square brackets are in reference to Figure 3.5. The same adoption thresholds are used for the semi-automated mediation as the automated mediation (33% for minority adoption and 66% for majority adoption). Figure 3.7 shows the result of applying automated mediation on the ontology set. Note that new or moved concepts are made bold and those subject to pruning now have a dashed outline. Because we are at the semi-automated mediation stage, a manual ballot needs to take place. Table 3.3 illustrates an example outcome of a manual ballot. The first three rows of Table 3.3 show the results of the reinstate ballots, with the human input being fabricated for the purpose of demonstration. *van* is majority adopted [A→C→D], *aeroplane* is minority adopted [A→C→F] and *SUV* is removed [A→C→E].

The remaining mediation cases are dealt with using the protocol outlined in Figure 3.6 in Section 3.1.5. *Yacht*, which is already minority adopted, gains enough support to be majority adopted [A→C→D]. *Coupé*, which has a conflicting concept model is majority adopted (with the less-supported concept model being removed) [A→C→D]. Finally *hovercraft* remains in conflict and will therefore need expert adjudication [A→C→E].

### Example of Fully Mediated Ontologies

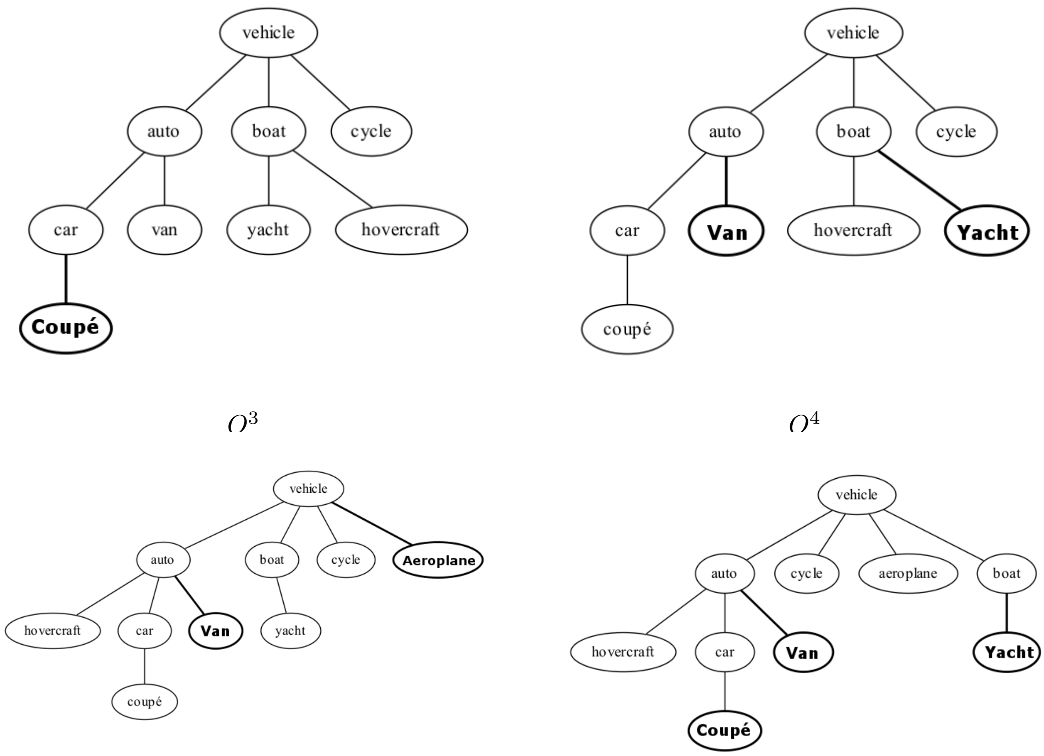


FIGURE 3.8: Mediation example: post semi-automated mediation

Figure 3.8 shows the final mediated ontologies. The concepts in bold have been added in accordance with the results of the manual ballot. Consensus has largely been achieved here. After mediation, the knowledge engineer would only be required to determine the concept models of *hovercraft* and *aeroplane*.

### 3.1.6 Enhanced Protocols for *WordNet* Incorporation

For Experiment 2 (See Section 5), additional processes were incorporated that require minor adjustments to the protocols described above. These adjustments were needed to reduce the number of reinstate questions generated (reducing the burden of human interaction) and to strengthen concept-matching. These additional processes are intended to allow this approach to be scaled without becoming too onerous and to enhance the semantic richness of the generated model. For an in-depth description of the processes outlined here, see Section 5.3.2. Below is a brief outline of how these additional processes are incorporated.

#### Stage 1: Formatting

Formatting is performed in the same way as described in Table 3.1.

#### Stage 2: Synonym Detection

*WordNet* is used to augment the process for matching equivalent concepts. It is recognised that label matching alone is not strong enough to determine all concept matches, this is mainly because equivalent terms are often used. *WordNet* is used in the following way to determine equivalency:

1. The synset of a single submitted concept is obtained from *WordNet*. *WordNet* synsets are a set containing the known equivalent terms for any given word.

2. The remaining submitted concepts are checked against this synset. If a match is made then the concepts are deemed to be equivalent.

3. Support for equivalent concepts is aggregated, with the combined support for both equivalent concepts being used to determine adoption. For example, if *car* was supported by one participant and *automobile* was supported by two participants; and both *car* and *automobile* were ruled to be equivalent – the two concepts would be considered as a single entity with the support of three participants when it came to mediation.

#### Stage 3: Automated Mediation I

Automated Mediation I (Adoption) is carried out in the same way as described in Table 3.1, using the aggregate support for each concept and their equivalencies.

#### Stage 4: Automated Mediation II

To deal with the greater number of pruning cases that are generated when the scale of participation is increased, *WordNet* is used to remove pruning cases where a specific criteria is met. The outline of this process is as follows:

1. Remove any concept that is not contained in the *WordNet* database

2. Remove any concept that is identified as being (or likely to be) an instance

An example of an entity being removed for not being contained within the *WordNet* database might be *counrty group* which is incorrectly spelt; an example of an instance being identified and removed might be *suomi* which is an instance of the concept *language* (both these examples are taken from Experiment 2 - see Chapter 5, Table 5.5). Any remaining concepts are used to produce reinstate questions in the same way as outlined in Section 3.1.5.

## 3.2 Method Evaluation

So far, this chapter has described, through a set of protocols, an experimental approach that could be used to acquire basic ontology structures from community crowds. The primary objective is to provide a re-usable and adaptable process through which crowd data can be semantically enhanced to provide a description of a given domain. Because this is an approach that aims to acquire knowledge from domains where no 'gold standard' exists, using such a standard for evaluation is not necessarily valid. Instead, the evaluation determines the level of consensus that is reached between the acquired knowledge structures. The argument is that if the community crowd is a good and reliable source of domain knowledge, then obtaining the commonly agreed upon knowledge from that crowd will provide an accurate representation of that domain. A secondary qualitative evaluation will also be used to determine the qualities of the produced consensual model.

Having acquired a set of crowdsourced ontologies and having processed those ontologies through a two-stage mediation process we need to measure the degree of convergence towards a consensus that we can achieve. To do this an evaluation method is needed to assess the similarity between set of ontologies.

The evaluation used for the experiments described is an adapted version of the cotopy-based approach offered by Dellschaft and Staab [22]. As this evaluation method was designed to compare a single learned ontology with a 'gold standard' we have adapted it so that it compares each ontology against all the others in a set, with the average of these measures being the determinant of the semantic similarity (and therefore the level of convergence). The final measure produced by this evaluation is called Cross-Compared Taxonomic F-Measure (CTF and $CTF^1$) for the concept structure. In this section there will be a description of the evaluation mechanism as defined by Dellschaft and Staab with a further description of how it is adapted.

### Determining Consensus between Two Ontologies

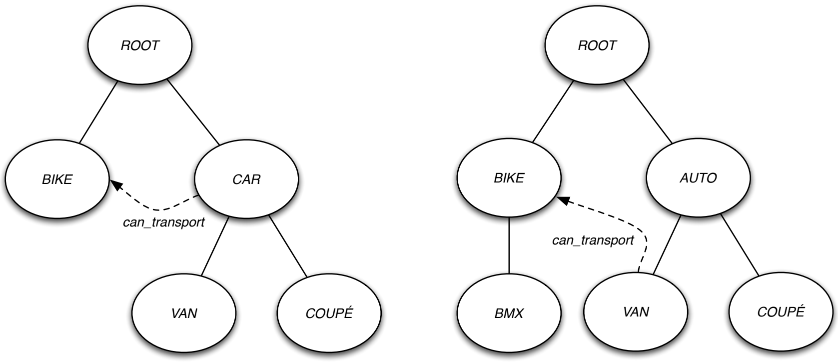


FIGURE 3.9: Example ontology structure (adapted from Dellschaft and Staab [22])

| $c$ | $sc(c, \mathcal{O}_{R1})$ | $sc(c, \mathcal{O}_{C1})$ |
|---|---|---|
| root | {root, bike, car, van, coupé} | {root, bike, BMX, auto, van, coupé} |
| car | {root, car, van, coupé} | – |
| auto | – | {root, auto, van, coupé} |
| van | {root, car, van} | {root, auto, van} |
| coupé | {root, car, coupé} | {root, auto, coupé} |
| bike | {root, bike} | {root, bike, BMX} |
| BMX | – | {root, bike, BMX} |

TABLE 3.4: Semantic cotopies derived from figure 3.9(from Dellschaft and Staab [22])

There are two ontology structures that are taken into account in the method described by Dellschaft and Staab: a 'gold standard' or reference ontology ($\mathcal{O}r$) and an acquired (or learned) ontology ($\mathcal{O}c$). Each concept entity within the pair of ontologies has a semantic cotopy which is defined as a sub-graph of the ontology concept structure that contains all the super classes and sub classes of the concept in question (see Figure 3.9 and Table 3.4 for an illustrative example). So where we have a given concept ($c$) and the ontology structure containing the concept ($\mathcal{O}$) and the set of concepts ($C_i$) that that ontology contains, the semantic cotopy is defined as follows:

$$sc(c, \mathcal{O}) := \{c_i \mid c_i \in \ C \wedge (c_i \leq c \vee c \leq c_i)\}$$

Having determined the semantic cotopy of identical concepts (concept-pairs) that exist in both the reference and learned ontology we can then determine the *local taxonomic precision* (tp) which is the intersection of concepts contained within the cotopies compared to the total concepts contained in the reference concept cotopy.

The local *taxonomic precision* is defined as follows:

$$tp_{sc}(c_1, c_2, \mathcal{O}_C, \mathcal{O}_R) := \frac{\mid sc(c_1, \mathcal{O}_c) \cap sc(c_2, \mathcal{O}_R) \mid}{\mid sc(c_1, \mathcal{O}_C) \mid}$$

The *global taxonomic precision* (TP) is the average of all the local *tp* measures that are performed. This means that the local taxonomic precision is calculated for each pair of matching concepts within the concept set ($C_c$). The resulting measure determines the average precision of concepts within the learned (acquired) ontology in relation to that of the reference ontology. TP roughly corresponds to the concept of *precision* in Information Retrieval and is a way of evaluating the effectiveness of retrieving information that is most relevant to the information request [101]. TP is defined as follows:

$$TP(\mathcal{O}_C, \mathcal{O}_R) := \frac{1}{C_C} \sum_{c \in C_c} \begin{cases} tp_{sc}(c, c, \mathcal{O}_C, \mathcal{O}_R) & \text{if c} \in C_R \\ 0 & \text{if c} \notin C_R \end{cases}$$

The *global taxonomic recall* (TR) can be calculated by reversing the arguments for TP. The *global taxonomic recall* determines the proportion of learned cotopies that match the reference ontology.

$$TR(\mathcal{O}_C, \mathcal{O}_R) := TP_{sc}(\mathcal{O}_R, \mathcal{O}_C)$$

It should be noted that both TR and TP are presented here to show the building blocks that make up this evaluation process. As they are being averaged out across a set of reference ontologies (each reference ontology becomes a learned ontology in the equation at some point) and therefore only the combined measure is indicative of the degree of consensus reached. The combined measure (TF) is defined as follows:

$$TF(\mathcal{O}_C, \mathcal{O}_R) := \frac{2 \cdot TP(\mathcal{O}_C, \mathcal{O}_R) \cdot TR(\mathcal{O}_C, \mathcal{O}_R)}{TP(\mathcal{O}_C, \mathcal{O}_R) + TR(\mathcal{O}_C, \mathcal{O}_R)}$$

The measures outlined above are not directly influenced by *lexical precision* (LP) and *lexical recall* (LR) which evaluate the similarity of lexical term layer without reference to the taxonomic structure. LP and LR reflect how well the learned lexical terms cover the reference domain's lexical layer. LP and LR are good indicators of the general coverage of a domain but do not take into account differences in models. A higher level of LP and LR would indicate that the ontology structures agree on the concepts that should be contained in the domain ontology, even if they disagree on where they are placed within the ontology structure. $\text{TF}^1$ is a further metric that factors in the lexical layer by using the lexical recall. It is defined as follows:

$$TF^1(\mathcal{O}_C, \mathcal{O}_R) := \frac{2 \cdot LR(\mathcal{O}_C, \mathcal{O}_R) \cdot TF(\mathcal{O}_C, \mathcal{O}_R)}{LR(\mathcal{O}_C, \mathcal{O}_R) + TF(\mathcal{O}_C, \mathcal{O}_R)}$$

Dellschaft and Staab suggest using *common semantic cotopy* (CSC) when the LP of the computed ontology is low, meaning that the reference ontology contains many terms that the computed ontology does not. While Dellschaft and Staab make no indication as to what level of LP should be considered either high or low, it is difficult to ascertain which measure to use. However, considering that the average lexical precision on the acquired ontology set from the implementation of this approach in Chapter 4.4 (Experiment 1) is 51.82% while the average lexical precision of the case studies used in Dellschaft and Staab is considerably lower at 9.35%, it felt reasonable to use SC rather than CSC.

### 3.2.1 CTF and $\text{CTF}^1$

As there is a need to compare more than two ontologies, this evaluation method needs to be adapted. This is done simply through an iterative process whereby all ontologies in the set ($\Theta$) are compared with each other. We call this measure *cross-referenced taxonomic f-measure* ($CTF$) and it is described as follows.:

$$CTF(\Theta_{\mathcal{O}}) = \frac{1}{\Theta_{\mathcal{O}}} \sum_{\mathcal{O} \in \Theta_{\mathcal{O}}} TF(\mathcal{O}, \Theta_n) \qquad \forall p \in \Theta_{\mathcal{O}} p, \in \Theta_n \text{ if } p \neq \mathcal{O}$$

$\text{CTF}^1$ uses the same iterative process to provide an average measure for the relative convergence of the ontology set using $\text{TF}^1$,

$$CTF^1(\Theta_{\mathcal{O}}) = \frac{1}{\Theta_{\mathcal{O}}} \sum_{\mathcal{O} \in \Theta_{\mathcal{O}}} TF^1(\mathcal{O}, \Theta_n) \qquad \forall p \in \Theta_{\mathcal{O}} p, \in \Theta_n \text{ if } p \neq \mathcal{O}$$

CTF and $\text{CTF}^1$ represents a small adaptation of the evaluation method offered by Dellschaft and Staab that enables multiple ontologies to be evaluated against the reference ontology. The way that CTF and $\text{CTF}^1$ work is perhaps best illustrated by the tables listing the detailed results found in Sections 4.4 and 6.7. Here, each table represents the CTF and $\text{CTF}^1$ calculations against one reference ontology within a set.

## 3.3 Qualitative Evaluation

The evaluation measures offered in the previous section will tell us the degree of convergence that is obtained across the ontology set. Of course, if we are removing and adopting entities the consensus will inevitably improve. This evaluation will provide a good indicator of the proportion of the domain that is 'common knowledge', and that about which differences of opinion exist. As there may be no 'gold standard' to compare against, a qualitative assessment of the consensual model can be performed. While this may prove subjective, a discussion on what has been incorporated and what has been excluded will give important insight into the effectiveness of this approach. In Chapter 7 a qualitative evaluation will take place which will look at the following questions:

- Are there any concepts and properties that have been erroneously endorsed?
- Are there any concepts or properties where the consensus model overrides a better model provided by a participant?

So, with a forensic look at the example implementations of this approach described in Chapters 4 and 6, the success of this approach can be established in terms of capturing crowd knowledge, removing erroneous input and identifying unresolvable conflicts.

## 3.4 Object Properties

Originally it was intended that the protocols above would include methods to mediate object properties along with concepts. While there was an attempt to achieve this in Experiments 1 and 3 (see Chapter 4 and 6), the amount and quality of the acquired object properties was severely lacking. This section briefly describes the approach that was taken and the evaluation method that was devised.

### 3.4.1 Automated Mediation

Table 3.5 indicates the rules that were applied to all the object properties obtained from participants.

| Stage | Rule |
|---|---|
| Formatting | **De-capitalise:** Property names that are not predefined have their first character converted to lower-case. **Composite Names:** Property names that contain multiple words are converted to lower-case strings with spaces being replaced with a underscore character (_). Where the user has used camel-case the initial capital is made lower-case and an underscore is inserted. **Edit Distance:** Where properties have names that are one edit-distance apart they are assumed to be equivalent. All subject properties are made uniform with the entity identified first being adopted. |
| Automated Mediation (I) | **Majority Adoption:** If an object property, including its domain and range specification is matched across the ontology set to the extent whereby it meets the majority adoption threshold, then it is universally adopted. **Minority Adoption:** If an object property, including its domain and range specification is matched across the ontology set to the extent whereby it meets the minority adoption threshold, then it is retained but not universally adopted. |
| Automated Mediation (II) | **Pruning:** Any object property that does not meet the minority adoption threshold is removed, it can be reinstated through manual mediation |
| Dependency | Additionally, if after manual mediation of the concept tree, a property exists that is dependent on a concept has been removed, the object property is also removed |

Table 3.5: Rules for automated mediation of Object Properties

A conflicting object property model is where a label match has been made, but the domain and range specification differ between knowledge models. Figure 3.10 illustrates a typical object property conflict.

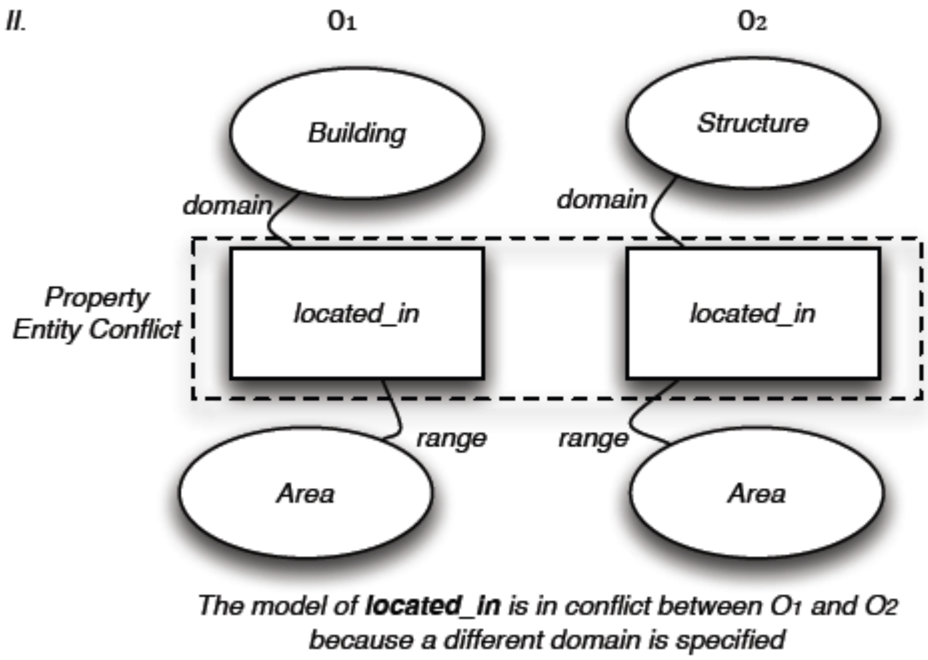


FIGURE 3.10: Example of a object property conflict

### 3.4.2 Automated Mediation

#### Stage 1: Automatic Mediation over Object Properties

For all object properties, majority adoption and minority adoption is determined based on whether the object property and its model (domain and range specification) are present in enough of the ontologies to meet the adoption thresholds. If there is a conflict between object property models and neither object property meets the majority adoption threshold, then the object property models in question will be manually mediated.

It should be noted that the above criteria may be too stringent as it not only requires a label match on the object property, but also that similar matches be made on the domain and range concepts.

#### Stage 2: Removal of Redundant Properties

The final stage of the process is to remove all those object properties that no longer have a domain or range concept defined (this is because these concepts will have been removed during concept mediation). Figure 3.11 describes the process. For all properties across the ontology set, if the domain and range concepts still exist in the ontology, then the property is retained (A→C); otherwise it is removed (A→B).

### 3.4.3 Semi-Automated Mediation

#### Stage 1: Questions

Questions are generated in the same way as for concept models for outcomes where there is a conflict, a minority adoption and where a reinstatement is possible. The problem here was that in the Experiments 1 and 3, no object property achieved majority adoption so all required manual mediation. This put too heavy a burden on the participants so it was decided that this stage of the object property mediation should not be conducted.

### 3.4.4 Evaluation using CDRF

To adapt the evaluation method described in Section 3.2 for determining convergence towards a consensus so that it can be used with object properties, a replacement for semantic cotopy needs to be found. Although object properties can exist in hierarchical super-property/sub-property structures, these are much less common than concept hierarchies in general. Therefore evaluating the convergence of object property descriptions needs to rely on alternative criteria. For this purpose a method of evaluation is devised in which we follow the evaluation steps described above, but replace the semantic cotopy with the domain and range and all the subclasses of the domain and range. The combined aggregate f-measure using the domain and range as a substitute for cotopy is referred to as CDRF.

Using the example ontology given in figure 3.9, an object property is specified (*can_transport*). In table 3.6, the *dr* measure is shown. This gives a flat representation of the domain and range that includes those classes that would inherit the object property. In this way we have a measure for object properties that is similar to the cotopy for concepts.

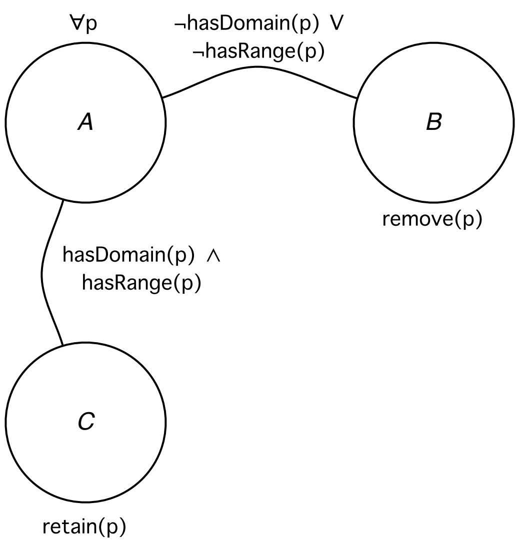


FIGURE 3.11: Removal of Redundant Properties

| $p$ | $dr(p,\mathcal{O}r)$ | $dr(p,\mathcal{O}l)$ |
|---|---|---|
| can_transport | {car, van, coupé, bike} | {auto, bike, BMX} |

TABLE 3.6: Domain and range as derived from figure 3.9

## 3.5 Summary

A process for knowledge acquisition is presented that is suitable for being incorporated into online community activity. A mechanism for eliciting concepts from community crowds and then using the same crowds to perform mediation is outlined. An evaluation which can determine the convergence towards a consensus is described, along with the criteria for a qualitative evaluation. Finally, an approach to acquiring and mediating object properties, which proved unsuccessful in its current form, is outlined.

# Chapter 4

# Experiment 1: Building a Consensual Model from Task-Aware Crowds

## 4.1 Description

Many studies of crowdsourcing tend to look at pre-existing crowdsourced resources that are built from the input of many thousands, or even millions, of users. Indeed, the notion of the 'crowd' tends to engender an image of huge numbers of people. The definition of the 'crowd' that is relevant here, particularly if looking at online community crowds that cover niche domains, is not of the quantity of users but the qualities they possess and the challenges that these groups present. The qualities are that they are motivated and have some expertise; the challenges are that the expertise levels are different and may be communicated in an inconsistent way.

The experiments described in this chapter use pre-existing collaborative ontology engineering tools to implement the approach outlined in Chapter 3. Here we will test the workflow and the response of the participants and provide a baseline for comparison with the experiment described in Chapter 6 that incorporates this approach into a community crowd, game motivated scenario. The first experiment described, establishes the viability of the approach by using manual quality control over the acquired models (by rejecting incomplete and poor quality submissions) and by incorporating an extension phase to ensure reasonable coverage. The second experiment is conducted in the same subject domain as the first, but has significantly more participants and does not rely on manual quality control or artificially extending the acquired ontologies.

## 4.2 Objective

The objectives of these experiments are:

- To test the effectiveness of the protocols described in Chapter 3 in terms of gaining a consensus
- To provide a comparison for later experimentation which performs a similar task but without the participants being aware of the model building process
- To determine the scope of conflicting knowledge that will require a knowledge engineer to resolve
- To see if the mediated models converge with a 'gold standard'
- To see if the approach is scalable and provides increased coverage given a larger number of models to mediate over

In these experiments, the participants are representative of an online community crowd whose membership are *task-aware*. That is to say, they are aware that the objective of participation is to create a knowledge representation and they have some basic understanding of this process. These experiments will show positive results if a significant degree of convergence towards a consensus, as measured by CTF and $CTF^1$ (see Section 3.2), is achieved.

The expected outcome is a convergence towards a conceptual consensus across the ontology set. Any knowledge remaining in the ontology set that is not universally supported could have a degree of validity, meaning that it should be considered for incorporation by a knowledge engineer when building a final model. This means that models which are not fully adopted should also reveal the marginal models that a knowledge engineer would be able to adjudicate over.

Ideally the convergence towards a consensus should be measured in conjunction with measuring the convergence towards a gold standard. Such a dual evaluation would allow this approach to be evaluated both in terms of creating agreed upon knowledge models through consensus and on the consistency and utility of the model through comparison with the gold standard. While the inclusion of a gold standard evaluation would be ideal, in most instances where this approach would be employed, the domain in question would not have been mapped by a knowledge engineer and therefore achieving consensus is the primary goal. For the purposes of demonstrating how a gold standard evaluation might be performed and what such an evaluation might look like, convergence towards a gold standard has been measured in Experiment 1. This is possible because a gold standard ontology exists for this domain. A high degree of similarity with a gold standard is not anticipated at this point due to the size of the crowd, however, some degree of convergence towards the gold standard would be indicative that useful knowledge is being modelled.

### Devising Competency Questions

Both experiments require that participants are able to devise, understand and translate 'competency questions' in order to build appropriate structures. Competency questions provide queries which the ontology must be able to answer [41]. Competency questions are usually expressed in informal language at the beginning of a development cycle, but may be made formal at a later stage. They represent queries that would be typically asked of an ontology and can be used to measure the scope, efficiency and expressiveness of an ontology. Well-constructed competency questions require more than a standard look-up query; they will test the structure of the ontology by requiring appropriate constraints to be present in order to provide the correct answer. Competency questions are usually used for ontology validation to help determine how well an ontology functions given a 'real world' scenario [36]. Competency questions are also used to determine the expressiveness of the target ontology by providing a set of problems for which the ontology would be able to express a solution [41]. While there is no test to ensure the completeness of an ontology [36] by providing as broad a set of appropriate competency questions, we can aim to identify incomplete definitions and modify the ontology as required. With this in mind the participants are required to evaluate a knowledge domain and devise competency questions that can be asked of it. For Experiment 1, they will also be required to use competency questions devised by other members of the crowd and to modify their own knowledge representation to answer the aggregate set of competency questions.

## 4.3 Method

1. Participants provided with a domain description
2. Participants devise a set of competency questions
3. Model produced to solve competency questions
4. Competency questions distributed between participants
5. Models extended to accommodate new competency questions
6. Automated mediation applied
7. Semi-Automated mediation applied

### Participation

The knowledge domain used for this experiment was a *Music Store*. Therefore, the ontology would have to be able to describe the processes, categorization and objects present in a typical music store with a view to supporting an application in this area. The acquired ontologies may include classifications for concepts such as *genres* and *artists* along with properties to determine relationships such as song authorship and group membership. This choice was made because our participants are likely to have some familiarity with the domain, yet we cannot assume that this knowledge is uniform or complete. Therefore we have a crowd where knowledge is very likely to exist but which will need to be processed to obtain consensus and to remove error.

As we have pre-conditions on the composition of the crowd, the question is whether this really is a 'crowd' in the sense that we have some idea of the expertise. However, the domain knowledge is unknown as there is no assessment of the participants' knowledge of music stores or music in general. Therefore this experiment should test whether we can obtain a useful and consensual model of a domain from a crowd whose domain expertise is unknown without having to take misunderstanding of knowledge engineering principles into account.

There is also a pre-existing ontology (or 'gold standard' ontology) for this domain which has been appropriately designed by a knowledge engineer with which we can make a comparison. This will provide a useful comparison particularly in determining if the consensual crowdsourced ontology has significant gaps in its coverage. We should see some convergence towards the 'gold standard' using this approach. However, as one of the advantages of using crowdsourcing in this way is to elicit alternative models of the subject domain, any acquired ontology may be significantly different from the gold standard. Therefore a high degree of similarity to the 'gold standard should not necessarily be seen as an indicator of the usefulness or completeness of an acquired ontology.

Given the domain and the type of crowd required by the experiment, it was appropriate that we presented this experiment to undergraduate and masters students enrolled on a semantic web module titled Advanced Web Technologies, run by the Computer Science department of the University of Liverpool. While most of the students were from Computer Science there were also students with background in other disciplines including maths and psychology. The students were presented this as an exercise in learning about knowledge acquisition and it took place three weeks into the module. At this point in the course the students would be expected to understand some of the fundamentals of ontology development including the idea of concept hierarchies, object properties and domain and range restrictions.

We divided the class into five groups, each consisting of between 3 and 5 students. In total there were 8 groups and 29 students. Of the 8 groups, 3 failed to meet the minimum criteria by not completing the tasks given to them in the set time (groups 1,2 and 7).

### Restrictions

The participants, having been familiarised with the idea of competency questions, were asked as a group to formulate five competency questions that a music store ontology should satisfy. Participants were also provided with a set of pre-defined base concepts; *Record*, *Song*, *Person*, *Genre*, *Gender*, *Group*. The purpose of providing these concepts was to define a common lexicon for users and to ensure that some entity matches would be made even if the acquired ontology structures deviated considerably. This pre-defined concept set was excluded from the evaluation calculation so that only the acquired concepts were measured.

### 4.3.1 Obtaining Competency Questions and Initial Ontology structures

To reiterate, the objective of the approach used was to acquire a set of basic ontology structures from the participants which, through a process of automated and semi-automated mediation, converge into a single consensual ontology that would provide a representation of the target domain. This acquired ontology would also contain concepts for which there was some, but not unanimous support or which have concept models that conflict; these concepts would be noted so that a knowledge engineer could make the final adjudication. At this stage the participants are only required to define a concept hierarchy and the object properties needed to define the relationships between these concepts.

Using the competency questions they were then asked to determine the concepts and object properties that would be needed to model the domain, and in particular to answer their particular set of competency questions. Then each group was tasked with engineering the ontology that would be needed to answer their set of competency questions, using *MoKi*[1] [25]. The complete instructions given to the participants are provided in Appendix B.

*MoKi* was chosen because it is available through a web-browser, meaning that it was readily available, and because for the tasks required it felt more intuitive than alternatives such as *WebProtege*.

After collecting the competency questions and provisional ontology structures a representative from each group was chosen to extend and refine their ontology. To achieve this they were provided with additional competency questions obtained from the filtered aggregate of other groups. The idea behind this was that, given a broader set of competency questions, the participants would have to produce more complete and therefore more expressive ontologies. While this extension might appear contrived as the filtering was based on a deliberate attempt to restrict the scope of the model, this helped to simulate a broader and more in-depth participation that would be achieved from having a greater participation. While this was a satisfactory solution for Experiment 1, it is acknowledged that this has the potential to introduce bias, therefore this expansion step was not repeated in Experiment 2 where a greater level of participation is achieved.

The informal competency questions obtained from the participants are listed in Table 4.1. It is fairly evident that the quality of these questions varied significantly, which is to be expected as the participants are, in terms of their knowledge of the domain in question, representative of crowd input. The task that presents itself was how to remove

[1] `https://moki.fbk.eu/website/index.php`

those questions which do not assist in the defining of an adequate ontology structure. The final competency question-set was manually chosen by filtering questions according to the following criteria:

- Does the question intuitively require a constraint or particular concept structure to be specified in order to answer?
- Would the competency question be better answered in an alternative domain?
- Is there considerable ambiguity in the phrasing of the question such that the competency which is being tested is unclear?

Additionally, some competency questions could have been better expressed as two or more separate queries, typically where the participant had used the 'or' clause. In these cases the competency question is replaced by two or more queries that represented all the possible outcomes of the original. Additionally, where instances were referenced directly the competency question is altered to provide a concept as variable (denoted by square brackets). Given these criteria, the refined competency questions were then re-circulated to the groups in order to refine their models, as listed in Table 4.2.

| **Question** | **Group(s)** |
|---|---|
| What is the most popular song this week? | 1,2 |
| What was the album published in [year] by [artist]? | 1 |
| Which female artist won an award this year? | 1 |
| Which classic English album cost is less that 10? | 1 |
| What is the most popular single? | 2 |
| What is the most popular group? | 2 |
| Does the store have a search function? | 2 |
| Does the store have a download function? | 2 |
| Which male artist sang *one love*? | 3 |
| What albums is artist best known for? | 3 |
| List the new rock songs? | 3 |
| How many members are in the *wailers*? | 3 |
| What was the No1 album of 2012? | 3 |
| This group/person belongs to which genre? | 4 |
| This song belongs to which record? | 4 |
| This person belongs to which group? | 4 |
| This record belongs to which artist? | 4 |
| Which album has [artist] released? | 5 |
| Which genre is [artist] best known for? | 5 |
| Which artist won an award last year? | 5 |
| What song charted last week? | 5 |
| What song is free this week? | 5 |

Table 4.1: Unprocessed competency questions obtained from participants.

The revised list of competency questions was then re-circulated amongst the participants and they were asked to augment their models according to the extended list. At this point five ontology structures had been obtained from our participants, each designed to answer an aggregate set of competency questions, which itself, had been obtained from the crowd of participants. The number of entities (concepts and object properties) that had been obtained at this point is summarised in Table 4.3.

| **Group No.** | **Concept Count** | **Property Count** |
|---|---|---|
| Group 3 | 18 | 18 |
| Group 4 | 8 | 6 |
| Group 5 | 9 | 7 |
| Group 6 | 18 | 10 |
| Group 8 | 10 | 9 |

Table 4.3: Summary of concepts and properties obtained from combined competency question list.

| **Question** | **Action** |
|---|---|
| *What is the most popular song this week?* | Ambiguity |
| What was the album published in [year] by [artist]? | - |
| Which female artist won an award this year? | - |
| *Which classic English album cost is less that 10?* | Domain |
| *What is the most popular single?* | Ambiguity |
| *What is the most popular group?* | Ambiguity |
| Does the store have a search function? | - |
| Does the store have a download function? | - |
| Which male artist sang [song]? | - |
| What albums is artist best known for? | - |
| *List the new rock songs?* | Ambiguity |
| How many members are in [group]? | - |
| What was the No1 album of [year]? | - |
| This group belongs to which genre? | Split |
| This person belongs to which genre? | Split |
| This song belongs to which record? | - |
| This person belongs to which group? | - |
| This record belongs to which artist? | - |
| *Which album has [artist] released?* | Ambiguity |
| Which genre is [artist] best known for? | - |
| Which artist won an award last year? | - |
| What song charted last week? | - |
| What song is free this week? | - |

Table 4.2: Competency questions filtered according to stated criteria. Entries *emphasised* were removed from consideration from second stage refinement

### 4.3.2 Mediation Results

#### Formatting

Table 4.4 shows the number of edits that were made across the ontology set. The most common modifications were made using the de-capitalisation rule. No edit distance changes were required. The formatting rules were applied using a small program written in Java which automatically updated the ontology files.

Table 4.4: Formatting Edits

| Group | De-Capitalisation | Composite | Edit Distance |
|---|---|---|---|
| Group 3 | 30 | 12 | 0 |
| Group 4 | 12 | 0 | 0 |
| Group 5 | 15 | 0 | 0 |
| Group 6 | 25 | 5 | 0 |
| Group 8 | 15 | 0 | 0 |

### 4.3.3 Automated Mediation

**Thresholds**

| | |
|---|---|
| Minority Adoption Threshold (Automated Mediation) | 20% |
| Majority Adoption Threshold (Automated Mediation) | 50% |
| Minority Adoption Threshold (Reinstate) | 50% |
| Majority Adoption Threshold (Reinstate / mA $\rightarrow$ MA) | 80% |

Table 4.5: Adoption Thresholds for Experiment 1

Table 4.5 gives an overview of the thresholds used in Experiment 1. The majority adoption threshold for reinstatement also applies to the majority adoption questions that determine if a minority adoption case should be upgraded. The thresholds have been set according to the results of the initial data acquisition from the crowd. To ensure that the users are not overwhelmed with mediation questions the adoption thresholds can be changed to suit the natures of crowd and the data it is contributing. Where matches are common the threshold can be higher, where matches are less frequent the thresholds should be lowered.

**Automated Mediation Overview**

Table 4.6 gives an overview of the automatic mediation process as it affects the concept hierarchy. All the concepts are listed in the first column, and their presence in each ontology listed in the next five columns. The *Conflict* column indicates if there was a conflict which could be resolved through auto-balloting to try and gain a resolution (see Section 3.1.5). If a conflict exists, it is marked with a '✓, otherwise a ' ×' symbol is used to indicate that no conflict exists. For each conflict the symbol in brackets indicates whether or not that conflict was resolved using automatic mediation.

The next column indicates the confidence, which is a measure of the proportion of ontologies that contain the given ontology at this stage. If there is a conflict in the model then the confidence measure is based on the most common model. The next three columns indicate what action was taken during automatic mediation. Column *mA* indicates that a minority adoption occurred. Column *MA* indicates that a majority adoption occurred and the concept was added to all ontologies. The final column indicates that a prune action was performed.

Table 4.7 indicates which object properties were majority adopted (see section 3.1.5). As mentioned in Section 3.4, the mediation of object properties proved difficult, therefore this is included for information only and for the purposes of disclosing all aspects of the experiment.

| **Concept** | $\mathcal{O}^3$ | $\mathcal{O}^4$ | $\mathcal{O}^5$ | $\mathcal{O}^6$ | $\mathcal{O}^8$ | **No conflict?** | **Conf.** | **mA** | **MA** | **P** |
|---:|---|---|---|---|---|---|---|---|---|---|
| album | | | | • | | | | | | • |
| artist | • | • | | • | | × | 0.6 | | • | |
| award | • | | • | • | • | × | 0.8 | | • | |
| best_album | | | | • | | | | | | • |
| best_group | | | | • | | | | | | • |
| best_record | | | | • | | | | | | • |
| best_single | | | | • | | | | | | • |
| chart | | • | | | | | | | | • |
| click_rate | • | | | | | | | | | • |
| composer | • | | | • | | ✓(×) | 0.2 | | | |
| continent | • | | | | | | | | | • |
| country | • | | | | • | × | 0.4 | • | | |
| covers | | | | | • | | | | | • |
| download_times | • | | | | | | | | | • |
| e_p | | | | • | | | | | | • |
| false | • | | | | | | | | | • |
| female | • | | | | | | | | | • |
| first | | | | | • | | | | | • |
| Gender* | • | • | • | • | • | ✓(✓) | 0.8 | | • | |
| Genre* | • | • | • | • | • | × | 1 | | • | |
| Group* | • | • | • | • | • | ✓(✓) | 0.8 | | • | |
| label | | | | • | | | | | | • |
| male | • | | | | | | | | | • |
| member | • | | | | | | | | | • |
| month | | • | • | • | • | ✓(✓) | 0.4 | • | | |
| name | • | | | | | | | | | • |
| no._1 | | | | | • | | | | | • |
| no1 | | • | | | | | | | | • |
| offer | | | | • | | | | | | • |
| Person* | • | • | • | • | • | × | | | • | |
| price | • | | | • | | × | 0.4 | • | | |
| producer | | | • | • | | × | 0.4 | • | | |
| record_label | | | • | | | | | | | • |
| Record* | • | • | • | • | • | × | 1 | | • | |
| recording_artist | | | • | | | | | | | • |
| release_date | • | | | | | | | | | • |
| sales | | | | | • | | | | | • |
| singer | • | | | | | | | | | • |
| Song* | • | • | • | • | • | × | 1 | | • | |
| store | • | • | • | • | • | × | 1 | | • | |
| time | | | | • | | | | | | • |
| time_period | | • | | | | | | | | • |
| true | • | | | | | | | | | • |
| week | | • | • | • | • | ✓(✓) | 0.4 | • | | |
| writer | | | • | | | | | | | • |
| year | • | • | • | • | • | ✓(✓) | 0.4 | • | | |

TABLE 4.6: Overview of Automated Mediation for Concepts

| Properties | $\mathcal{O}^3$ | $\mathcal{O}^4$ | $\mathcal{O}^5$ | $\mathcal{O}^6$ | $\mathcal{O}^8$ | Con. | Conf. | MA |
|---:|---|---|---|---|---|---|---|---|
| achieved | | | | | • | | | |
| belongs_to | | • | • | | | × | 0.4 | |
| belongs_to_country_of | • | | | | | | | |
| best_known_for | • | | | • | | × | 0.4 | |
| compose | | | | • | | | | |
| for_time | | | | • | | | | |
| has_chart | | • | | | | | | |
| has_click_rate | • | | | | | | | |
| has_download_function | • | | | | | | | |
| has_download_times | • | | | | | | | |
| has_gender | • | • | • | | • | ✓ | 0.6 | • |
| has_genre | • | • | | | | × | 0.4 | |
| has_member | • | | | | | | | |
| has_name | • | | | | | | | |
| has_price | • | | | • | | × | 0.4 | |
| has_record | • | | | | | | | |
| has_release_date | • | | | | | | | |
| has_released | | | | • | | | | |
| has_search_function | • | | | | | | | |
| has_song | • | | | | | | | |
| has_time_period | | • | | | | | | |
| has_year | | • | | | | | | |
| included | | | | • | | | | |
| is_on | | | • | | | | | |
| label_of | | | | • | | | | |
| number_one | • | | | | | | | |
| owns | | | • | | | | | |
| produce | | • | | • | | ✓ | 0.4 | |
| produce | | | • | | | | | |
| provide | | | | • | | | | |
| sings | | | • | | | | | |
| sings_song | • | | | | • | ✓ | 0.4 | |
| sung_by | • | | | | | | | |
| won_award | • | | | • | • | ×(✓) | 0.6 | • |
| writes | • | | • | | | ✓ | 0.4 | |

Table 4.7: Overview of Automated Mediation for Properties

### Semi-Automated Mediation

Table 4.8 shows the questions generated by the reinstate protocol (see Section 3.1.5). The first column contains the generated question, the second column indicates the confidence (i.e. the proportion of the participants who answered the question in support of the proposition). The next three columns indicate the action taken. *mA* indicates that a minority adoption took place, meaning that the concept was adopted by the set proportion of the ontologies that indicated support. *MA* indicates that a majority adoption took place, meaning that the concept was adopted across the ontology set. *C* indicates that the removed concept was culled as it was not supported by a majority of

| Reinstate Questions | Con. | mA | MA | C |
|---:|---|---|---|---|
| Is album, which is a type of thing, a valid concept? | 0.8 | | • | |
| Is best_album, which is a type of award, a valid concept? | 0.8 | | • | |
| Is best_group, which is a type of award, a valid concept? | 0.8 | | • | |
| Is best_record, which is a type of award, a valid concept? | 0.8 | | • | |
| Is best_single, which is a type of award, a valid concept? | 0.6 | • | | |
| Is chart, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is click_rate, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is continent, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is covers, which is a type of song, a valid concept? | 1 | | • | |
| Is download_times, which is a type of thing, a valid concept? | 0.4 | | | • |
| Is e_p, which is a type of Record, a valid concept? | 0.8 | | • | |
| Is ×, which is a type of thing, a valid concept? | 0.2 | | | |
| Is female, which is a type of gender, a valid concept? | 0.6 | • | | |
| Is first, which is a type of no._1, a valid concept? | 0.2 | | | • |
| Is label, which is a type of Person, a valid concept? | 0.2 | | | • |
| Is male, which is a type of gender, a valid concept? | 0.6 | • | | |
| Is member, which is a type of artist, a valid concept? | 0.4 | | | • |
| Is name, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is no._1, which is a type of Song, a valid concept? | 0.4 | | | • |
| Is no1, which is a type of Song, a valid concept? | 0.2 | | | • |
| Is offer, which is a type of store, a valid concept? | 0.2 | | | • |
| Is record_label, which is a type of thing, a valid concept? | 0.4 | | | • |
| Is recording_artist, which is a type of Person, a valid concept? | 0.8 | | • | |
| Is release_date, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is sales, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is singer, which is a type of Person, a valid concept? | 1 | | • | |
| Is time, which is a type of thing, a valid concept? | 0.4 | | | • |
| Is time_period, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is true, which is a type of thing, a valid concept? | 0.2 | | | • |
| Is writer, which is a type of Person, a valid concept? | 1 | | • | |

TABLE 4.8: Semi-Automated Mediation: Reinstate Questions

the participants. The thresholds used here were 50% for minority adoption and 80% for majority adoption - higher than in Experiment 1 as more matches have been made.

| Majority Adopt Question | Con. | mA | MA |
|---:|---|---|---|
| Is country, which is a type of thing, a valid concept? | 0.6 | • | |
| Is month, which is a type of thing, a valid concept? | 0.8 | | • |
| Is price, which is a type of thing, a valid concept? | 0.6 | • | |
| Is producer, which is a type of person, a valid concept? | 1 | | • |
| Is week, which is a type of thing, a valid concept? | 0.8 | | • |
| Is year, which is a type of thing, a valid concept? | 0.8 | | • |

TABLE 4.9: Automated Mediation: Minority to Majority Adopt Questions

Table 4.9 shows the questions generated from the minority adoption cases. The ballot confidence threshold of 80% was needed for majority adoption.

## 4.4 Results

This section contains the full tables of results obtained for the convergence analysis, followed by an overview in which any trends are illustrated. The measure used to calculate the CTF is the TF and the measure used to calculate the CTF$^1$ is TF$^1$. The lexical recall (LR) and lexical precision (LP) are still recorded as they might give some insight into the effect of the mediation process — in particular, these measures might indicate how much information is lost during the pruning/reinstate process.

**Unprocessed**

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF$^1$ |
|---|---|---|---|---|---|---|
| 4 | 66.67% | 33.33% | 91.67% | 76.39% | 83.33% | 47.62% |
| 5 | 60.00% | 33.33% | 87.50% | 76.39% | 81.57% | 47.33% |
| 6 | 36.84% | 38.89% | 84.29% | 83.33% | 83.81% | 53.13% |
| 8 | 58.33% | 38.89% | 57.62% | 79.76% | 66.91% | 49.19% |
| | CTF of $\mathcal{O}^3$ | 78.91% | CTF$^1$ of $\mathcal{O}^3$ | 49.32% | | |

TABLE 4.10: Unprocessed: $\mathcal{O}^3$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF$^1$ |
|---|---|---|---|---|---|---|
| 3 | 33.33% | 66.67% | 76.39% | 91.67% | 83.33% | 74.07% |
| 5 | 60.00% | 66.67% | 87.50% | 91.67% | 89.53% | 76.43% |
| 6 | 36.84% | 77.78% | 84.29% | 100.00% | 91.47% | 84.07% |
| 8 | 50.00% | 66.67% | 50.56% | 91.67% | 65.17% | 65.91% |
| | CTF of $\mathcal{O}^4$ | 82.38% | CTF$^1$ of $\mathcal{O}^4$ | 75.12% | | |

TABLE 4.11: Unprocessed: $\mathcal{O}^4$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF$^1$ |
|---|---|---|---|---|---|---|
| 3 | 33.33% | 60.00% | 76.39% | 87.50% | 81.57% | 69.14% |
| 4 | 66.67% | 60.00% | 91.67% | 87.50% | 89.53% | 71.85% |
| 6 | 36.84% | 70.00% | 84.29% | 92.86% | 88.36% | 78.12% |
| 8 | 50.00% | 60.00% | 50.56% | 87.50% | 64.08% | 61.98% |
| | CTF of $\mathcal{O}^5$ | 80.89% | CTF$^1$ of $\mathcal{O}^5$ | 70.27% | | |

TABLE 4.12: Unprocessed: $\mathcal{O}^5$

**After Extension and Formatting**

The extension and formatting should have a positive impact on the lexical precision and recall of the ontologies, as aggregating the competency questions should prompt the participants to create similar entities. This is true in general: the lexical precision improves on average by 9%, indicating that more entity matches are made. However, this improvement is not uniform across the ontology set: $\mathcal{O}^5$ sees a reduction in lexical

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 38.89% | 36.84% | 83.33% | 84.29% | 83.81% | 51.18% |
| 4 | 77.78% | 36.84% | 100.00% | 84.29% | 91.47% | 52.53% |
| 5 | 70.00% | 36.84% | 92.86% | 84.29% | 88.36% | 52.00% |
| 8 | 58.33% | 36.84% | 57.62% | 70.00% | 63.21% | 46.55% |
| | CTF of $\mathcal{O}^6$ | 81.71% | $CTF^1$ of $\mathcal{O}^6$ | 50.57% | | |

TABLE 4.13: Unprocessed: $\mathcal{O}^6$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 38.89% | 58.33% | 79.76% | 57.62% | 66.91% | 62.33% |
| 4 | 66.67% | 50.00% | 91.67% | 50.56% | 65.17% | 56.59% |
| 5 | 60.00% | 50.00% | 87.50% | 50.56% | 64.08% | 56.17% |
| 6 | 36.84% | 58.33% | 70.00% | 57.62% | 63.21% | 60.67% |
| | CTF of $\mathcal{O}^8$ | 64.84% | $CTF^1$ of $\mathcal{O}^8$ | 58.94% | | |

TABLE 4.14: Unprocessed: $\mathcal{O}^8$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 4 | 64.29% | 37.50% | 88.89% | 79.26% | 83.80% | 51.81% |
| 5 | 60.00% | 37.50% | 91.67% | 83.70% | 87.50% | 52.50% |
| 6 | 50.00% | 50.00% | 77.50% | 83.33% | 80.31% | 61.63% |
| 8 | 62.50% | 41.67% | 70.33% | 85.33% | 77.11% | 54.10% |
| | CTF of $\mathcal{O}^3$ | 82.18% | $CTF^1$ of $\mathcal{O}^3$ | 55.01% | | |

TABLE 4.15: Extended and Formatted: $\mathcal{O}^3$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 37.50% | 64.29% | 79.26% | 88.89% | 83.80% | 72.76% |
| 5 | 66.67% | 71.43% | 92.50% | 75.00% | 82.84% | 76.71% |
| 6 | 45.83% | 78.57% | 71.82% | 81.82% | 76.49% | 77.52% |
| 8 | 62.50% | 71.43% | 70.33% | 75.00% | 72.59% | 72.01% |
| | CTF of $\mathcal{O}^4$ | 78.93% | $CTF^1$ of $\mathcal{O}^4$ | 74.75% | | |

TABLE 4.16: Extended and Formatted: $\mathcal{O}^4$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 37.50% | 60.00% | 83.70% | 91.67% | 87.50% | 71.19% |
| 4 | 71.43% | 66.67% | 75.00% | 92.50% | 82.84% | 73.88% |
| 6 | 50.00% | 80.00% | 67.50% | 95.83% | 79.21% | 79.60% |
| 8 | 68.75% | 73.33% | 73.03% | 93.18% | 81.88% | 77.37% |
| | CTF of $\mathcal{O}^5$ | 82.86% | $CTF^1$ of $\mathcal{O}^5$ | 75.51% | | |

TABLE 4.17: Extended and Formatted: $\mathcal{O}^5$

precision indicating that there are fewer entity matches as a proportion of the concepts compared across the ontologies. This probably indicates that this group performed poorly in the extension phase.

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF$^1$ |
|---|---|---|---|---|---|---|
| 3 | 50.00% | 50.00% | 83.33% | 77.50% | 80.31% | 61.63% |
| 4 | 78.57% | 45.83% | 81.82% | 71.82% | 76.49% | 57.32% |
| 5 | 80.00% | 50.00% | 95.83% | 67.50% | 79.21% | 61.30% |
| 8 | 68.75% | 45.83% | 73.03% | 62.73% | 67.49% | 54.59% |
| | CTF of $\mathcal{O}^6$ | 75.88% | CTF$^1$ of $\mathcal{O}^6$ | 59.71% | | |

TABLE 4.18: Extended and Formatted: $\mathcal{O}^6$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF$^1$ |
|---|---|---|---|---|---|---|
| 3 | 41.67% | 62.50% | 85.33% | 70.33% | 77.11% | 69.04% |
| 4 | 71.43% | 62.50% | 75.00% | 70.33% | 72.59% | 67.17% |
| 5 | 73.33% | 68.75% | 93.18% | 73.03% | 81.88% | 74.74% |
| 6 | 45.83% | 68.75% | 62.73% | 73.03% | 67.49% | 68.11% |
| | CTF of $\mathcal{O}^8$ | 74.77% | CTF$^1$ of $\mathcal{O}^8$ | 69.77% | | |

TABLE 4.19: Extended and Formatted: $\mathcal{O}^8$

### Automated Mediation

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 4 | 66.67% | 41.67% | 90.00% | 81.33% | 85.45% | 56.02% |
| 5 | 62.50% | 41.67% | 94.00% | 81.33% | 87.21% | 56.39% |
| 6 | 50.00% | 50.00% | 77.50% | 83.33% | 80.31% | 61.63% |
| 8 | 64.71% | 45.83% | 88.18% | 83.03% | 85.53% | 59.68% |
| | CTF of $\mathcal{O}^3$ | 84.63% | $CTF^1$ of $\mathcal{O}^3$ | 58.43% | | |

Table 4.20: Automated Mediation: $\mathcal{O}^3$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 41.67% | 66.67% | 81.33% | 90.00% | 85.45% | 74.90% |
| 5 | 75.00% | 80.00% | 95.00% | 83.33% | 88.79% | 84.16% |
| 6 | 50.00% | 80.00% | 67.50% | 83.33% | 74.59% | 77.20% |
| 8 | 70.59% | 80.00% | 89.17% | 83.33% | 86.15% | 82.96% |
| | CTF of $\mathcal{O}^4$ | 81.25% | $CTF^1$ of $\mathcal{O}^4$ | 79.81% | | |

Table 4.21: Automated Mediation: $\mathcal{O}^4$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 41.67% | 62.50% | 81.33% | 94.00% | 87.21% | 72.82% |
| 4 | 80.00% | 75.00% | 83.33% | 95.00% | 88.79% | 81.31% |
| 6 | 54.17% | 81.25% | 71.54% | 96.92% | 82.32% | 81.78% |
| 8 | 70.59% | 75.00% | 89.17% | 95.00% | 91.99% | 82.63% |
| | CTF of $\mathcal{O}^5$ | 87.58% | $CTF^1$ of $\mathcal{O}^5$ | 79.64% | | |

Table 4.22: Automated Mediation: $\mathcal{O}^5$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 50.00% | 50.00% | 83.33% | 77.50% | 80.31% | 61.63% |
| 4 | 80.00% | 50.00% | 83.33% | 67.50% | 74.59% | 59.87% |
| 5 | 81.25% | 54.17% | 96.92% | 71.54% | 82.32% | 65.34% |
| 8 | 70.59% | 50.00% | 89.17% | 67.50% | 76.84% | 60.58% |
| | CTF of $\mathcal{O}^6$ | 78.52% | $CTF^1$ of $\mathcal{O}^6$ | 61.86% | | |

Table 4.23: Automated Mediation: $\mathcal{O}^6$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 45.83% | 64.71% | 83.03% | 88.18% | 85.53% | 73.67% |
| 4 | 80.00% | 70.59% | 83.33% | 89.17% | 86.15% | 77.60% |
| 5 | 75.00% | 70.59% | 95.00% | 89.17% | 91.99% | 79.88% |
| 6 | 50.00% | 70.59% | 67.50% | 89.17% | 76.84% | 73.58% |
| | CTF of $\mathcal{O}^8$ | 85.13% | $CTF^1$ of $\mathcal{O}^8$ | 76.18% | | |

Table 4.24: Automated Mediation: $\mathcal{O}^8$

### Semi-Automated Mediation

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 4 | 83.33% | 76.92% | 100.00% | 93.33% | 96.55% | 85.63% |
| 5 | 76.92% | 76.92% | 96.67% | 93.33% | 94.97% | 85.00% |
| 6 | 80.00% | 92.31% | 97.92% | 94.44% | 96.15% | 94.19% |
| 8 | 84.62% | 84.62% | 90.91% | 93.94% | 92.40% | 88.34% |
| | CTF of $\mathcal{O}^3$ | 95.01% | $CTF^1$ of $\mathcal{O}^3$ | 88.29% | | |

TABLE 4.25: Semi-Automated Mediation: $\mathcal{O}^3$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 76.92% | 83.33% | 93.33% | 100.00% | 96.55% | 89.46% |
| 5 | 92.31% | 100.00% | 97.22% | 100.00% | 98.59% | 99.29% |
| 6 | 80.00% | 100.00% | 95.83% | 100.00% | 97.87% | 98.92% |
| 8 | 92.31% | 100.00% | 91.67% | 100.00% | 95.65% | 97.78% |
| | CTF of $\mathcal{O}^4$ | 97.17% | $CTF^1$ of $\mathcal{O}^4$ | 96.36% | | |

TABLE 4.26: Semi-Automated Mediation: $\mathcal{O}^4$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 76.92% | 76.92% | 93.33% | 96.67% | 94.97% | 85.00% |
| 4 | 100.00% | 92.31% | 100.00% | 97.22% | 98.59% | 95.35% |
| 6 | 86.67% | 100.00% | 98.08% | 100.00% | 99.03% | 99.51% |
| 8 | 92.31% | 92.31% | 91.67% | 97.22% | 94.36% | 93.32% |
| | CTF of $\mathcal{O}^5$ | 96.74% | $CTF^1$ of $\mathcal{O}^5$ | 93.30% | | |

TABLE 4.27: Semi-Automated Mediation: $\mathcal{O}^5$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 92.31% | 80.00% | 94.44% | 97.92% | 96.15% | 87.33% |
| 4 | 100.00% | 80.00% | 100.00% | 95.83% | 97.87% | 88.04% |
| 5 | 100.00% | 86.67% | 100.00% | 98.08% | 99.03% | 92.44% |
| 8 | 92.31% | 80.00% | 91.67% | 95.83% | 93.70% | 86.31% |
| | CTF of $\mathcal{O}^6$ | 96.43% | $CTF^1$ of $\mathcal{O}^6$ | 88.53% | | |

TABLE 4.28: Semi-Automated Mediation: $\mathcal{O}^6$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | $TF^1$ |
|---|---|---|---|---|---|---|
| 3 | 84.62% | 84.62% | 93.94% | 90.91% | 92.40% | 88.34% |
| 4 | 100.00% | 92.31% | 100.00% | 91.67% | 95.65% | 93.95% |
| 5 | 92.31% | 92.31% | 97.22% | 91.67% | 94.36% | 93.32% |
| 6 | 80.00% | 92.31% | 95.83% | 91.67% | 93.70% | 93.00% |
| | CTF of $\mathcal{O}^8$ | 94.03% | $CTF^1$ of $\mathcal{O}^8$ | 92.15% | | |

TABLE 4.29: Semi-Automated Mediation: $\mathcal{O}^8$

### 4.4.1 Convergence toward Consensus

This section describes the effect of the mediation protocols on the consensus achieved between the ontology set in Experiment 1.

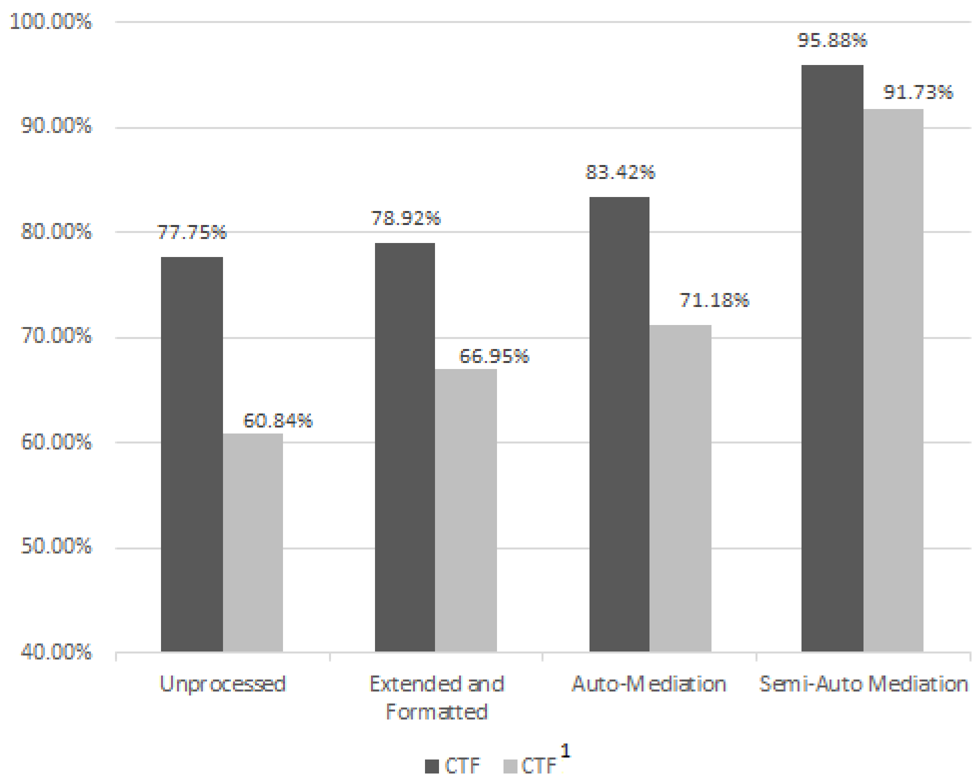


FIGURE 4.1: Experiment 1: convergence towards consensus

Figure 4.1 shows the general convergence of the ontologies towards a consensus as measured by CTF and CTF$^1$. This shows that the acquired ontologies are in general agreement after all the processes are complete. The most effective process, in terms of improving consensus, was the semi-automated mediation which improved the consensus by 12.46% for CTF and 20.55% for CTF$^1$. The pronounced increase in the CTF$^1$ measure during semi-automated mediation (as opposed to the more even improvement evident for CTF over the mediation process) indicates that the lexical recall increases proportionately more than the other metrics used to calculate CTF. Generally speaking, good results were recorded, indicating that a consensus has largely been achieved.

Figure 4.2 illustrates the consensus convergence of the models by individual ontology. Interestingly, the most conforming ontology is $\mathcal{O}^3$, which started off as the furthest away from consensus with the rest of the ontology set.

#### Evaluation against Gold Standard

Tables 4.30, 4.31 and 4.32 show the level of convergence of the ontology set towards a 'gold standard'. Each table lists the metrics used for determining the CTF as well as the lexical precision and recall.

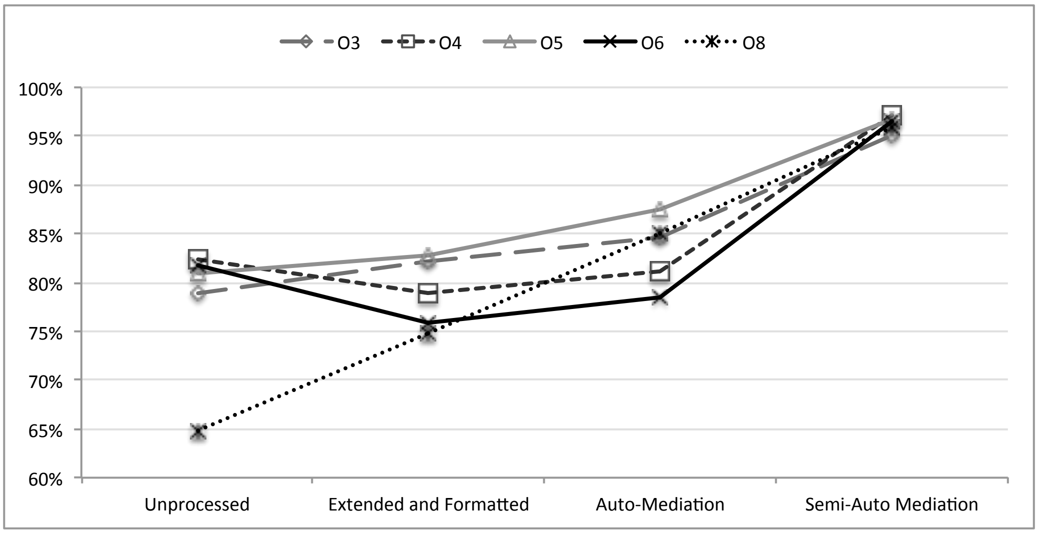


Figure 4.2: Convergence towards consensus by Ontology (CTF)

The 'gold standard' was written for a class exercise by a knowledge engineer and was used to provide a basic abstract description of the music store domain. This particular gold standard ontology is appropriate to use because it was designed as a general description of a music store and not to support any particular application. Therefore the gold standard is less likely to contain the nuanced and idiosyncratic entities and models which an application ontology might need. As expected, there is a modest degree of convergence (4.63%). All the ontologies converge towards the gold standard, which may indicate that no knowledge representation that conforms to the gold standard is lost at any stage. If the mediation had removed some entities or changed any models that were supported by the gold standard, then the individual ontology would be likely to diverge from the gold standard, even if the overall measure across the ontology set improves. $\mathcal{O}^5$ does see a reduction in the TF measure from the initial structure, to the post-automated mediation. However, after semi-automated mediation the convergence is an improvement on the initial structure. This indicates that a concept or concepts that correlate to the 'gold standard' are removed during the pruning stage, but then correctly reinstated during semi-automated mediation.

The comparison of lexical precision across the ontology set shows a general improvement obtained through mediation. This indicates that after mediation, there are fewer concepts in the acquired ontologies that fail to match concepts in the gold standard.

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF[1] |
|---|---|---|---|---|---|---|
| 3 | 41.67% | 50.00% | 89.33% | 82.50% | 85.78% | 63.18% |
| 4 | 50.00% | 35.00% | 92.86% | 70.24% | 79.98% | 48.69% |
| 5 | 40.00% | 30.00% | 87.50% | 74.31% | 80.36% | 43.69% |
| 6 | 33.33% | 40.00% | 88.75% | 71.35% | 79.11% | 53.13% |
| 8 | 43.75% | 35.00% | 57.62% | 77.98% | 66.27% | 45.81% |
| | CTF | 78.30% | CTF[1] | 50.90 | | |

Table 4.30: Vs. Gold Standard: Unprocessed

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF[1] |
|---|---|---|---|---|---|---|
| 2 | 41.67% | 50.00% | 89.33% | 82.50% | 85.78% | 63.18% |
| 4 | 46.67% | 35.00% | 92.86% | 70.24% | 79.98% | 48.69% |
| 5 | 43.75% | 35.00% | 91.43% | 70.24% | 79.44% | 48.59% |
| 6 | 33.33% | 40.00% | 88.75% | 71.35% | 79.11% | 53.13% |
| 8 | 47.06% | 40.00% | 83.75% | 73.96% | 78.55% | 53.01% |
| | CTF | 80.57% | CTF[1] | 53.32 | | |

Table 4.31: Vs. Gold Standard: Automatic Mediation

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF[1] |
|---|---|---|---|---|---|---|
| 3 | 69.23% | 45.00% | 100.00% | 80.09% | 88.95% | 59.76% |
| 4 | 58.33% | 35.00% | 100.00% | 70.24% | 82.52% | 49.15% |
| 5 | 53.85% | 35.00% | 95.24% | 70.24% | 80.85% | 48.85% |
| 6 | 53.33% | 40.00% | 96.88% | 71.35% | 82.18% | 53.81% |
| 8 | 61.54% | 40.00% | 87.50% | 73.96% | 80.16% | 53.37% |
| | CTF | 82.93% | CTF[1] | 52.99% | | |

Table 4.32: Vs. Gold Standard: Semi-Automatic Mediation

### 4.4.2 Object Properties

| Stage | TP (DR) | TR(DR) | CDRF |
|---|---|---|---|
| Pre-mediation | 38.18% | 82.18% | 52.14% |
| Auto-mediation | 39.74% | 74.62% | 51.86% |
| Semi-Automated Mediation | 48.81% | 81.83% | 61.14% |

Table 4.33: Effect of Mediation on Object Properties

Table 4.33 shows convergence towards consensus for the object properties in the acquired ontologies. While there is a degree of convergence (9%), it is significantly less than the consensus achieved between the concepts. This is included here for information only as no detailed discussion will be made of the object property mediation (for reasons outlined in Section 3.4).

## 4.5 Summary

Experiment 1 was carried out to acquire knowledge from participants divided into five groups. The participants were all given some basic understanding of knowledge modeling. Each group devised competency questions which were, filtered, processed and aggregated before being redistributed so that the participants could extend their models. From the groups, five ontologies were collected, each representing the group's knowledge of the music domain. These were first automatically, and then semi-automatically mediated to improve consensus. Significant convergence towards a consensus was achieved for the concept definitions.

# Chapter 5

# Experiment 2: Eliciting Concept Hierarchies from a Task-Aware crowd with *WordNet* support

## 5.1 Description

The objective of this experiment is to see if the approach being offered can be scaled up to accommodate input from a greater number of participants, while still identifying a consensus for a given domain. In Experiment 1, convergence towards a consensus was achieved by using the protocols outlined in Chapter 3, however there is a degree of contrivance in terms of a) the nature of the crowd and b) how we obtain coverage of the domain. The crowd in Experiment 1 is small, so to provide coverage the models have been artificially extended through the distribution of competency questions across the groups. In a real-world crowd scenario, competency questions are unlikely to be distributed in such a way. Indeed, competency questions may not exist in explicit form, with participants relying on general beliefs about how the domain should be described. For Experiment 2 the participants are required to devise competency questions, however these are not distributed and there is no requirement for extending the models after the initial collection period. Additionally, while the initial high-level concepts are suggested in the instructions, no explicit seed ontology is provided (unlike Experiment 1). The reason not to include an explicit seed ontology was that Experiment 3 seemed to suggest that similar convergence is achieved no matter what seed was used so the necessity of using seed at all should be investigated. As with Experiment 1, the chosen domain is the *Music Store*.

This experiment includes two additional processes that could reduce the number of mediation cases that require manual adjudication.

## 5.2 Objective

The objectives of these experiments are:

- To test the effectiveness of the protocols described in Chapter 3 in terms of gaining a consensus
- To provide a comparison for later experimentation which performs a similar task but without the participants being aware of the model building process
- To determine the scope of conflicting knowledge that will require a knowledge engineer to resolve
- To test the effectiveness of the enhanced protocols described in Chapter 3 by incorporating *WordNet*

## 5.3 Method

1. Participants provided with a domain description
2. Participants devise a set of competency questions
3. Model produced to solve competency questions
4. Invalid concepts removed using WordNet
5. Automated mediation applied
6. Semi-Automated mediation applied with additional question types generated

### Participation

A total of 22 participants took part in the model submission and mediation process during Experiment 2. Participants worked as individuals rather than groups. The primary model-building and submission was performed on February 12$^{\text{th}}$ 2014. The manual input needed for semi-automated mediation was performed over two sessions (March 6$^{\text{th}}$ and March 10$^{\text{th}}$). As with Experiment 1, the participants comprised final year BSc students in Computer Science (or a closely related field) and MSc students in Computer Science from the University of Liverpool.

In Experiment 1 we asked our participants to use *Moki* to produce the model; however, as there was time to offer additional technical tuition to students we instructed them to use *Protégé* instead, using the well-known tutorial by Horridge et. al. [49] (participants were instructed to follow the tutorial up to, and including Section 4.8). The reasons for using *Protégé* were twofold; firstly the coordination was simplified because no external set-up was needed; secondly, the participants were required to use *Protégé* as part of their standard assessment. Unlike Experiment 1, all submitted ontologies that are syntactically correct (i.e. did not produce errors when being processed) have been used. This means that all models are used, even when it is clear that the participant did not understand the given task.

### Restrictions

There were no restrictions placed on the participants in terms of how they could model the domain, however the following high level concepts were suggested: *Record*, *Song*, *Person*, *Genre*, *Gender*, *Group*. Due to the findings of Experiment 3, which suggested that useful models could be created without a seed ontology being provided, no seed ontology is used. [1]

### 5.3.1 Challenges of Scale

With the greater number of submitted models, and the removal of many of the restrictions and controls in place during Experiment 1, various challenges become apparent. Firstly, the manual input required for the semi-automated mediation phase could increase to a point whereby it becomes impractical. Secondly, with each participant using their own set of competency questions, there is a greater probability of unique concepts and concept relationships being present. To some degree, this can be addressed by adjusting the adoption thresholds, thus influencing the composition of the mediation outcomes. However, only adjusting the majority adoption threshold will reduce the number of mediation questions produced as this is the only automated mediation outcome that does not lead to manual input being required. Adjusting the minority adoption threshold only influences the potential for unresolved concept models being present in the final ontology (i.e. influences the potential number of concepts which the knowledge engineer will have to adjudicate over), and therefore does not reduce the number of mediation questions.

Given the results of Experiment 1, it is evident that the biggest burden on the participants is the number of *reinstate* questions produced; this is clear from Tables 4.8 and 4.9 which show that 30 reinstate questions are produced, compared with just 6 minority adoption questions. Therefore the best strategy to reduce the burden on the participants is to remove pruning/reinstate cases which are clearly invalid. From the outcome of Experiment 1, it is evident that certain concept modeling errors are common, and if identified and removed, could reduce the number of mediation questions considerably. With a small scale experiment it is possible to handle these errors through the pruning process; with 22 participants it is essential that some of the pruning cases are resolved automatically.

An analysis of the raw input from Experiment 1 shows that there are a number of cases where the validity of the concept is clearly incorrect. Table 5.1 outlines the three most common forms of invalid input

[1] Experiment 2 was chronologically carried out after Experiment 3, but as Experiment 2 was primarily aimed at addressing the shortcomings of Experiment 1 it made sense to order them in this way

| | |
|---|---|
| Instance as Concept | Where a user describes an individual instance as a concept |
| Spelling | Where, despite the edit distance allowance, the spelling is not a recognised word or phrase |
| Malformed | Where the use of non-conforming character sets are used (e.g. Chinese or Arabic letters) |

Table 5.1: Types of invalid concept labels

### 5.3.2 Using WordNet to Reduce Mediation Cases

*WordNet* is a lexical database that groups words into synsets. These synsets are interlinked by means of conceptual-semantic and lexical relations. As mentioned earlier (Section 2.5.1), *WordNet* has been used with some success to determine specific semantic relationships between concepts in a given model — in particular hyponym/hypernym relationships. Given the scope (in terms of words, phrases and concepts described) and accessibility of *WordNet* it was a relatively simple decision to use it to filter the pruning cases.

For the *spelling* and *malformed* invalidity types, it is simple matter of checking whether the terms exist in the WordNet dictionary. As WordNet also lists many phrases, such as *dance music* or *concept album*, so it will identify where a valid phrase is used and allow this to be used to create a reinstate question.

Identifying *instance as concept* is a more nuanced process. *WordNet* does list some instances, however it would be impossible to note all instances of all concepts. The methods used to identify instances in *WordNet* is based on a few basic rules, which are outlined by Miller and Hristea [74]. Firstly instances are always nouns; secondly they are proper nouns, which means that they should be capitalised; finally, they should be a unique entity, which implies that they should not have hyponyms (it is meaningless to have an instance of an instance) (ibid.). It should be stressed that these are basic rules which, by themselves and applied to a general text corpus, would produce many false positives. However, because our pruning cases are already of suspect validity due to their lack of support across the submitted models, and because of the need to reduce the number of reinstate cases, the kind of generalising bias of using such rules can be tolerated. On this basis, it is possible to formulate some basic rules in order to remove invalid concepts.

1. Check if the candidate concept is a valid spelling (dictionary check). If the concept is not contained in the *WordNet* dictionary then it is removed.

2. Check if the candidate concept is listed as a instance. If the concept is listed then it is removed.

3. Check if the candidate concept has a known instance as hyponym. If a hyponym exists, then the concept must be approved (the concept must be valid if it has known instances).

This process can be illustrated by looking at three candidate concepts: 'countrie', 'Belgium' and 'country' . By processing these three candidate concepts, 'countrie' would be rejected because it is not in the *WordNet* dictionary (because it is misspelled); 'Belgium' would be rejected as it is in the list of known instances; and 'country' would be accepted as it has known instances as hyponyms (e.g. Belgium).

It is accepted that useful concept descriptions may be lost at this point; however, this is felt to be a necessary measure to ensure the workability of the mediation protocols in terms of reducing the number of mediation questions.

### 5.3.3 Using *WordNet* to identify Synonyms

Along with using *WordNet* to remove invalid concepts, it has also been used to determine equivalence relationships between two submitted concepts. The premise is that concepts discovered to be within the same synset can be recognised as equivalent if given support through semi-automated mediation. While this will generate additional mediation questions, it is valuable as it can help identify where submitted ontologies are in agreement over a concept model, but where they have used different labels. This is intended to address the weakness of using concept labels exclusively to determine conceptual matches. It may also help reach a consensus over a particular concept model because we can aggregate the mediation support for equivalent concept models that share the same superclass, meaning that the minority and majority adoption thresholds may be met. Table 5.2 illustrates the possible outcomes if two concept models are deemed equivalent given the adoption thresholds outlined in Section 5.4.1. In the first line of Table 5.2, both concept models are majority adopted because their combined support meets the threshold; in the second example both concept models are majority adopted because they are equivalent and concept model A meets the majority adoption threshold; in the third example, neither concept the models nor their combined aggregate support, meets the minority adoption threshold, so both remain as pruning cases. The remaining two examples (in italics) show where equivalence could also be used to provide more minority adoption cases, but which were not considered here as this would require two rounds of questioning, first to validate the equivalence and then to determine if the resulting minority adoption cases should be majority adopted.

| Concept Model A | Spt. | Concept Model B | Spt. | Result |
|---|---|---|---|---|
| CD $\sqsubseteq$ Media | 55% | Compact Disc $\sqsubseteq$ Media | 12% | (A$\wedge$B) MA |
| Singer $\sqsubseteq$ Person | 80% | Vocalist $\sqsubseteq$ Person | 1% | (A$\wedge$B) MA |
| Compilation $\sqsubseteq$ Record | 1% | Collection $\sqsubseteq$ Record | 1% | (A$\wedge$B) pruned |
| *Player $\sqsubseteq$ Person* | *10%* | *Instrumentalist $\sqsubseteq$ Person* | *1%* | *(A$\wedge$B) mA* |
| *Cover $\sqsubseteq$ Song* | *1%* | *Cover Version $\sqsubseteq$ Song* | *3%* | *(A$\wedge$B) mA* |

Table 5.2: Example Equivalence Cases. Note that in each line of the table the left-hand side of the equations indicate those concepts contained within the same WordNet synset. For example in line 1 *CD* and *Compact Disc* have been identified as synonyms in *WordNet*. The right hand side of the equations indicate the specified parent concept according to the crowd input

The inclusion of equivalence validation will produce more mediation questions and therefore is counter to the objectives outlined in Section 5.3.2. However, Experiment 1 indicates that the biggest burden on the user during mediation is likely to be the production of reinstate questions during automated mediation — which this process will not effect. Moreover, the inclusion of this process is to show how we can go beyond the basic mechanism of label matching to achieve a more sophisticated notion as to what comprises a concept within a knowledge model, in this case identifying where two distinct labels in reality represent the same concept. It would be possible to simply set an equivalency rule that would avoid the need for generating these questions, however, using human input is in keeping with the approach in general and would help prevent the loss of valid concepts.

### 5.3.4 Obtaining Competency Questions and Initial Ontology structures

Participants were asked to formulate six competency questions that could be used to query the music store domain. No further instructions were given. No seed ontology was provided, although several base classes were suggested in the instructions given to students, see Appendix B. As the competency questions are not being analysed or processed, there is little value in presenting them here.

## 5.4 Results

#### Formatting

Table 5.3 shows the number of edits that were made across the ontology set for Experiment 2. The same formatting rules as outlined in Section 4.3.2 were used. The formatting was performed prior to using WordNet.

| $\mathcal{O}$ | De-Cap. | Comp. | Edit Dist. | $\mathcal{O}$ | De-Cap. | Comp. | Edit Dist. |
|---|---|---|---|---|---|---|---|
| $\mathcal{O}$ 1 | 11 | 5 | 0 | $\mathcal{O}$ 12 | 9 | 3 | 0 |
| $\mathcal{O}$ 2 | 15 | 0 | 1 | $\mathcal{O}$ 13 | 77 | 21 | 2 |
| $\mathcal{O}$ 3 | 14 | 1 | 0 | $\mathcal{O}$ 14 | 10 | 1 | 0 |
| $\mathcal{O}$ 4 | 17 | 0 | 0 | $\mathcal{O}$ 15 | 5 | 1 | 0 |
| $\mathcal{O}$ 5 | 22 | 4 | 1 | $\mathcal{O}$ 16 | 9 | 1 | 0 |
| $\mathcal{O}$ 6 | 11 | 2 | 0 | $\mathcal{O}$ 17 | 25 | 2 | 0 |
| $\mathcal{O}$ 7 | 34 | 8 | 0 | $\mathcal{O}$ 18 | 7 | 1 | 0 |
| $\mathcal{O}$ 8 | 13 | 0 | 0 | $\mathcal{O}$ 19 | 2 | 0 | 0 |
| $\mathcal{O}$ 9 | 15 | 1 | 0 | $\mathcal{O}$ 20 | 12 | 0 | 0 |
| $\mathcal{O}$ 10 | 14 | 3 | 0 | $\mathcal{O}$ 21 | 21 | 0 | 0 |
| $\mathcal{O}$ 11 | 9 | 0 | 1 | | | | |

Table 5.3: Formatting Edits

### 5.4.1 Automated Mediation

#### Thresholds

| | |
|---|---|
| Minority Adoption Threshold (Automated Mediation) | 10% |
| Majority Adoption Threshold (Automated Mediation) | 33% |
| Minority Adoption Threshold (Reinstate) | 50% |
| Majority Adoption Threshold (Reinstate / mA →MA) | 80% |

Table 5.4: Adoption Thresholds for Experiment 2

Table 5.4 gives an overview of the thresholds used in Experiment 2. As noted earlier, in order to increase the scale and in order to cope with a less restricted input, the thresholds need to be altered. Because it was not practical to undertake the experiment multiple times to determine the ideal thresholds (pruning threshold and majority adoption threshold), thresholds were chosen so that the outcomes (i.e. proportion of concepts being pruned, the proportion minority adopted and the proportion being majority adopted) in automated mediation were as similar to Experiment 1 as possible. In Experiment 1 8.69% of conflicts were minority adopted and 13.04% were majority adopted (before conflict resolution). The adoption thresholds for automated mediation were therefore set at 10% for minority adoption, and at 33% for majority adoption; this resulted in 8.33% of concepts being minority adopted and 10.41% being majority adopted in Experiment 2. Given the 96 distinct concepts acquired in Experiment 2 and not removed using WordNet, the set thresholds mean that the concept will be pruned if two or less submitted ontologies agree on the concept; it is majority adopted if more than seven ontologies agree; otherwise it is minority adopted. There was no reason to adjust the reinstate adoption thresholds.

#### Filtering Pruning Cases Using WordNet

Having determined the thresholds, the first step in Experiment 2 was to identify the pruned cases, and remove as many as possible from consideration using *WordNet* (see

Section 5.3.2). Table 5.5 lists all the cases where the *WordNet* was used to remove submitted concepts from the submitted models. In the majority of these cases the Dictionary Check (DC) was used. Where the '?' character is present indicates that a non-standard character was used. In total 157 concepts were removed, leaving just 96 to be further mediated. This may seem excessive; however, consider that in one of the submitted models a list of 66 languages is incorporated which included various non-latin symbols — how they managed this is a mystery, but illustrative of the kind of input that might constitute an outlier in a crowdsourcing context. The Instance Check (IC) also proved useful in reducing the number of mediation questions.

While eliminating concepts through *WordNet* should prove effective in reducing the number of mediation cases (meaning that fewer reinstate questions are generated) the amount of useful concepts which are removed needs to be considered. It should also be noted that the non-latin entities had to be manually removed in order to run the convergence evaluation program; therefore the unprocessed set does not include these concepts in Section 6.7. In Chapter 7 an assessment of how much useful input is lost during this process will be made.

| **Pruning Candidates** | **DC** | **IC.** |
|---|---|---|
| zazaki | ✓ | |
| nedersaksies | ✓ | |
| vro | ✓ | |
| aragons | ✓ | |
| ????? | ✓ | |
| ???????? | ✓ | |
| sardu | ✓ | |
| ???????? | ✓ | |
| ?????????? | ✓ | |
| caribbean and latin american 4.1 brazilian | ✓ | |
| south and southeast asian | ✓ | |
| eesti | ✓ | |
| franais | ✓ | |
| ?????? | ✓ | |
| interlingua | | ✓ |
| crecord | ✓ | |
| sicilianu | ✓ | |
| bn-lm-g | ✓ | |
| dj | | ✓ |
| ????????? | ✓ | |
| ?????????? | ✓ | |
| ????? | ✓ | |

Table 5.5 – *Continued on next page*

| **Pruning Candidates** | **DC** | **IC** |
|---|---|---|
| magyar | ✓ | |
| collection record type | ✓ | |
| u.s billboard | ✓ | |
| galego | ✓ | |
| popular record | ✓ | |
| hip-hop | | ✓ |
| ??????? | ✓ | |
| espaol | ✓ | |
| slenska | ✓ | |
| plattdtsch | ✓ | |
| singer songwriter | ✓ | |
| writing award | ✓ | |
| boarisch | ✓ | |
| cd | | |
| afrikaans | ✓ | |
| igbo | | ✓ |
| other latin | ✓ | |
| romn? | ✓ | |
| girl band | ✓ | |
| boy band | ✓ | |
| dansk | ✓ | |
| cymraeg | ✓ | |
| ??afrikaans | ✓ | |
| trke | ✓ | |
| sloven??ina | ✓ | |
| hiphop | ✓ | |
| instrumentalist songwriter | ✓ | |
| malti | ✓ | |
| ?????? | ✓ | |
| ???????? | ✓ | |
| best singer award | ✓ | |
| special record type | ✓ | |
| tatara | ✓ | |
| ??????? | ✓ | |
| ltzebuergesch | ✓ | |
| dvd | | ✓ |
| ?????? | ✓ | |
| avant-garde | ✓ | |
| limburgs | ✓ | |

Table 5.5 – *Continued on next page*

| **Pruning Candidates** | **DC** | **IC** |
|---|---|---|
| ripoarisch | ✓ | |
| extended play | ✓ | |
| ?????? | ✓ | |
| singing award | ✓ | |
| latina | ✓ | |
| sloven?ina | ✓ | |
| ?????????? | ✓ | |
| ?emait??ka | ✓ | |
| ?????? | ✓ | |
| norsk nynorsk | ✓ | |
| hrvatski | ✓ | |
| ?????? | ✓ | |
| easy listening | ✓ | |
| az?rbaycanca | ✓ | |
| lietuvi? | ✓ | |
| bosanski | ✓ | |
| ?? | ✓ | |
| kiswahili | | ✓ |
| celtic traditional music | ✓ | |
| international | | ✓ |
| ?????????????? | ✓ | |
| suomi | | ✓ |
| hip hop | | ✓ |
| ????? | ✓ | |
| euskara | ✓ | |
| balkan music | ✓ | |
| tagalog | | ✓ |
| ??? | ✓ | |
| basa jawa | ✓ | |
| cebuano | | ✓ |
| counrty group | ✓ | |
| ?e?tina | ✓ | |
| furlan | ✓ | |
| ???? | ✓ | |
| ??????? | ✓ | |
| simple english | ✓ | |
| basa sunda | ✓ | |
| bahasa melayu | | ✓ |
| east asian | ✓ | |

Table 5.5 – *Continued on next page*

| **Pruning Candidates** | **DC** | **IC** |
|---|---|---|
| soomaaliga | ✓ | |
| ?????? | ✓ | |
| alemannisch | ✓ | |
| bamanankan | ✓ | |
| ??????? | ✓ | |
| ?????????? | ✓ | |
| occitan | | ✓ |
| brezhoneg | ✓ | |
| asturianu | ✓ | |
| ??? | ✓ | |
| norsk bokml | ✓ | |
| ??????? | ✓ | |
| nederlands | | ✓ |
| polski | ✓ | |
| art music | ✓ | |
| folk music of china | ✓ | |
| mixed band | ✓ | |
| lumbaart | ✓ | |
| sax | | ✓ |
| shqip | ✓ | |
| malagasy | ✓ | |
| nordic folk music | ✓ | |
| latvie?u | ✓ | |
| yorb | ✓ | |
| ti?ng vi?t | ✓ | |
| singer instrumentalist | ✓ | |
| nouormand | ✓ | |
| ilokano | ✓ | |
| portugus | ✓ | |
| live record type | ✓ | |
| srpski | ✓ | |
| ??????? | ✓ | |
| band member | ✓ | |
| arecord | ✓ | |
| deutsch | ✓ | |
| winaray | ✓ | |
| runa simi | ✓ | |
| gaeilge | ✓ | |
| classic record | ✓ | |

Table 5.5 – *Continued on next page*

| **Pruning Candidates** | **DC** | **IC** |
|---|---|---|
| n?huatl | ✓ | |
| corsu | ✓ | |
| catal | ✓ | |
| svenska | ✓ | |
| italiano | ✓ | |
| froyskt | ✓ | |
| brecord | ✓ | |
| rhythm and blues | | ✓ |
| esperanto | | ✓ |
| abassist | ✓ | |
| agroup | ✓ | |
| asinger | ✓ | |
| asong | ✓ | |
| bbassist | ✓ | |
| bsong | ✓ | |
| bgroup | ✓ | |
| csinger | ✓ | |
| csong | ✓ | |
| dsong | ✓ | |

TABLE 5.5: Entities removed through WordNet filter

### Automated Mediation Overview

Table 5.6 shows the outcome for automated mediation on the remaining 96 distinct concepts. The automated mediation was performed with a pruning threshold of 3% (greater than 2 concept models in agreement) and a majority adoption threshold of 33% (greater than 7 concept models in agreement). Column $f$ denotes the confidence or frequency with which the concept appears across the ontology set. The next four columns (*MA, mA, p*) denote the outcome of the mediation where *MA* denotes the majority adoption outcome, *mA* the minority adoption outcome, *p* a pruning outcome and *c* a conflict case.

| | | **Outcome** | | | |
|---|---|---|---|---|---|
| **Concept** | ***f*** | **MA** | **mA** | **p** | **c** |
| african | 1 | × | × | ✓ | × |
| album | 2 | × | × | × | ✓ |
| animal | 1 | × | × | ✓ | × |
| artist | 5 | × | ✓ | × | × |
| asian | 1 | × | × | ✓ | × |
| author[1] | 1 | × | × | ✓ | × |

Table 5.6 – *Continued on next page*

| | | Outcome | | | |
|---|---|---|---|---|---|
| **Concept** | ***f*** | **MA** | **mA** | **p** | **c** |
| award | 1 | × | × | ✓ | × |
| b | 4 | × | ✓ | × | × |
| bahasa indonesia | 1 | × | × | ✓ | × |
| band | 2 | × | × | ✓ | × |
| bassist | 1 | × | × | ✓ | × |
| beatles | 1 | × | × | ✓ | × |
| blue | 1 | × | × | ✓ | × |
| blues | 3 | × | × | × | ✓ |
| classical | 1 | × | × | ✓ | × |
| comedy | 1 | × | × | ✓ | × |
| company | 1 | × | × | ✓ | × |
| composer | 4 | × | ✓ | × | × |
| country | 4 | × | × | × | ✓ |
| country music | 1 | × | × | ✓ | × |
| country record | 1 | × | × | ✓ | × |
| country singer | 1 | × | × | ✓ | × |
| drummer | 1 | × | × | ✓ | × |
| duo | 3 | × | ✓ | × | × |
| electronic | 2 | × | × | × | ✓ |
| ep | 2 | × | × | × | ✓ |
| era | 1 | × | × | ✓ | × |
| european | 1 | × | × | ✓ | × |
| female | 1 | × | × | ✓ | × |
| folk | 3 | × | ✓ | × | × |
| gender | 9 | ✓ | × | × | × |
| genre | 22 | ✓ | × | × | × |
| group | 21 | ✓ | × | × | × |
| guitarist | 1 | × | × | ✓ | × |
| ha | 1 | × | × | ✓ | × |
| hip-hop | 1 | × | × | ✓ | × |
| instrument | 1 | × | × | ✓ | × |
| instrumentalist | 1 | × | × | ✓ | × |
| jazz | 4 | × | × | × | ✓ |
| john lennon | 1 | × | × | ✓ | × |
| language | 2 | × | × | ✓ | × |
| lp | 3 | × | × | × | ✓ |
| lyric | 1 | × | × | ✓ | × |
| lyricist | 3 | × | ✓ | × | × |

Table 5.6 – *Continued on next page*

| | | Outcome | | | |
|---|---|---|---|---|---|
| **Concept** | ***f*** | **MA** | **mA** | **p** | **c** |
| male | 1 | × | × | ✓ | × |
| melody | 1 | × | × | ✓ | × |
| member | 1 | × | × | ✓ | × |
| metal | 1 | × | × | ✓ | × |
| musician | 1 | × | × | ✓ | × |
| other | 1 | × | × | ✓ | × |
| person | 22 | ✓ | × | × | × |
| piano | 1 | × | × | ✓ | × |
| pop | 8 | × | × | × | ✓ |
| popular music | 1 | × | × | ✓ | × |
| producer | 3 | × | ✓ | × | × |
| record | 18 | ✓ | × | × | × |
| rock | 10 | ✓ | × | × | × |
| rock music[1] | 1 | ✓ | × | × | × |
| singer | 14 | ✓ | × | × | × |
| single | 3 | × | × | × | ✓ |
| ska | 2 | × | × | × | ✓ |
| song | 21 | ✓ | × | × | × |
| tik tok | 1 | × | × | ✓ | × |
| trio | 3 | × | ✓ | × | × |
| vocalist | 1 | × | × | ✓ | × |
| writer | 8 | ✓ | × | × | × |
| year | 3 | × | × | × | ✓ |
| zither | 1 | × | × | ✓ | × |

TABLE 5.6: Automated Mediation Overview

### Note on thresholds

In Experiments 1 the minority adoption threshold is set at a minimum of two matching concepts across the ontology set, any conflict would also be a minority adopted. In Experiment 2, where the participation is much greater, the minority threshold is set at three matching concepts, therefore it is possible that a conflict could exist in the concept model of two concepts that would normally be pruned. If this is the case, then each should be considered a pruning case.

[1] *Author* and *Rock* was determined to be equivalent to the majority adopted concepts *rock* and *Writer* respectively. Therefore both of these equivalence were majority adopted also as equivalent concepts (see Table 5.7)

### 5.4.2 Semi-Automated Mediation

To deal with the increased participation, the composition of the mediation questions has been altered from Experiment 1. Below is a description of the two new sets of questions which have been generated. The 'either/or' questions are intended to reduce the burden on the participants, while the purpose of the equivalence questions is to include more of the submitted knowledge into the final knowledge representation.

#### Equivalence Questions

Using the WordNet, a synset of each concept is generated and matched against all the other submitted concepts. If a match is made, and the concepts share the same direct superclass, then the particpants are asked to whether the matched concepts are in fact equivalent. If equivalence is confirmed then the support for both concept models is aggregated, meaning that the minority adoption threshold or the majority adopted threshold may be reached.

It is acknowledged that by generating the equivalence questions we may increase the question burden on the participants (thus countering the benefit of producing either/or questions and other measures to reduce mediation questions).

#### WordNet Equivalency

Checks were made for equivalent concepts, where a reinstate candidate is deemed to be equivalent if the terms being compared are contained in the same WordNet synset *and* if the equivalency is endorsed by two-thirds of the particpants.

#### Either/Or Questions

Table 5.8 outlines the conflict resolution questions generated to resolve conflict cases in Experiment 2. Where two or more instances of a concept exist with differing concept models (a *conflict* outcome in Table 5.6), a composite question can be generated that determines which concept model (if any) should be majority adopted. In Experiment 1, the only concept that remained in conflict after automated mediation was *composer*, and no attempt to resolve this was made through generating mediation questions. In Experiment 2, many more conflicts were generated; therefore, by generating some additional questions, we can attempt to both resolve conflict and reach consensus on adoption. The 'either / or' questions present both concept models to the participants who are invited to choose 'A' or 'B', or alternatively they can choose not to commit to either of the concepts models. At present these questions are generated even if the concept does not meet the minority adoption threshold. Also, there is no mechanism for generating questions for three or more competing concept models; where this has occurred (once in Experiment 2), the concept model with the least support is dropped. The conflict is resolved if the reinstate majority adoption threshold (80%) is met, otherwise it remains in conflict and is therefore sent for expert adjudication. Table 5.8 also contains information

TABLE 5.7: Experiment 2: Equivalency Question Overview

| ≡ | $\mathcal{O}1$ | $\mathcal{O}2$ | $\mathcal{O}3$ | $\mathcal{O}4$ | $\mathcal{O}5$ | $\mathcal{O}6$ | $\mathcal{O}7$ | $\mathcal{O}8$ | $\mathcal{O}9$ | $\mathcal{O}10$ | $\mathcal{O}11$ | $\mathcal{O}12$ | $\mathcal{O}13$ | $\mathcal{O}14$ | $\mathcal{O}15$ | $\mathcal{O}16$ | $\mathcal{O}17$ | $\mathcal{O}18$ | $\mathcal{O}19$ | $\mathcal{O}20$ | $\mathcal{O}21$ | $\mathcal{O}22$ | ✓ | × | Outcome |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| writer ≡ author | ✓ | × | × | ✓ | × | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | × | ✓ | × | ✓ | ✓ | ✓ | ✓ | ✓ | 77.27% | 22.73% | ✓ |
| singer ≡ vocalist | ✓ | | × | × | × | ✓ | ✓ | ✓ | ✓ | ✓ | × | ✓ | ✓ | ✓ | | × | × | ✓ | ✓ | ✓ | ✓ | × | 59.09% | 31.82% | |
| instrumentalist ≡ musician | ✓ | × | ✓ | ✓ | ✓ | × | ✓ | ✓ | × | × | × | × | × | × | ✓ | × | × | × | ✓ | ✓ | ✓ | × | 45.45% | 54.55% | |
| lyrics ≡ language | × | | × | ✓ | × | × | ✓ | × | × | × | ✓ | × | × | × | × | × | × | × | ✓ | × | × | × | 18.18% | 77.27% | × |
| rock music ≡ rock | × | | ✓ | ✓ | × | ✓ | ✓ | × | ✓ | ✓ | × | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | × | × | ✓ | ✓ | ✓ | 68.18% | 27.27% | ✓ |

on whether or not each participant changed their mind in relation to their initial concept model specification. The outcomes circled (L) show where a participant has stuck with the concept model submitted in their original ontology. Those outcomes squared ($\boxed{L}$) show where a participant has changed their mind, if there is an empty square a participant has changed their positions so that they no longer support their original concept model, but have not committed to a new concept model. A more complete discussion of participants' behaviour during mediation can be found in Chapter 7.

**Reinstate Questions**

The outcome of the manual mediation over the reinstatement candidates and the minority adoption to majority adoption candidates are listed in 5.9. The minority adoption threshold was set at 50% and the majority adoption threshold was set at 80%. These thresholds are the same as those outlined in Experiment 1 (see Table 4.8).

### 5.4.3 Convergence toward Consensus

This Section describes the effect of the mediation protocols on the convergence towards a consensus achieved between the ontology set in Experiment 2. Due to space considerations, the raw data used to calculate the CTF and $CTF^1$ has been put online and can be found at `https://github.com/roscminni/crowdsourcing-semantic-resources/blob/master/EXP-1B.xlsx`.

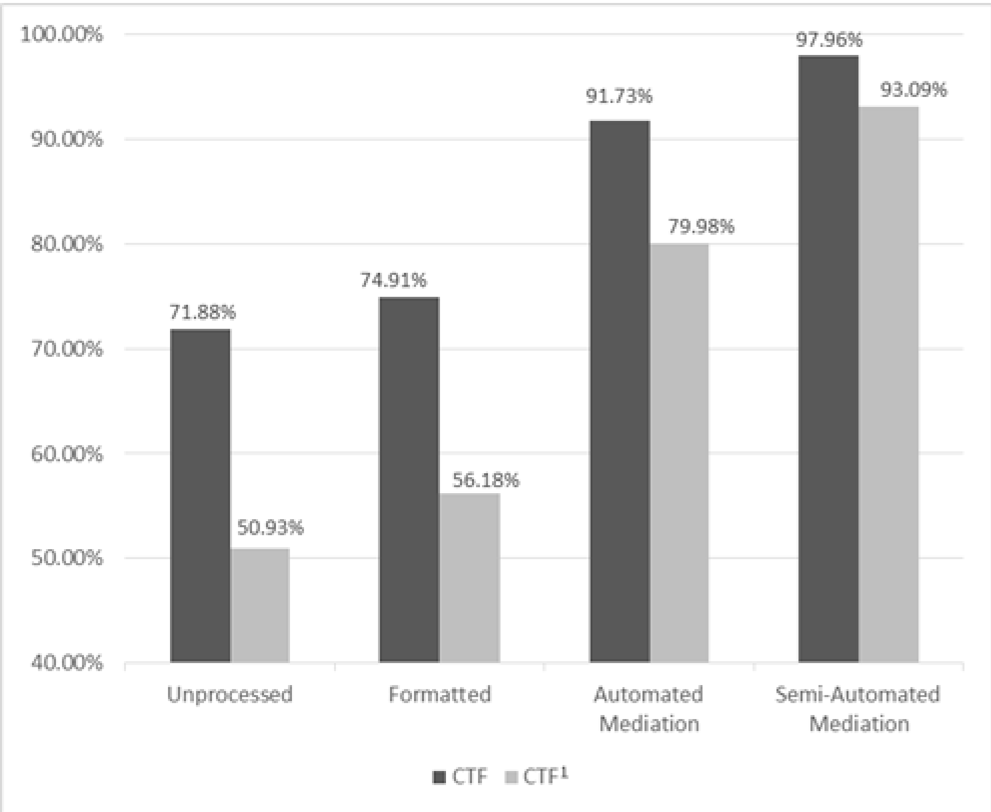


FIGURE 5.1: Experiment 2: convergence towards consensus

TABLE 5.8: Semi-Automated Mediation Overview for Experiment 2 - ‘Either / Or’ Majority Adoption Questions

| $A \vee$ B | $\mathcal{O}^1$ | $\mathcal{O}^2$ | $\mathcal{O}^3$ | $\mathcal{O}^4$ | $\mathcal{O}^5$ | $\mathcal{O}^6$ | $\mathcal{O}^7$ | $\mathcal{O}^8$ | $\mathcal{O}^9$ | $\mathcal{O}^{10}$ | $\mathcal{O}^{11}$ | $\mathcal{O}^{12}$ | $\mathcal{O}^{13}$ | $\mathcal{O}^{14}$ | $\mathcal{O}^{15}$ | $\mathcal{O}^{16}$ | $\mathcal{O}^{17}$ | $\mathcal{O}^{18}$ | $\mathcal{O}^{19}$ | $\mathcal{O}^{20}$ | $\mathcal{O}^{21}$ | $\mathcal{O}^{22}$ | $A$ | B | Res. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (album $\sqsubseteq$ record) $\vee$ (album $\sqsubseteq$ recording) | A | A | $\boxed{A}$ | B | A | A | | A | A | A | B | | | B | A | B | B | A | A | B | B | B | 54.55% | 36.36% | |
| (blues $\sqsubseteq$ genre) $\vee$ (blues $\sqsubseteq$ pop._music) | A | | A | A | A | B | | A | A | A | A | A | A | A | A | A | A | $\boxed{A}$ | A | A | A | A | 86.36% | 4.55% | A |
| (electronic $\sqsubseteq$ genre) $\vee$ (electronic $\sqsubseteq$ pop._music) | A | | A | A | A | A | B | A | A | A | A | A | A | A | A | B | A | $\boxed{A}$ | A | A | A | A | 86.36% | 9.09% | A |
| (ep $\sqsubseteq$ record) $\vee$ (ep $\sqsubseteq$ recording) | A | | $\boxed{A}$ | | A | A | B | | A | | | A | | $\boxed{B}$ | A | B | B | | A | B | B | B | 36.36% | 31.82% | |
| (jazz $\sqsubseteq$ genre) $\vee$ (jazz $\sqsubseteq$ pop._music) | A | | A | A | A | B | A | A | A | A | B | A | A | A | A | A | $\boxed{A}$ | A | A | A | A | A | 86.36% | 9.09% | A |
| (lp $\sqsubseteq$ record) $\vee$ (lp $\sqsubseteq$ recording) | B | | $\square$ | A | A | A | A | | A | | | A | | $\boxed{B}$ | A | B | B | | A | $\boxed{B}$ | B | B | 36.36% | 31.82% | |
| (single $\sqsubseteq$ record) $\vee$ (single $\sqsubseteq$ recording) | A | | B | | A | A | A | A | A | A | B | | | B | A | B | | A | A | $\boxed{B}$ | B | B | 45.45% | 31.82% | |
| (ska $\sqsubseteq$ genre) $\vee$ (ska $\sqsubseteq$ pop._music) | B | | A | | A | B | B | | | | | | | A | A | B | A | A | A | A | A | A | 45.45% | 18.18% | |
| (song $\sqsubseteq$ release) $\vee$ (song $\sqsubseteq$ thing) | $\boxed{A}$ | B | B | $\square$ | B | $\boxed{A}$ | $\boxed{A}$ | $\boxed{A}$ | $\boxed{A}$ | $\square$ | $\boxed{A}$ | B | $\boxed{A}$ | B | $\square$ | B | $\square$ | $\square$ | $\boxed{A}$ | $\square$ | $\square$ | $\square$ | 36.36% | 27.27% | |
| (year $\sqsubseteq$ thing) $\vee$ (year $\sqsubseteq$ date) | B | B | B | B | B | $\boxed{B}$ | A | $\boxed{B}$ | B | B | B | B | B | | | | | B | B | B | | B | 4.55% | 72.73% | B |
| (country $\sqsubseteq$ thing) $\vee$ (country $\sqsubseteq$ genre) | B | B | B | B | B | B | B | B | $\boxed{B}$ | B | B | B | A | | B | | | B | B | B | | B | 4.55% | 77.27% | |
| (pop $\sqsubseteq$ popular_music) $\vee$ (pop $\sqsubseteq$ genre) | B | B | B | B | B | B | B | B | B | B | B | B | $\boxed{A}$ | B | B | | A | $\boxed{B}$ | B | B | | B | 9.09% | 81.82% | B |

| Concept | *MA* | *mA* | Concept | *MA* | *mA* |
|---|---|---|---|---|---|
| african | × | × | guitarist | × | ✓ |
| animal | × | × | ha | × | × |
| artist | ✓ | × | b | × | ✓ |
| asian | × | × | hip-hop | × | ✓ |
| award | × | ✓ | instrument | × | ✓ |
| bahasa indonesia | × | × | instrumentalist | × | × |
| trio | × | ✓ | language | × | ✓ |
| band | × | ✓ | john lennon | × | × |
| bassist | × | ✓ | lyric | × | × |
| producer | ✓ | × | lyricist | × | ✓ |
| beatles | × | × | male | × | ✓ |
| blue | × | × | melody | × | × |
| classical | × | ✓ | member | × | × |
| company | × | × | metal | × | ✓ |
| country music | × | × | musician | × | ✓ |
| comedy | × | × | composer | ✓ | × |
| country record | × | × | other | × | × |
| country singer | × | × | piano | × | × |
| drummer | × | ✓ | popular music | ✓ | × |
| duo | × | ✓ | folk | × | ✓ |
| era | × | × | tik tok | × | × |
| european | × | × | vocalist | × | × |
| female | × | × | zither | × | × |

TABLE 5.9: Overview: Reinstate and minority adoption to majority adoption questions

Figure 5.1 shows the general convergence of the ontologies towards a consensus in Experiment 2, as measured by CTF and CTF[1]. As with Experiment 1, the acquired ontologies show a good convergence towards a consensus after all the processes are complete. The most effective process, in terms of improving consensus, was the automated mediation (as opposed to the Experiment 1 where the semi-automated mediated does the most work). This difference is illustrated in Figures 5.2 and 5.3 which compare the effect of each mediation stage on the CTF and CTF[1] convergence measure. Given the inclusion of the additional WordNet processes (See Section 5.4.1), this is what we would expect as more concepts are resolved (by being removed) in automated mediation and therefore proportionately less concepts require manual mediation. So, in terms of reducing the burden of reinstatement, incorporating the WordNet processes has been successful. What remains to be seen is how much valid knowledge is discarded.

## 5.5 Summary

Experiment 2 significantly expanded the participation to 22 participants, each submitting an ontology to the mediation process. To accommodate this expansion a method to reduce the burden of manual mediation was found that exploited *WordNet*. The degree

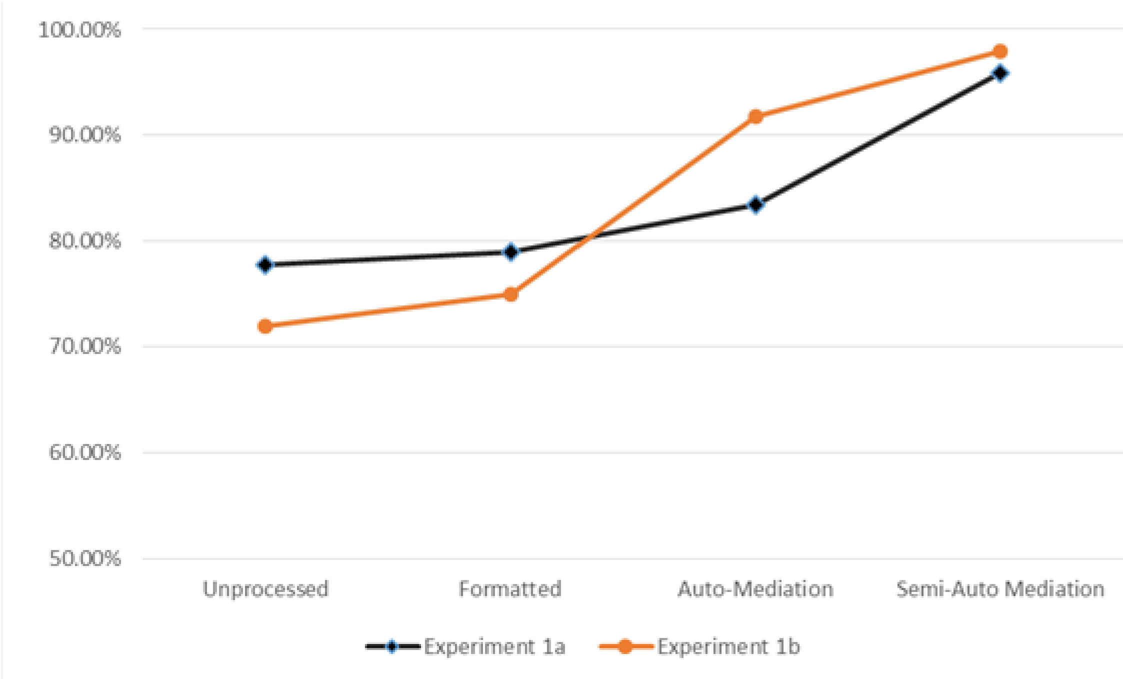


FIGURE 5.2: Comparison of CTF convergence

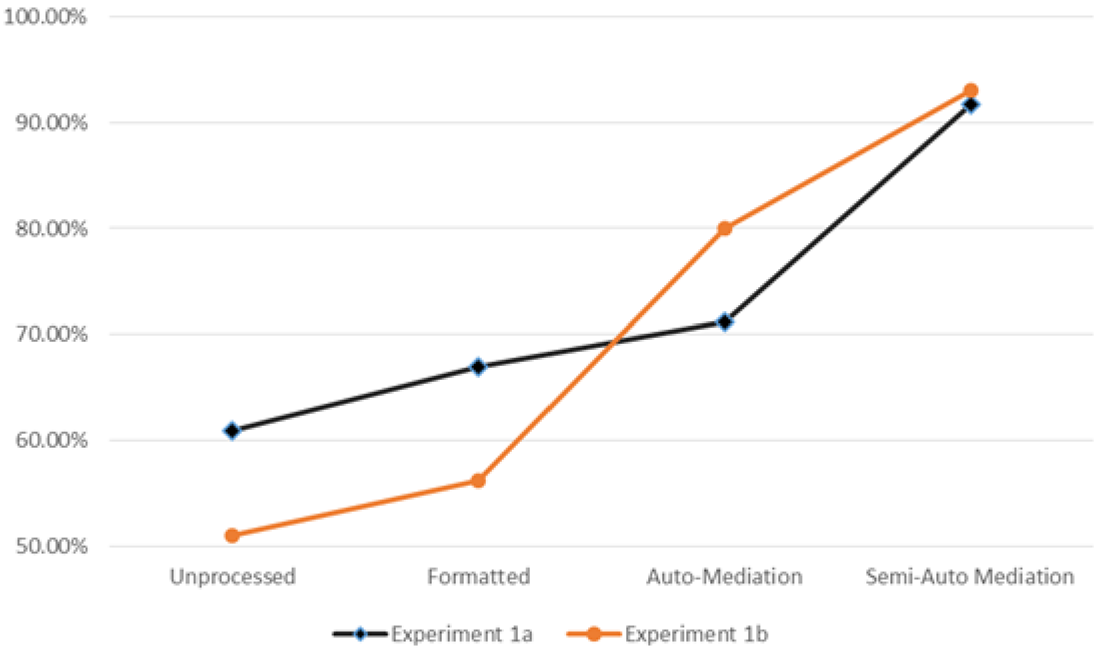


FIGURE 5.3: Comparison of CTF[1] convergence

of convergence towards a consensus achieved in Experiment 2 was similar to that of Experiment 1. A full qualitative analysis of these results is given in Chapter 7.

# Chapter 6

# Experiment 3: Building a Consensual Models from Task-Unaware Crowds

## 6.1 Description

The experiment described in this chapter provides an implementation of the crowdsourcing protocols described in Chapter 3. For this experiment, a crowd that has no awareness of the knowledge engineering task is used. The premise tested here is that, given a compatible motivation and an appropriate user interface, a crowd that is unaware of the knowledge modelling activity they are engaged in can still be leveraged to produce useful representation of a given domain. As with the previous experiment, the convergence towards consensus will be measured according to the methods outlined in section 3.2.

A key aspect of this experiment is test whether a crowd of users whose understanding of knowledge engineering is unknown can be motivated to produce useful information which can subsequently be used to build a consensual model of the domain. As discussed in Chapter 2 there are many successful applications of crowdsourcing that acquire knowledge or perform a knowledge engineering task as a by-product of the primary activity carried out by the user. The experiment presented here aims to obtain crowdsourced information indirectly by inviting users to contribute towards a digital mapping resource (henceforth referred to as the *campus map*). If this can be done in such a way that the information they give can be efficiently translated into a basic but useful ontology structure, then the knowledge acquisition of using this approach will be proven.

In this experiment, the individual task asked of each participant is to contribute objects to the campus map that represent the buildings and services that are relevant to student life. The collective objective of this crowd is to build a mapping resource aimed at helping new students and visitors familiarise themselves with formal and informal services within and around a university campus. By choosing a themed map we are simulating an online community crowd by providing a qualification for input. Anyone can contribute to the map resource, but they should only be motivated to do so if they have some knowledge to give. While the participants have the objective of building the campus map, the owner, that is to say the person or group who sets-up the experiment, has the alternative objective of obtaining the knowledge representation from the participant input.

Each participant is assigned one of three 'seed ontologies' when they sign up to take part. By placing map objects, each participant extends the seed ontology. Automated mediation is performed on the seed ontologies. While it would have been ideal to provide all our participants with the same level of seed ontology, there was a risk that providing one which was too abstract might lead to chaotic input, while providing a developed structure might stifle independent and alternative models being acquired and severely limit the number of new concepts being submitted. Therefore, the first part of our experiment is aimed at determining the right level of seed ontology. The most appropriate level will be the most abstract seed ontology that can still be automatically mediated so that average CTF compares with that achieved in the earlier experiment (83.42%). A full description of the seed ontologies used and the impact which each might have will be provided in Section 6.2.

The second part of the experiment is to take the group of ontologies deemed to be the best and then perform the semi-automated mediation to see if we can achieve a similar convergence towards consensus to that demonstrated in the earlier experiment.

## 6.2 Objective

The objectived of this experiment is the same as experiment in terms of achieving a good level of convergence towards a consensus within the ontology set. The two additional objectives are:

- To determine the appropriate seed ontology for the given scenario
- To perform semi-automated mediation and compare the results with that obtained from the task-aware crowd

## 6.3 Method

1. The campus map is used to collect map objects and extend the ontology
2. Automated mediation then applied
3. Best seed ontology chosen
4. Semi-Automated mediation is applied to best seed

### Motivations

As illustrated in figure 6.1 the participant is first asked to specify a map object that represents an area or structure that would be of interest to a new student or visitor. For example, a user might want to describe a local bakery which is a good place to have lunch. To do this they will plot the outline of the bakery on the map and then input some metadata about the bakery including a name and brief description. The map object then needs to be classified so that it belongs to a particular concept. To do this the user is required to place a submitted instance entity into a folder which translates to a entity model for the owner. The user is supplied with a set of basic folders to help maintain some control on the scope of the submitted entities. These base folders are defined by the seed ontology that is pre-defined by the owner.

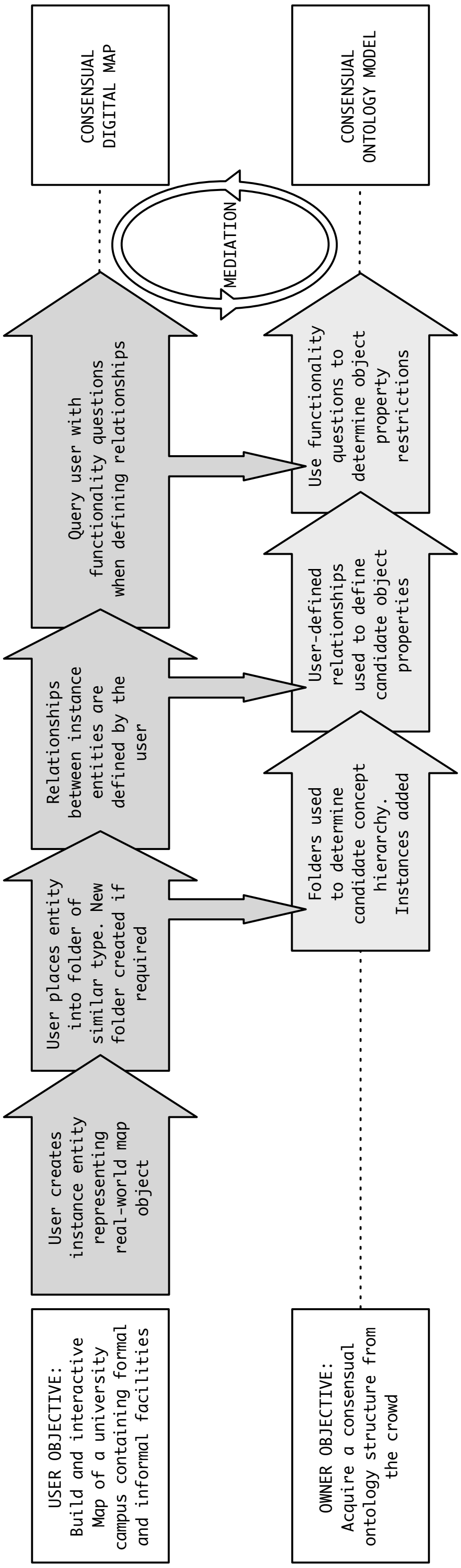


FIGURE 6.1: Parallel motivations, how building a digital map resource can be leveraged to acquire an ontology structure

| Participant | University Link |
|---|---|
| A1 | Current undergraduate student |
| A2 | Former undergraduate student |
| A3 | Current undergraduate student |
| A4 | Lives locally |
| B1 | Current Ph.D. student |
| B2 | Current undergraduate student |
| B3 | Former undergraduate student |
| B4 | Works locally |
| C1 | Current Ph.D. student |
| C2 | Current undergraduate student |
| C3 | Current undergraduate student |
| C4 | Current masters student |
| A5 | Current Ph.D. student |

Table 6.1: Summary of Participants

## 6.4 Experimental Conditions

### Participation

While there is no pre-requisite knowledge engineering competency, it is important that participants are both motivated and capable of providing the information needed to build the collaborative campus map. The participants were all students, former students of the University of Liverpool or people who have knowledge of the university campus. A community crowd is being stimulated based on the knowledge of a locality. No knowledge engineering skills can be assumed of this crowd and, unlike Experiment 1 or 2, the knowledge acquisition objective is not made explicit.

An email appeal for volunteers was made and circulated across various lists. The only details required from participants was their email and a brief description as to what relationship the participant had to the University of Liverpool. Respondents were then given a username to login with.

In total 12 participants volunteered information for the initial part of this experiment, the details of which are summarised in table 6.1. A further participant was obtained for the second part of this experiment to make it analogous with the earlier experiment. The volunteers were divided into 3 groups each using one of the 3 seed ontologies. In the analysis of the results a separate analysis of the effectiveness of each seed ontology will be provided along with analysis of the effectiveness of the experiment as a whole.

### Seed Ontologies

The concepts provided in the seed ontologies are summarised in figure 6.2. The *Shallow Structure* seed ontology describes only the most abstract domain concepts. The *Intermediate Structure* provides an additional level of descriptions with the *Deep Structure* seed ontology representing the highest degree of description.

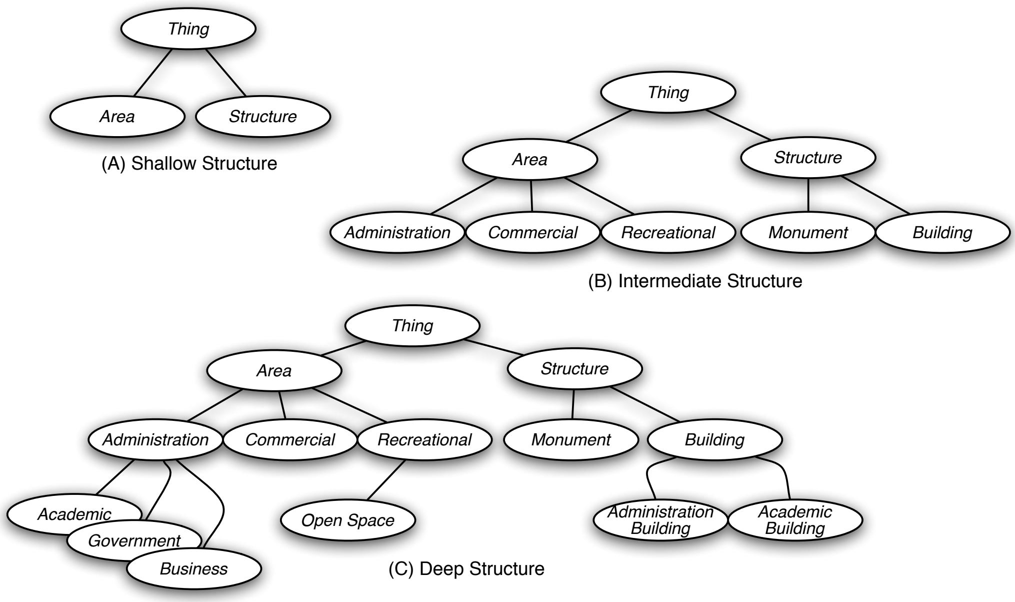


FIGURE 6.2: Seed Ontology Structures

### 6.4.1 User Interface

For this experiment, using a general collaborative ontology development tool like *Moki* was unsuitable. While our computer science students were, mostly, able to understand and use *MoKi* to build a knowledge model, we could not assume this knowledge of the campus map crowd. Instead a map-based interface is used for our participants to engage in. A user interface was designed and built that could be accessed by our participants via a web-browser. This section will provide a description of the user interface including its design and technical information.

The primary objective of the user interface design was to facilitate not only general participation, but also to encourage our participants to use the tool in the best possible way in order to generate useful ontology structures. The map also provided a useful visual aid for the participants who were able to use it to place objects in relation to the local geographic environment.

### 6.4.2 User Experience

The user is first presented with a introduction window that included a link to a full set of instructions. The participant then had to login using a username provided to identify them and that linked them to a particular ontology file. After login, the participant was given the option of entering three modes. The default browse mode allowed the participants to view the map and look at any entities that they had already created. The input mode allowed participants to create instance entities. Finally, the edit mode could be used to edit any existing object. The user experience could be described as instance-orientated, as their primary focus was on manipulating and describing the

map objects which are viewed as instances of concepts when placed into the ontology structures.
The full instructions offered to participants can be found in appendix C.

### Creating Instance Entities

Having entered the input mode, Figure 6.3 shows the initial stage of the instance entity creation process. The participant uses the cursor to map out the object shape, when the shape is finished the participant uses the 'Complete Shape' button to move on to the next stage. The participant can then provide a name for the instance entity along with any additional comments.


FIGURE 6.3: Screenshot: Entity Creation

Once an object is placed on the map the participant is prompted to move that object into a relevant folder, a process analogous to organising files on a computer (Figure 6.4). Each folder corresponded to a concept from the seed ontology provided or concepts already created by the participant themselves. The participant is then encouraged to create folders where appropriate, thus unwittingly extending the ontology concept structure. A participant is able to delete or move folders that they had created when in edit mode, but could not modify the pre-defined folders in any way. The folder metaphor is used to encourage contributors to build a hierarchy as required for ontology building.

### Defining Relationships between Instance Entities

There is also a mode for defining object property relationships. This allows a participant to link objects together by selecting two map objects and either defining a new object property, or choosing an existing property from a drop-down menu. Unfortunately, only a few participants contributed with object properties, meaning that a meaningful analysis cannot take place.

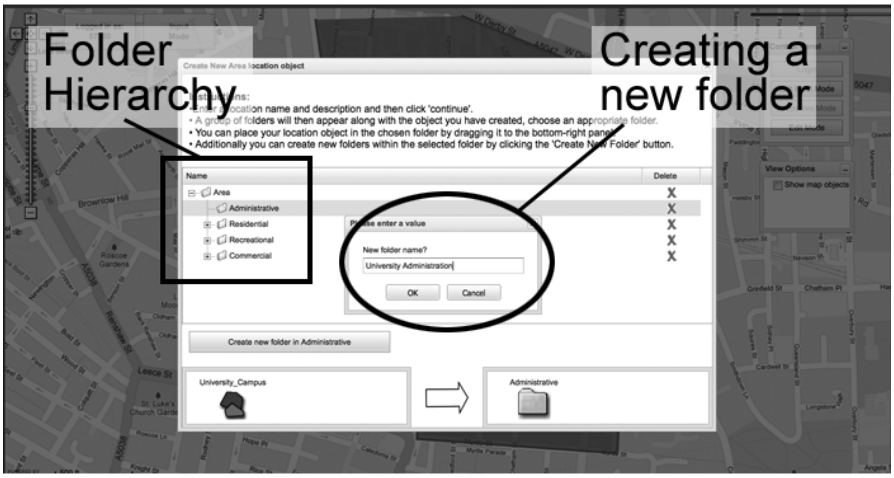


FIGURE 6.4: Screenshot: Classifying

**Technology Overview**

Figure 6.5 illustrates the various technologies which were incorporated into the pilot implementation. Here can be seen the linear process through which information is passed into the system via a web-browser. In this case the interface uses tools provided by the mapping software (*GoogleMaps*) to enable input. The client-server communication is handled using Remote Procedure Calls (RPC) which passes serialized strings from our server to the client machine. The server-side code incorporates Apache Jena[1], a Java-based framework which can be used to develop semantic web applications and which is capable of building domain models and outputting them as OWL files. The system was built using *Google Web Toolkit* (GWT) which has provided a framework for communicating between our Java code and our clientside code.

A working version of this interface is no longer live. Please contact the author for a copy of this software which can be hosted on a Apache Tomcat Server.

[1]`http://http://jena.apache.org/`

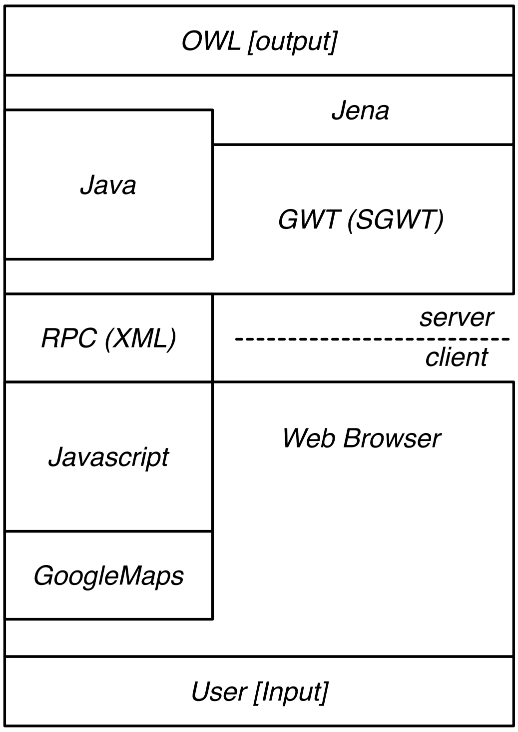


Figure 6.5: Technology Outline

## 6.5 Results: Part 1

### 6.5.1 Concept Elicitation from the Crowd

Table 6.2 provides details on the number of entities that comprise each of the submitted ontologies before mediation. The results are broken down into each of the seed ontology sets.

| | **Total Entities** | | | | **New Entities as % of Ontology** | | | |
|---|---|---|---|---|---|---|---|---|
| *Ontology* | **1** | **2** | **3** | **4** | **1** | **2** | **3** | **4** |
| Shallow Structure | 33 | 37 | 48 | 30 | 90.9% | 91.9% | 93.7% | 90.0% |
| Intermediate Structure | 42 | 45 | 39 | 26 | 73.8% | 75.5% | 71.7% | 57.7% |
| Deep Structure | 42 | 43 | 40 | 42 | 42.8% | 44.2% | 40.0% | 42.8% |

Table 6.2: Summary of Submitted Entities

### 6.5.2 Convergence towards consensus achieved through Automated Mediation

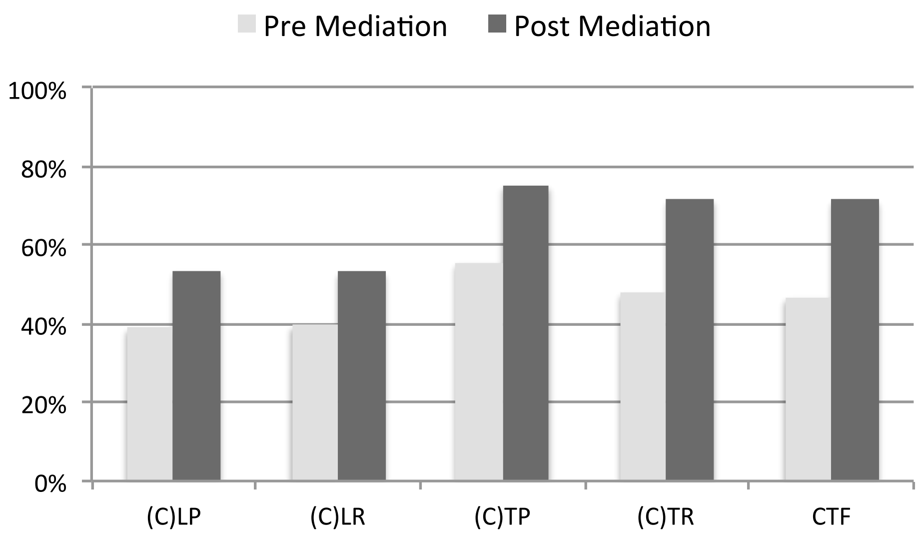


Figure 6.6: Mediation of Shallow Set

| | **Shallow** | | **Intermediate** | | **Deep** | |
|---|---|---|---|---|---|---|
| | **Pre** | **Post** | **Pre** | **Post** | **Pre** | **Post** |
| CTF | 46.77% | 71.54% | 62.54% | 70.38% | 90.56% | 91.81% |

Table 6.3: Effect of Automated Mediation on CTF

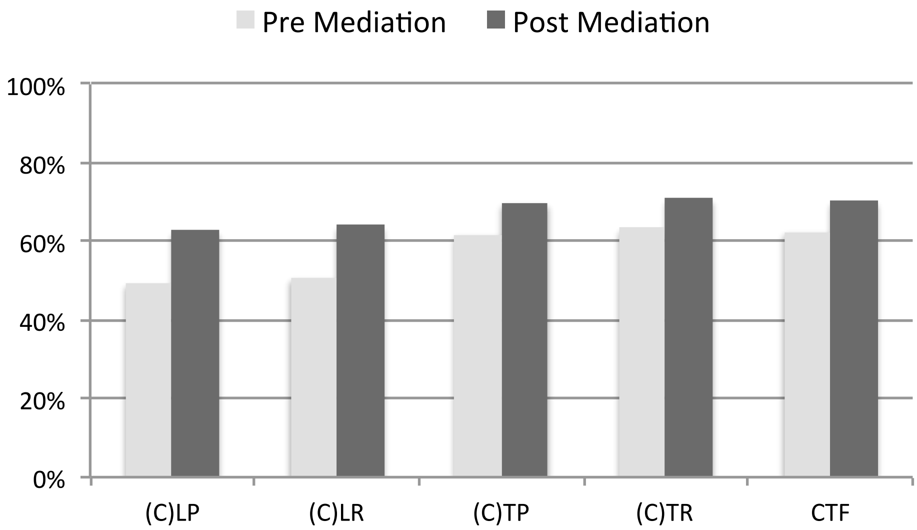


Figure 6.7: Mediation of Intermediate Set

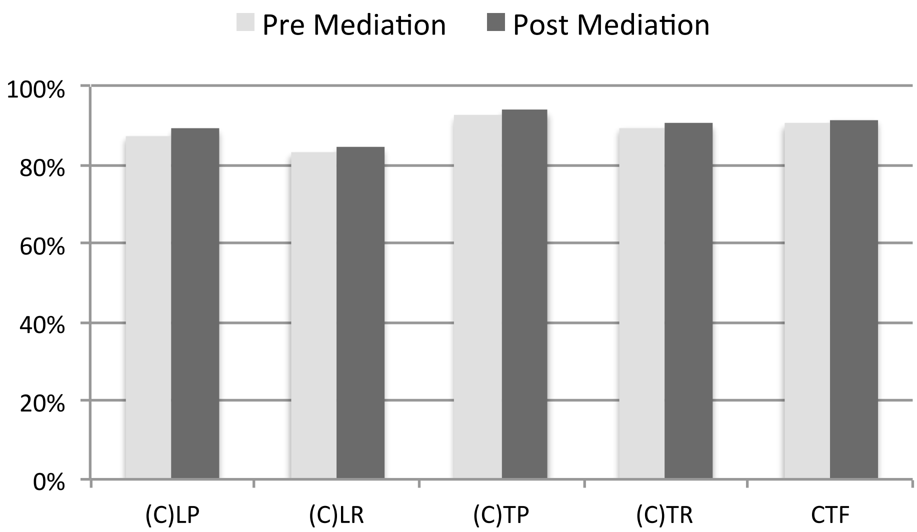


Figure 6.8: Mediation of Deep Set

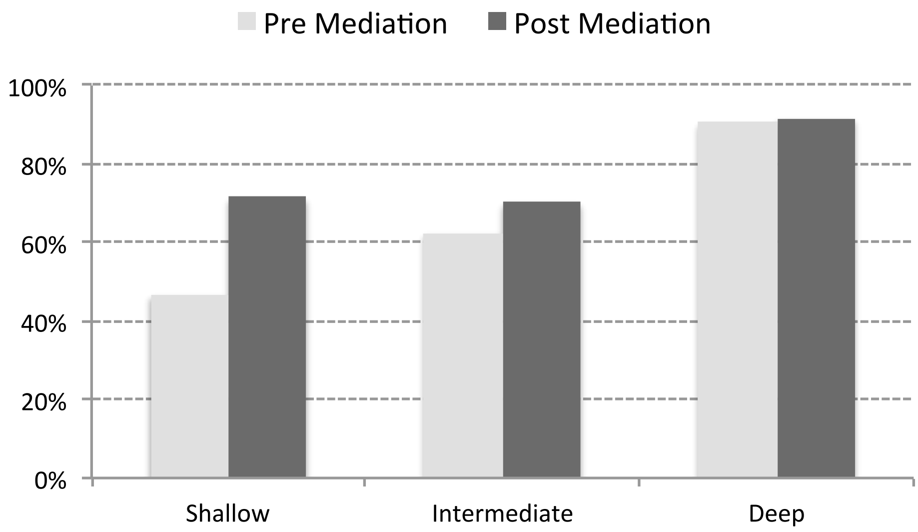


FIGURE 6.9: Mediation Effect on CTF

## 6.6 Choice of Ontology Set for Semi-Automated Mediation

Figures 6.6, 6.7 and 6.8 show the convergence towards a greater consensus across the ontology sets achieved through automated mediation. As expected the shallower the seed ontology structure the greater the performance of the mediation. It can be seen that mediation has a positive impact of the lexical precision and recall, indicating that there is an a greater agreement regarding the entities that should be contained in the ontology after automated mediation. This improvement is fairly uniform across all three ontology sets.

The most interesting result is the improvement in consensus of the shallow seed ontology set as measured by CTF. As figure 6.9 and table 6.3 shows, the CTF measure sees an improvement of 31.83% for the shallow set, as compared to the intermediate set which shows an improvement of just 7.84%. While we might expect to see the greatest improvement in the shallow set because of the greater number of mediation cases it provides, the improvement is such that the post mediation of the shallow set (71.53%) shows greater consensus after mediation than in the intermediate set (70.38%). This suggests that to move forward with this approach we should use the shallow seed ontology set. This is good news for reducing the requirement for input from knowledge engineers as they only have to specify abstract seed ontologies.

## 6.7 Results: Part II

This section contains the full tables of results obtained, followed by an overview in which any trends are illustrated. The measure used to calculate the CTF is the TF, which excludes the influence of the lexical recall. The LR, LP and the dependant TF1 metric are still recorded as they might give some insight into the effect of the mediation process.

### 6.7.1 Results: Pre Mediation

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 2 | 33.33% | 44.44% | 45.83% | 73.81% | 56.55% | 49.77% |
| 3 | 22.22% | 44.44% | 33.33% | 73.81% | 45.93% | 45.17% |
| 4 | 50.00% | 33.33% | 60.00% | 71.43% | 65.22% | 44.12% |
| 5 | 33.33% | 55.56% | 51.29% | 88.57% | 64.96% | 59.89% |
| | CTF | 58.16% | | | | |

TABLE 6.4: Pre-Mediation: $\mathcal{O}^1$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 2 | 33.33% | 44.44% | 45.83% | 73.81% | 56.55% | 49.77% |
| 3 | 22.22% | 44.44% | 33.33% | 73.81% | 45.93% | 45.17% |
| 4 | 50.00% | 33.33% | 60.00% | 71.43% | 65.22% | 44.12% |
| 5 | 33.33% | 55.56% | 51.29% | 88.57% | 64.96% | 59.89% |
| | CTF | 58.16% | | | | |

Table 6.5: Pre-Mediation: $\mathcal{O}^2$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 44.44% | 22.22% | 73.81% | 33.33% | 45.93% | 29.95% |
| 2 | 50.00% | 33.33% | 58.33% | 48.15% | 52.75% | 40.85% |
| 4 | 50.00% | 16.67% | 60.00% | 18.52% | 28.30% | 20.98% |
| 5 | 40.00% | 33.33% | 57.34% | 45.37% | 50.66% | 40.21% |
| | CTF | 44.41% | | | | |

Table 6.6: Pre-Mediation: $\mathcal{O}^3$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 33.33% | 50.00% | 71.43% | 60.00% | 65.22% | 56.60% |
| 2 | 33.33% | 66.67% | 45.83% | 66.67% | 54.32% | 59.86% |
| 3 | 16.67% | 50.00% | 18.52% | 60.00% | 28.30% | 36.14% |
| 5 | 33.33% | 83.33% | 64.29% | 92.00% | 75.69% | 79.33% |
| | CTF | 55.88% | | | | |

Table 6.7: Pre-Mediation: $\mathcal{O}^4$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 55.56% | 33.33% | 88.57% | 51.29% | 64.96% | 44.06% |
| 2 | 58.33% | 46.67% | 61.90% | 60.46% | 61.17% | 52.94% |
| 3 | 33.33% | 40.00% | 45.37% | 57.34% | 50.66% | 44.70% |
| 4 | 83.33% | 33.33% | 92.00% | 64.29% | 75.69% | 46.28% |
| | CTF | 63.12% | | | | |

Table 6.8: Pre-Mediation: $\mathcal{O}^5$

### 6.7.2 Results: Automated Mediation

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 2 | 50.00% | 80.00% | 60.00% | 83.33% | 69.77% | 74.53% |
| 3 | 44.44% | 80.00% | 58.33% | 83.33% | 68.63% | 73.88% |
| 4 | 60.00% | 60.00% | 66.67% | 77.78% | 71.79% | 65.37% |
| 5 | 45.45% | 100.00% | 61.00% | 100.00% | 75.78% | 86.22% |
| | CTF | 71.49% | | | | |

Table 6.9: Automated Mediation: $\mathcal{O}^1$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 80.00% | 50.00% | 83.33% | 60.00% | 69.77% | 58.25% |
| 3 | 66.67% | 75.00% | 66.67% | 68.89% | 67.76% | 71.20% |
| 4 | 80.00% | 50.00% | 87.50% | 80.00% | 83.58% | 62.57% |
| 5 | 63.64% | 87.50% | 74.29% | 82.86% | 78.34% | 82.67% |
| | CTF | 74.86% | | | | |

Table 6.10: Automated Mediation: $\mathcal{O}^2$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 80.00% | 44.44% | 83.33% | 58.33% | 68.63% | 53.95% |
| 2 | 75.00% | 66.67% | 68.89% | 66.67% | 67.76% | 67.21% |
| 4 | 60.00% | 33.33% | 66.67% | 50.00% | 57.14% | 42.11% |
| 5 | 54.55% | 66.67% | 63.33% | 63.89% | 63.61% | 65.10% |
| | CTF | 64.28% | | | | |

Table 6.11: Automated Mediation: $\mathcal{O}^3$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 60.00% | 60.00% | 77.78% | 66.67% | 71.79% | 65.37% |
| 2 | 50.00% | 80.00% | 80.00% | 87.50% | 83.58% | 81.75% |
| 3 | 33.33% | 60.00% | 50.00% | 66.67% | 57.14% | 58.54% |
| 5 | 45.45% | 100.00% | 75.33% | 100.00% | 85.93% | 92.43% |
| | CTF | 74.61% | | | | |

Table 6.12: Automated Mediation: $\mathcal{O}^4$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 100.00% | 45.45% | 100.00% | 61.00% | 75.78% | 56.82% |
| 2 | 87.50% | 63.64% | 82.86% | 74.29% | 78.34% | 70.23% |
| 3 | 66.67% | 54.55% | 63.89% | 63.33% | 63.61% | 58.73% |
| 4 | 100.00% | 45.45% | 100.00% | 75.33% | 85.93% | 59.46% |
| | CTF | 75.91% | | | | |

Table 6.13: Automated Mediation: $\mathcal{O}^5$

### 6.7.3 Results: Semi-Automated Mediation

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 2 | 86.67% | 81.25% | 92.67% | 94.38% | 93.52% | 86.95% |
| 3 | 76.47% | 81.25% | 90.17% | 93.42% | 91.77% | 86.19% |
| 4 | 86.67% | 81.25% | 97.86% | 96.95% | 97.40% | 88.60% |
| 5 | 80.00% | 75.00% | 94.64% | 93.95% | 94.29% | 83.55% |
| | CTF | 94.25% | | | | |

Table 6.14: Semi-Automated Mediation: $\mathcal{O}^1$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 81.25% | 86.67% | 94.38% | 92.67% | 93.52% | 89.96% |
| 3 | 82.35% | 93.33% | 87.10% | 88.69% | 87.89% | 90.53% |
| 4 | 80.00% | 80.00% | 96.30% | 93.65% | 94.96% | 86.84% |
| 5 | 80.00% | 80.00% | 91.17% | 86.31% | 88.67% | 84.11% |
| | CTF | 91.26% | | | | |

Table 6.15: Semi-Automated Mediation: $\mathcal{O}^2$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 81.25% | 76.47% | 93.42% | 90.17% | 91.77% | 83.42% |
| 2 | 93.33% | 82.35% | 88.69% | 87.10% | 87.89% | 85.03% |
| 4 | 80.00% | 70.59% | 95.72% | 87.04% | 91.17% | 79.57% |
| 5 | 80.00% | 70.59% | 89.63% | 79.63% | 84.34% | 76.85% |
| | CTF | 88.79% | | | | |

Table 6.16: Semi-Automated Mediation: $\mathcal{O}^3$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 81.25% | 86.67% | 96.95% | 97.86% | 97.40% | 91.72% |
| 2 | 80.00% | 80.00% | 93.65% | 96.30% | 94.96% | 86.84% |
| 3 | 70.59% | 80.00% | 87.04% | 95.72% | 91.17% | 85.22% |
| 5 | 73.33% | 73.33% | 95.29% | 93.18% | 94.23% | 82.48% |
| | CTF | 94.44% | | | | |

Table 6.17: Semi-Automated Mediation: $\mathcal{O}^4$

| $\mathcal{O}$ | LP | LR | TP | TR | TF | TF1 |
|---|---|---|---|---|---|---|
| 1 | 75.00% | 80.00% | 93.95% | 94.64% | 94.29% | 86.56% |
| 2 | 80.00% | 80.00% | 86.31% | 91.17% | 88.67% | 84.11% |
| 3 | 70.59% | 80.00% | 79.63% | 89.63% | 84.34% | 82.11% |
| 4 | 73.33% | 73.33% | 93.18% | 95.29% | 94.23% | 82.48% |
| | CTF | 90.38% | | | | |

Table 6.18: Semi Automated Mediation: $\mathcal{O}^5$

Semi-automated mediation is less successful in improving the consensus at a lexical level than in Experiment 1 or 2. The comparison between the effect of mediation on lexical precision and recall in the two experiments is illustrated in figure 6.10.

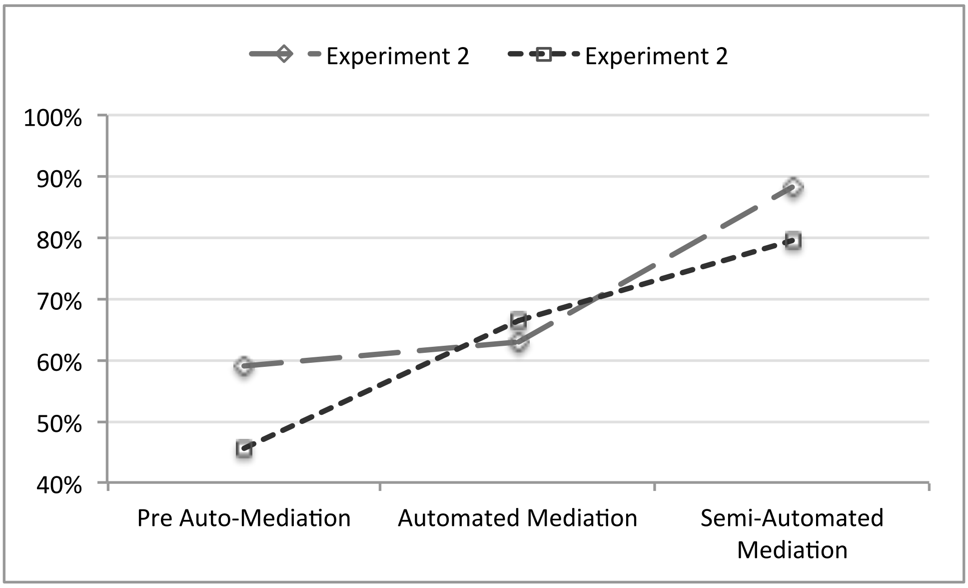


Figure 6.10: Effect of Mediation on Lexical Precision and Recall

### 6.7.4 Automated Mediation of Shallow Ontology Set

For the reasons listed above, the shallow ontology set was selected and the automated mediation was reapplied. The mediation had to be reapplied because the ontology set was augmenting with an additional ontology to conform with the experiment in Chapter 4 in which five ontologies were used. The automated mediation of this ontology set is summarised in table 6.19. In this table, the concept and its presence in each ontology are listed in the first six columns. The next column indicates if there was model conflict (Y or N). If there was a conflict, the success of the mediation in reaching a decision as to the correct model is indicated in brackets. The next column indicates the confidence. This is the number of models for the given concept that are identical across the ontology set. If the model has been altered during the mediation process, then the pre-mediation confidence is recorded here. The next three columns indicate the outcome of the mediation: *mA* indicates that a Minority Adopt was performed, *MA* indicates that a Majority Adopt was performed and *P* indicates that the concept was pruned from the list.

The questions produced in order to carry out the ballots on Reinstate and Majority Adopt are listed in tables 6.20 and 6.21 respectively. The participants were all emailed this list and asked to indicate their support for the concepts.

| Concept | 1 | 2 | 3 | 4 | 5 | Conflict? | Conf. | mA | MA | P |
|---|---|---|---|---|---|---|---|---|---|---|
| academic | | | • | | • | $\times$ | 0.4 | • | | |
| academic_building | | | • | | • | $\times$ | 0.4 | • | | |
| administration | | | • | | | | | | | • |
| administration_building | | | • | | | | | | | • |
| Area* | • | • | • | • | • | | | | • | |
| borough | | | • | | | | | | | • |
| building | • | • | • | • | • | | | | • | |
| bus_stop | | | | | • | | | | | • |
| campus | | • | • | | • | $\checkmark(\times)$ | | | | |
| city | | • | • | | | $\checkmark(\times)$ | | | | |
| commercial_building | | | • | | | | | | | • |
| department | | • | • | | • | $\checkmark(\times)$ | | | | |
| department_building | | | • | | | | | | | • |
| drinking | • | | | | | | | | | • |
| eating | • | | | | | | | | | • |
| enterprize_zone | | | • | | | | | | | • |
| faculty_building | | • | | | | | | | | • |
| food | | | • | | | | | | | • |
| library | | | | • | | | | | | • |
| museum | | | | • | • | $\times$ | 0.4 | • | | |
| office | | | | | • | | | | | • |
| open_space | | • | | | | | | | | • |
| other | | • | | | | | | | | • |
| park | | | • | | | | | | | • |
| pub | | | | | • | | | | | • |
| recreation | • | | | | | | | | | • |
| region | | • | | | • | $\times$ | 0.4 | • | | |
| retail | | | • | | | | | | | • |
| school | | | | | • | | | | | • |
| shop | | • | | • | • | $\checkmark(\checkmark)$ | 0.4 | • | | |
| shopping | • | | | | | | | | | • |
| Structure* | • | • | • | • | • | $\times$ | | | • | |
| support_building | | | • | | | | | | | • |
| university | • | | | | • | $\times$ | 0.4 | • | | |
| university_building | | • | | | | | | | | • |

TABLE 6.19: Summary of Automated Mediation

| Reinstate Questions | Conf. | mA | MA | Rmv |
|---|---|---|---|---|
| Is administration_building, which is a type of building, a valid concept? | 0.8 | | • | |
| Is administration, which is a type of area, a valid concept? | 0.8 | | • | |
| Is borough, which is a type of administration, a valid concept? | 0.4 | | | • |
| Is bus_stop, which is a type of structure, a valid concept? | 0.6 | • | | |
| Is commercial_building, which is a type of building, a valid concept? | 0.8 | | • | |
| Is department_building, which is a type of academic_building, a valid concept? | 1 | | • | |
| Is drinking, which is a type of shopping, a valid concept? | 0.2 | | | • |
| Is eating, which is a type of shopping, a valid concept? | 0.4 | | | • |
| Is enterprise_zone which is a type of administration, a valid concept? | 0.4 | | | • |
| Is faculty building, which is a type of university_building, a valid concpet? | n/a | | | • |
| Is food, which is a type of retail, a valid concept? | n/a | | | • |
| Is library, which is a type of building, a valid concept? | 1 | | • | |
| Is office, which is a type of building, a valid concept? | 0.8 | | • | |
| Is open_space, which is a type of area, a valid concept? | 0.4 | | | • |
| Is other, which is a type of structure, a valid concept? | 0.4 | | | • |
| Is park, which is a type of administration, a valid concept? | 0.2 | | | • |
| Is pub, which is a type of building, a valid concept? | 1 | | • | |
| Is recreation, which is a type of area, a valid concept? | 0.8 | | • | |
| Is retail, which is a type of commercial building, a valid concept? | 0.4 | | | • |
| Is school, which is a type of university, a valid concept? | 0.2 | | | • |
| Is shopping, which is a type of area, a valid concept? | 0.6 | • | | |
| Is support_building, which is a type of academic_building, a valid concept? | 0.4 | | | • |
| Is university_building, which is a type of building, a valid concept? | 0.4 | | | • |

Table 6.20: Semi-Automated Mediation: Reinstate Questions

| Majority Adopt Question | Confidence | mA | MA |
|---|---|---|---|
| Is academic, which is a area a valid concept? | 0.6 | * | |
| Is academic_building, which is a building, a valid concept? | 1 | | * |
| Is museum, which is a type of building, a valid concept? | 1 | | * |
| Is region, which is a type of area, a valid concept? | 0.8 | | * |
| Is shop, which is a type of building, a valid concept? | 1 | | * |
| Is university, which is a type of area, a valid concept? | 0.4 | * | |

Table 6.21: Summary of Majority Adoption Questions

### 6.7.5 Convergence towards consensus achieved through Mediation

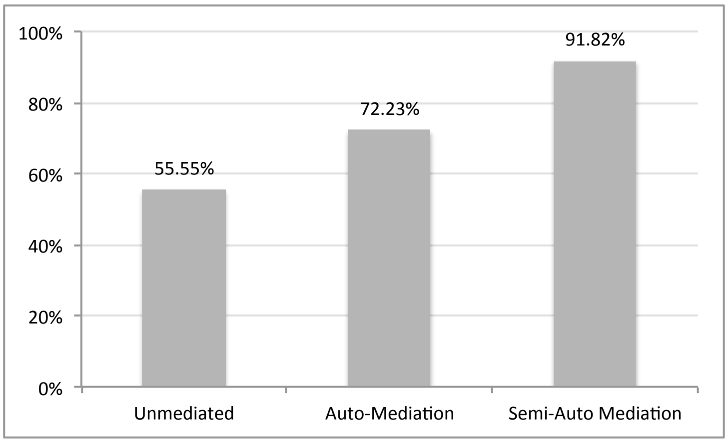


FIGURE 6.11: Average convergence towards consensus (CTF)

Figure 6.11 illustrates the average convergence of the ontology set as measured by CTF. Overall, a similar average CTF is achieved to that obtained in the first experiment. In the first experiment an average CTF of 95.88% was obtained, while in this experiment it was an average of 91.82%. The convergence achieved from the task-unaware crowd is particularly impressive considering that the initial level of consensus was 55.5% as compared with the 78.92% CTF achieved at the comparable stage of the task-aware experiment. That means that a convergence of 16.96% was achieved in the first experiment, while a convergence of 36.32% was achieved in the second experiment.

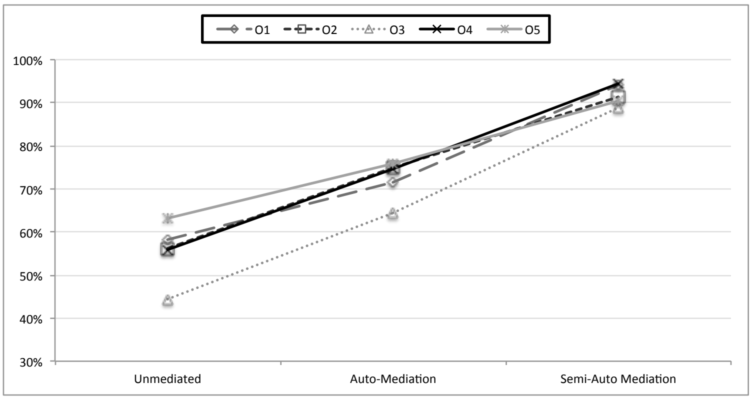


FIGURE 6.12: Convergence towards consensus by Ontology (CTF)

Figure 6.12 illustrates the convergence achieved by ontology. Compared to the equivalent result from the first experiment the automated mediation process appears to have a greater effect. This is most likely because the proportion of automated-mediation outcomes resulting in the temporary removal of concepts was higher in the second experiment.

## 6.8 Summary

The results of the experiments described in this chapter indicate that the automated mediation process is effective at improving consensus between ontologies, even if the seed ontology given to the participants contains only an abstract description. The second part of the experiment shows that a similar level of convergence can be obtained using the processes outlined from task-unaware crowds as to the results achieved from task-aware crowds. These results show great potential for the adoption of this approach. However, the usefulness of the acquired knowledge contained in the mediated models needs to be assessed. If, during the mediation process too much valid knowledge is being discarded, or if irrelevant or contradictory knowledge is being universally adopted, then the effectiveness of this approach may be compromised.

# Chapter 7

# Discussion

## 7.1 Quality of Knowledge Acquisition

Having elicited and processed knowledge from the crowd, a set of ontology models have been acquired. All the concepts and object properties acquired will now be categorised into four distinct types:

- Majority adopted: fully consensual knowledge
- Minority adopted: consensus reached across two or more ontologies, but not universally adopted
- Conflicted: existence of entity supported, but model conflict unresolved
- Rejected: entity rejected as invalid

In minority adopted and conflicted cases further adjudication is needed; in the case of minority adoption a knowledge engineer should consider if the concept should be included in the final knowledge representation; in conflicted cases the knowledge engineer should consider whether the concept should be included and what the concept model should be.

### 7.1.1 Quality of Acquired Knowledge from Experiment 1

This section of the discussion will look at these categories of entities and will try to identify how effectively the acquired entities have been classified. Table 7.1 summarises the outcomes of the mediation process obtained from the first experiment. Figure 7.2 illustrates the consensual hierarchy. Those concepts marked (*) are those which only have a Minority Adoption. Those marked twice (**) are the concepts where no resolution was found to the model conflict. In the case of unresolved concepts, multiple versions of that concept are included to show any different conceptualisations.

| | Majority Adopt | Minority Adopt | In Conflict | Rejected |
|---|---|---|---|---|
| | Artist | Best_single | Composer | Chart |
| | Award | Female | | Click_rate |
| | Gender* | Male | | Continent |
| | Genre* | Country | | Download_times |
| | Group* | Price | | False |
| | Person* | | | Label |
| | Song* | | | Member |
| | Store | | | Name |
| | Album | | | No._1 |
| | Best_Group | | | No1 |
| | Best_record | | | Offer |
| | Best_album | | | Record_label |
| | Covers | | | Release_date |
| | E_p | | | Sales |
| | Recording_artist | | | Time |
| | Singer | | | Time_period |
| | Writer | | | True |
| | Month | | | |
| | Year | | | |
| | Week | | | |
| | Producer | | | |
| **Total** | 21 | 5 | 1 | 17 |

Table 7.1: Experiment 1: Mediation Outcomes
(Concepts marked with * are contained in the seed ontology)


Figure 7.1: Experiment 1: Consensual Model Concept Hierarchy

### Majority Adoption

Generally speaking, the concepts and models that have been majority adopted are valid. There are no instances of concepts whose presence in the model can be disputed, however there are some concepts that could be better described. The concepts *Year*, *Month* and *Week* are all modelled as being subclasses of *thing*. However, in two of the pre mediation ontologies these concepts are more usefully conceptualised as subclasses of *Time_Period* and *Time*. Had the equivalence of *Time_Period* and *Time* been established so that it became a Minority Adoption case, then a more useful model would have been recorded or even majority adopted at a later stage. Similarly, the subclasses of *Person* could be better modelled. *Recording_Artist* and *Singer* would be better placed as subclasses of *Artist*.

The distinction between *Record* and *Album* is also problematic, as it is not clear if they are equivalent. The fact that at least one group felt that *Record* was not descriptive enough to encompass the concept of Album shows that *Record* is a problematic concept. However, this is possibly a failure of the experimental conditions which determined *Record* as a fundamental concept.

### Minority Adoption and Conflicted Concept Models

For the Minority Adoption and conflicted cases it is less important that the conceptualisation is correct. It is more important that these cases represent knowledge which might be more difficult to conceptualise due to the differing perspectives of the crowd. In other words, these should be cases the knowledge engineer should legitimately have to adjudicate over. The concepts *Male* and *Female*, while making sense given the question presented ("Is Male, which is a type of Gender, a valid concept?"), does not make much sense in an ontology where they would be instances of *Gender*. This is a case where a knowledge engineer would be able to make a quick decision, removing the entity from the model. In the case of *Country* and *Price* the validity of their inclusion as concepts is less clear, and therefore adjudication by the knowledge engineer is a valid outcome. Given that *Best_album*, *Best_artist* and *Best_group* were majority adopted, the fact that *Best_single* is not majority adopted indicates some inconsistency in the adoption mechanism. Given the nature of crowd input, this is perhaps to be expected. However, this could be problematic to the knowledge engineer who would need to have knowledge of the majority adoption cases to obtain the context in which to make a validity decision. Overall, there appears to be no major concerns regarding the Minority Adoption cases as they can all me manifested as legitimate validity decisions to the knowledge engineer.

### Removed Concepts

By removing concepts and their dependent object properties, there is a danger that the expressiveness of the acquired consensual knowledge model may be compromised. The best way to test this is to evaluate whether the competency questions used to build the individual models can still be answered given the consensual model. Table 7.2 and 7.3 indicate which of the competency questions submitted to the participants can be considered to be answered by the produced models. The average percentage of competency questions answered by each individual ontology is also indicated in the two tables. Where some aspect of the competency question has been answered, but where it is difficult to see how the question as a whole can be answered a 0.5 score is given. For example, if the question is whether a female award winner can be identified, if a property to determine gender is present, but not property to indicate award-winning, then a score of 0.5 is awarded. Otherwise if the question can be considered fully answered it is awarded a score 1 and if it cannot be answered the model will be awarded a zero. Acknowledging that this analysis is fairly subjective, requiring a manual translation of the intent of the participants in specifying concepts and object properties, it may still provide some understanding how the mediation process has effected the expressiveness of the acquired model. This analysis does rely on object properties which were not mediated, therefore this analysis will should only be seen as being indicative of the concept coverage of the competency questions.

It should be noted that some questions are answered, correctly, through the use of data type properties. Data type properties link concepts to a data type such as an integer, string or boolean. While the participants were not instructed to use data type properties, they were not prevented from doing so.

| Before Mediation | $\mathcal{O}3$ | $\mathcal{O}\mathcal{O}4$ | $\mathcal{O}5$ | $\mathcal{O}6$ | $\mathcal{O}8$ |
|---|---|---|---|---|---|
| What was the album published in [year] by [artist]? | 0 | 0.5 | 0 | 0 | 0 |
| Which female artist won an award this year? | 1 | 0.5 | 0 | 0 | 1 |
| Does the store have a search function? | 1 | 1 | 1 | 0 | 1 |
| Does the store have a download function? | 1 | 1 | 1 | 0 | 0 |
| Which male artist sang [song]? | 1 | 0 | 1 | 0 | 1 |
| What albums is [artist] best known for? | 1 | 0 | 0 | 1 | 1 |
| How many members are in [group]? | 1 | 0 | 1 | 0 | 0 |
| What was the No1 album of [year]? | 1 | 0.5 | 0 | 0 | 0 |
| This group belongs to which [genre]? | 0 | 0 | 0 | 1 | 0 |
| This song belongs to which [Record]? | 0 | 0 | 1 | 0 | 0 |
| This person belongs to which [Group]? | 0 | 1 | 1 | 0 | 0 |
| This record belongs to which [artist]? | 0 | 0 | 1 | 1 | 0 |
| Which genre is [artist] best known for? | 1 | 0 | 0 | 0.5 | 0 |
| Which artist won an award last year? | 1 | 0 | 1 | 1 | 0 |
| What song charted last week? | 0 | 0.5 | 0 | 0 | 0 |
| What song is free this week? | 1 | 0.5 | 1 | 1 | 0 |
| Average | 62.50% | 34.38% | 56.25% | 34.38% | 25.00% |

Table 7.2: Competency Questions answered by submitted models (Pre-Mediation)

From the analysis of tables 7.2 and 7.3 we can see that the number of competency questions that can be answered is reduced from an average of 42.50% prior to mediation, to 31.88% after mediation. This is a significant reduction, but not surprising considering that many object properties are removed in the final stage of mediation. A fairer comparison would be to divide the average percentage of competency questions answered by the number of object properties present in the ontology set. This results in the

| After Mediation | $\mathcal{O}3$ | $\mathcal{O}4$ | $\mathcal{O}5$ | $\mathcal{O}6$ | $\mathcal{O}8$ |
|---|---|---|---|---|---|
| What was the album published in [year] by [artist]? | 0 | 0 | 0 | 0 | 0 |
| Which female artist won an award this year? | 0.5 | 1 | 0.5 | 0.5 | 0.5 |
| Does the store have a search function? | 0 | 1 | 1 | 0 | 1 |
| Does the store have a download function? | 0 | 1 | 1 | 0 | 0 |
| Which male artist sang [song]? | 1 | 0.5 | 0.5 | 1 | 1 |
| What albums is [artist] best known for? | 1 | 0 | 0 | 1 | 0 |
| How many members are in [group]? | 0 | 1 | 1 | 0 | 0 |
| What was the No1 album of [year]? | 1 | 0 | 0 | 0 | 0 |
| This group belongs to which [genre]? | 1 | 0 | 0 | 0 | 0 |
| This song belongs to which [Record]? | 0 | 0 | 0 | 1 | 0 |
| This person belongs to which [Group]? | 0 | 1 | 1 | 0 | 0 |
| This record belongs to which [artist]? | 0 | 0 | 1 | 0 | 0 |
| Which genre is [artist] best known for? | 0 | 0 | 0 | 1 | 0 |
| Which artist won an award last year? | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| What song charted last week? | 0 | 0 | 0 | 0 | 0 |
| What song is free this week? | 0 | 0 | 0 | 0 | 0 |
| | 31.25% | 37.50% | 40.63% | 31.25% | 18.75% |

Table 7.3: Competency Questions answered by submitted models (Post-Mediation)

pre-mediation set answering 0.80 competency questions per object property specified, while the mediated set answer 0.86. This shows that the mediated object properties are marginally more effective at answering the competency questions.

### 7.1.2 Quality of Acquired Knowledge from Experiment 2

| | **MA** | **mA** | **In Conflict** | **Rejected** |
|---|---|---|---|---|
| | artist | award | album | african |
| | blues | b | ep | animal |
| | composer | band | lp | asian |
| | electronic | classical | single | bahasa indonesia |
| | gender | drummer | ska | beatles |
| | genre | duo | song | blue |
| | country | bassist | | comedy |
| | group | folk | | company |
| | jazz | guitarist | | country music |
| | person | hip-hop | | country record |
| | pop | instrument | | country singer |
| | popular music | language | | era |
| | producer | lyricist | | european |
| | record | male | | female |
| | rock / rock music | metal | | ha |
| | singer | musician | | instrumentalist |
| | writer / author | trio | | john lennon |
| | year | | | lyric |
| | | | | melody |
| | | | | member |
| | | | | other |
| | | | | piano |
| | | | | tik tok |
| | | | | vocalist |
| | | | | zither |
| Total | 18 | 17 | 6 | 25 |

Table 7.4: Experiment 2: Mediation Outcomes

Table 7.4 summarises the mediation outcomes for Experiment 2. Compared with Experiment 1 (See Table 7.1) fewer concepts are majority adopted, while there are a greater number of minority adoption, conflict and rejection cases. The difference in the size of the participation pool should be considered here, it may be that the adjustments to the adoption thresholds was not sufficient to allow for more majority adoption cases.

The crowd for Experiment 2 was four times the size, therefore you would expect a greater number of concepts to emerge, and this certainly is the case. If you include the concepts removed with WordNet, a total of 226 concepts are acquired from the crowd; 68 of which are mediated and 41 which remain valid and the end of the process. This compares with a total of 44 acquired concepts in Experiment 1, of which 27 remain valid at the end of the process. Therefore, ten times the number of concepts are discarded in Experiment 2 as is discarded in Experiment 1. Of course, with the addition of the greater number of participants combined with the removal of quality control and the inclusion of WordNet processing – this was always likely to be the case.

Experiment 2 has successfully constrained the number of mediation questions to a reasonable level (56 in total), and produced a final model that is of a reasonable composition meaning that no obvious invalid concepts are majority adopted. Also, like Experiment 1, the minority adopted concepts and those with that remain in conflict are reasonable. Less minority adoption outcomes would be an objective if repeating this experiment with different thresholds, as the proportion with this outcome is high and would require some work for a knowledge engineer to resolve.

The analysis of the either/or questions are also interesting as it appears to show that people will change their mind on some occasions when there is clearly a better option being presented. This is the case when participants were asked to choose between 'song' being a type of 'thing' or a type of 'release' (See table 5.8) – given this choice most participants went with the more specific 'release', in many cases changing their original concept model.

The final assessment that needs to be made of Experiment 2 is to determine whether useful knowledge is lost through the application of the WordNet processes. A glance at Table 5.5 will tell you that a majority of the concepts removed are of no value in building an accurate knowledge model of the domain. In most cases, the concept has been removed because it is formatted incorrectly, is clearly an instance, or it is a composite word created by the participant due to a misunderstanding of how to correctly specifying a concept (for example where a participant has specified a concepts called *ABassist* and *BBassist*, presumably to indicate distinct instances rather than a concept per se). However some concepts that have been removed could be useful in building a knowledge representation, these are listed in Table 7.5.

| Removed Concept |
| ---: |
| dj |
| popular record |
| singer songwriter |
| writing award |
| special record type |
| best singer award |
| girl band |
| boy band |
| instrumentalist songwriter |
| extended play |
| international |
| singer instrumentalist |
| mixed band |
| rhythm and blues |
| classic record |
| band member |
| easy listening |
| live record type |
| band member |

Table 7.5: Experiment 2: Potentially useful concepts which were removed

This shows that the processes used to reduce the mediation burden, while successful in managing the number of questions generated, has removed some useful concepts from the final knowledge representation. Despite this, considering the fact that there was far more data and less control of input, Experiment 2 has shown that the approach being offered can be scaled upwards and is adaptable.

### 7.1.3 Quality of Acquired Knowledge from Experiment 3

Table 7.6 summarises the outcomes of the mediation process obtained from the second experiment. Figure 7.2 illustrates the consensual hierarchy obtained from campus map. As with the earlier figure, those concepts marked (*) are those which only have a Minority Adoption. Those marked twice (**) are the concepts where no resolution was found to the model conflict. In the case of unresolved model conflict multiple versions of that concept are included to show the differences in the model.

#### Majority Adoption

The concepts that were majority adopted in the Experiment 3 were all valid with no major issues regarding the modelling. The adoption of the concept *Region*, which could

| | **Majority Adopt** | **Minority Adopt** | **In Conflict** | **Rejected** |
|---|---|---|---|---|
| | Structure* | Bus_stop | Campus | borough |
| | Building | University | Department | drinking |
| | Shop | Academic | City | eating |
| | Pub | Shopping | | enterprize_zone |
| | Office | | | faculty_building |
| | Library | | | food |
| | Academic_building | | | open_space |
| | Department_building | | | recreation |
| | Commercial_building | | | park |
| | Administration_building | | | retail |
| | Area* | | | school |
| | Region | | | support_building |
| | Recreation | | | university_building |
| | Administration | | | |
| Total | 14 | 4 | 3 | 13 |

TABLE 7.6: Experiment 3: Mediation Outcomes
(Concepts marked with * were predefined)

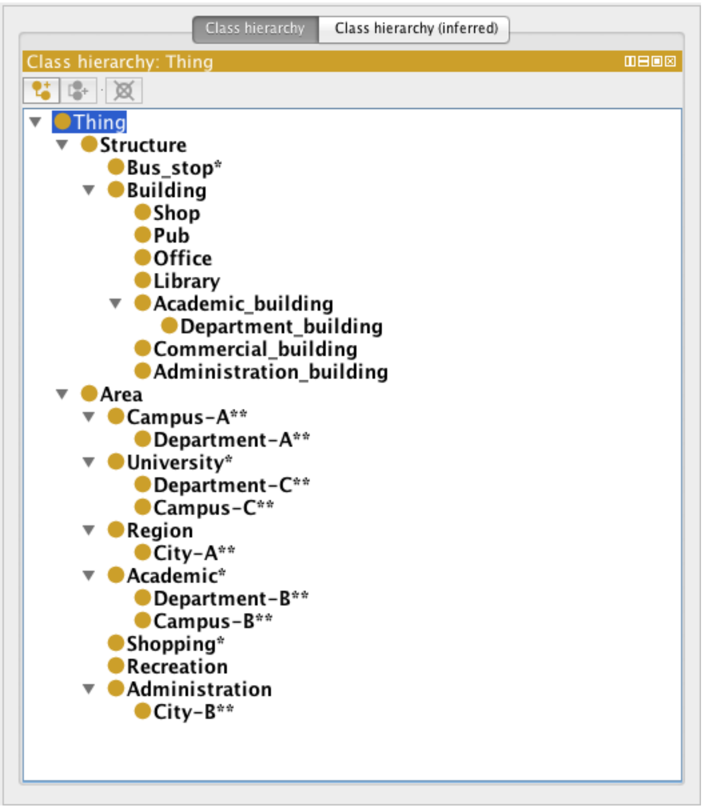


FIGURE 7.2: Experiment 3: Consensual Model Concept Hierarchy

be considered equivalent to the base concept *Area*, is the only slight concern.

**Minority Adoption**

*Bus_stop* as a type of structure, and *Academic*, *Shopping* and *University* as types of area are minority adopted. These are all excellent candidates for adjudication by the knowledge engineer. The distinction between *Academic* and *University* is subtle, if it exists at all. A knowledge engineer might be able to distinguish the two sets by determining that an *Academic* area can include colleges and other educational institutions; or they might simply define the two areas as equivalent. Needless to say, this is the type of decision that should be left to a knowledge engineer. Also, if the knowledge engineer were to assert that *Academic* and *University* areas were equivalent, then the subclasses of those concepts (*Campus* and *Department*), which are in conflict, would be reassigned to Minority Adoption status. A *Shopping* area as a concept may be considered valid, as would a *Bus_stop* structure. Here the knowledge engineer would incorporate these concepts according to the needs of the required model.

**Model Conflicts**

There are three concepts whose models are in conflict not resolved to a consensus: *Campus*, *Department* and *City*. Again, these concepts are good candidates for adjudication by the knowledge engineer.

## 7.2 Future Work

### 7.2.1 Retaining additional knowledge from the crowd

A persistent outcome the mediation process is the large amount of data that is discarded through pruning. The most common outcome of mediation is for a submitted concept model to be removed when it fails to gain the support needed to to be either minority or majority adopted. In general, the mediation protocols are biased towards the exclusion of concepts as opposed to inclusion. This is because the effect of a false-positive (invalid classes that are majority adopted and therefore not subject to knowledge engineer scrutiny) would undermine the entire approach by formalising erroneous concepts and producing flawed ontologies.

In its present form, the mediation process produces core components of the desired domain ontology consisting of the most easily agreed-upon concept models. For this reason, the mediation process is primarily useful for building the core module of an ontology. While the creation of comprehensive and robust domain ontologies is probably beyond the abilities of crowds at present, much work could be done towards obtaining greater coverage and identifying valid concepts before they are discarded. This is particularly relevant when additional processes are being used to reduce the number of mediation questions being produced. With the incorporation and development of

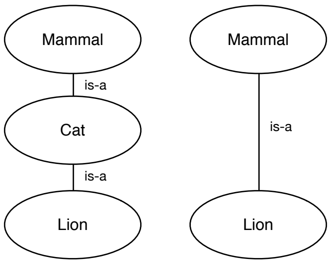


Figure 7.3: Concept Granularity

additional processes that automate the exclusion of concept models (such as the use of *WordNet* demonstrated in Experiment 2) parallel development of tools that help recognise potentially valid concept models is required.

Developing a two-stage application of this approach, in which a core ontology is produced first and then extended using the same approach at a later point, may also be possible. The core ontology created in first stage would perform a similar function in the second stage as the seed ontology in Experiment 3. In essence the first stage ontology would become the seed ontology of the second stage ontology. While some thought would have to be invested into how exactly this might work, the adaptability of the protocols would make this a promising area of development.

### 7.2.2 Determining types of conflict

A potential problem occurs when two competing concept models exist, both of which are semantically correct, yet one is discarded in favour of the other. For example, in reference to Figure 7.3, we can see that two competing concept models (*Lion is a Cat* and *Lion is a Mammal*) could be marked as being in conflict, yet no conflict really exists. This is an issue of granularity, whereby the correct concept model would depend on what level of detail is intended to be represented. To some extent, because both competing concept models are correct, they are both likely to receive support meaning that majority adoption thresholds would not be met and therefore both concept models would be retained for manual resolution (minority adoption). This is evident in the results of the either/or questions in Experiment 2 (see Table 5.8) where conflicts, such as "is *album* a *record* or *recording*" fail to gain majority adoption. While identifying these cases and retaining them is useful, more work needs to be done on automating this process to reduce the burden on the knowledge engineer. To achieve this, the precise nature of what constitutes a conflict would need to be developed further with additional protocols to determine how these conflicts are resolved and what information should be retained. By finding ways to do this, knowledge engineers will only be presented with conflicts that genuinely require expert-input to resolve.

### 7.2.3 Overloading the *is-a* relation

In Experiment 3 a folder analogy was used to help task-unaware crowds classify concepts. While this was useful as a device to lead unaware crowds towards performing classification, it has lead to a forced simplification over what relationships can be defined. For example, users are forced to classify mereological (part-of) relationships in the same way as they would *is-a* relationships. So, if developing a motoring domain ontology, an engine could reasonably be placed inside the folder representing the *automobile* concept, yet *engines* are not a type of *automobile*. This is not something that is easy to remedy, but in geospatial domains it may be possible to incorporate some simple rules that would better determine the precise nature of inter-concept relationships. Using Region Connection Calculus (RCC)[91], for example, could provide a standard way to describe the precise nature of the relationships between concepts that would typically have a spatial profile. RCC defines the types of spatial relationships that two regions can have. Using RCC you could, for instance, determine that a buildings is contained entirely within an area (such as a department building within a campus) and would therefore have a *within* relationship rather than a simplistic and erroneous *is-a* relationship.

As with all crowdsourcing endeavours, the user-experience is essential in obtaining ideal behaviour from the crowd. Further experimentation is required to determine what additional mechanisms could be incorporated into this approach that would encourage an improvement in the quality of crowd input. Using more sophisticated behavioural prompting to promote useful crowd behaviour, in conjunction with guidance from a set of logical relationship rules, would be one way of resolving this issue.

### 7.2.4 Too many reinstate / mediation questions?

While the use of *WordNet* successfully reduced the number of mediation and reinstate questions presented to the participants, with a larger crowd there would inevitably be a point where too many questions would be generated. To address this, it may be useful to segment the generated questions into manageable groups that would then be distributed to each user (rather than a full set of questions). This would mean that each user would only have to answer a subset of the generated questions. Given a big enough crowd, it may be possible to have enough participation in the mediation process for each group, that this type of distributed mediation might be successful.

## 7.3 Summary

In Experiment 1 the approach performs well in terms of convergence towards a consensual model, the price of this convergence is the loss of some of the acquired knowledge leading to a reduction in the reasoning capacity of the consensual model. This is illustrated by the reduced number of competency questions that the consensual model can answer. A possible reason for this is that less concept entity conflicts were identified. To improve this a more sophisticated formatting regime could be implemented that deals with the complex mismatch errors and which can identify synonyms (See Section 3.1.5).

Getting users to specify meaningful object properties proved a difficult task in both experiments, in Experiment 3, too few participants specified object properties to be worthy of reporting. Even in Experiment 1, where the participants should have had the required competency to specify effective object properties, the results were disappointing.

In Experiment 2, the approach was successfully used to converge a larger number of competing ontologies towards a consensus. This success was facilitated by the inclusion of additional processes and by creating more sophisticated mediation questions.

# Chapter 8

# Conclusion

The research described in this work attempts to test the viability of acquiring semantic resources, in the form of basic ontology structures, from online community crowds. The argument is made that a traditional view of crowdsourcing, which focuses on the scale of crowds, ignores the potential of smaller, more specialised crowds who represent online communities and could potentially represent the expertise needed to build domain ontologies. The stated advantages of using these community crowds are that the informal community membership requirements provide a minimum expertise level, and that community activity can be utilised to unwittingly build semantic resources. A method of knowledge acquisition is specified that defines how information elicited from these community crowds could be mediated. This includes a mechanism (in the form of a set of protocols) that allows the community themselves to validate the input of other users in order to build consensus and to organise information in such a way as it can be easily adapted to become a fully-fledged ontology. A formal evaluation method is then described the determines what levels of consensus can be achieved from this approach. This adapts an established ontology similarity evaluation method that determines the similarity between a reference and a learned ontology, so that a reference ontology can be compared to multiple ontologies. The adapted evaluation method determines the convergence of the community crowdsourced ontologies towards a consensus. It is argued that this is a good evaluation method for situations where no 'gold standard' exists, as it at least provides a representation of what is agreed upon amongst a community crowd, a crowd which is known to be knowledgeable of the target domain. The knowledge acquisition process is tested, firstly by using a community of students with knowledge engineering capability (Experiments 1 and 2) and, secondly by acquiring the knowledge through an interactive map interface aimed at encouraging a community crowd (Experiment 3). The results of the three experiments show that a convergence towards consensus is achieved in both experiments. These are promising results, as harnessing community crowds to elicit and validate the knowledge needed to build ontologies could dramatically broaden the coverage and use of domain ontologies.

# Appendix A

# Acquired Ontologies

This appendix describes the initial ontologies and the final moderated ontologies. The full set of ontologies in digital form (.owl) can be found at `https://github.com/roscminni/crowdsourcing-semantic-resources/blob/master/ontology-files.zip`.

## A.1 Unprocessed Ontologies acquired from Experiment 1

Group 3 - Initial Ontology Structure

### Classes

**Album**

Album $\equiv$ Record

**Awards**

**ClickRate**

**Composer**

Composer $\sqsubseteq$ Person

**Continent**

**Country**

**DownloadTimes**

**Gender**

**Genre**

**Group**

**Member**

Member $\sqsubseteq$ Person

**Name**

**Person**

**Price**

**Purchases**

**Record**

Record $\equiv$ Album

**ReleaseDate**

**Singer**

Singer $\sqsubseteq$ Person

**Song**

**Year**

## Object properties

**Artist_genre**

**BelongsToContinentOf**

$\sqsubseteq$ BelongsToCountryOf

$\exists$ BelongsToContinentOf Thing $\sqsubseteq$ Group

$\exists$ BelongsToContinentOf Thing $\sqsubseteq$ Singer

$\top \sqsubseteq \forall$ BelongsToContinentOf Continent

**BelongsToCountryOf**

$\exists$ BelongsToCountryOf Thing $\sqsubseteq$ Group
$\exists$ BelongsToCountryOf Thing $\sqsubseteq$ Singer
$\top \sqsubseteq \forall$ BelongsToCountryOf Country

**Has_DownloadTimes**

$\exists$ Has_DownloadTimes Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ Has_DownloadTimes DownloadTimes

**Has_clickRate**

$\exists$ Has_clickRate Thing $\sqsubseteq$ Singer
$\exists$ Has_clickRate Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ Has_clickRate ClickRate

**Has_gender**

$\exists$ Has_gender Thing $\sqsubseteq$ Singer
$\exists$ Has_gender Thing $\sqsubseteq$ Composer
$\exists$ Has_gender Thing $\sqsubseteq$ Member
$\top \sqsubseteq \forall$ Has_gender Gender

**Has_genre**

$\exists$ Has_genre Thing $\sqsubseteq$ Record
$\exists$ Has_genre Thing $\sqsubseteq$ Person
$\exists$ Has_genre Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ Has_genre Genre

**Has_member**

$\exists$ Has_member Thing $\sqsubseteq$ Group
$\top \sqsubseteq \forall$ Has_member Member

**Has_name**

$\exists$ Has_name Thing $\sqsubseteq$ Record
$\exists$ Has_name Thing $\sqsubseteq$ Composer
$\exists$ Has_name Thing $\sqsubseteq$ Group
$\exists$ Has_name Thing $\sqsubseteq$ Singer
$\exists$ Has_name Thing $\sqsubseteq$ Song
$\exists$ Has_name Thing $\sqsubseteq$ Member
$\top \sqsubseteq \forall$ Has_name Name

**Has_price**

$\exists$ Has_price Thing $\sqsubseteq$ Record
$\exists$ Has_price Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ Has_price Price

**Has_record**

$\exists$ Has_record Thing $\sqsubseteq$ Group
$\exists$ Has_record Thing $\sqsubseteq$ Singer
$\top \sqsubseteq \forall$ Has_record Record

**Has_releaseDate**

$\exists$ Has_releaseDate Thing $\sqsubseteq$ Song
$\exists$ Has_releaseDate Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ Has_releaseDate ReleaseDate

**Has_song**

$\exists$ Has_song Thing $\sqsubseteq$ Group
$\exists$ Has_song Thing $\sqsubseteq$ Singer
$\top \sqsubseteq \forall$ Has_song Song

**Sings_song**

¡http://MoKi_light#Sung_by¿ $\equiv$ ¡http://MoKi_light#Sings_song¿$^{-}$
$\exists$ Sings_song Thing $\sqsubseteq$ Group
$\exists$ Sings_song Thing $\sqsubseteq$ Singer
$\top \sqsubseteq \forall$ Sings_song Song

**Sung_by**

¡http://MoKi_light#Sung_by¿ $\equiv$ ¡http://MoKi_light#Sings_song¿$^{-}$

**Won_award**

$\exists$ Won_award Thing $\sqsubseteq$ Group
$\exists$ Won_award Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Won_award Awards

**Writes**

$\exists$ Writes Thing $\sqsubseteq$ Composer
$\top \sqsubseteq \forall$ Writes Song

### publishedYear

$\exists$ publishedYear Thing $\sqsubseteq$ Album
$\top \sqsubseteq \forall$ publishedYear Year

### signatureSong

$\exists$ signatureSong Thing $\sqsubseteq$ Group
$\exists$ signatureSong Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ signatureSong Song

## Data properties

### numberOfSales

$\exists$ numberOfSales Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Person
$\exists$ numberOfSales Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Song
$\exists$ numberOfSales Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Group
$\exists$ numberOfSales Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Album
$\top \sqsubseteq \forall$ numberOfSales Datatypehttp://www.w3.org/2001/XMLSchema#int

Group 4 - Initial Ontology Structure

## Classes

### Artist

Artist $\sqsubseteq$ Person

### Chart

### Gender

### Genre

### Group

### Person

### PopularGroup

PopularGroup $\sqsubseteq$ Group

### PopularSong

PopularSong $\sqsubseteq$ Song

### Record

### Song

### Year

## Object properties

### Belongs_to

$\exists$ Belongs_to Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Belongs_to Group

### Has_chart

$\exists$ Has_chart Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ Has_chart Chart

### Has_gender

$\exists$ Has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Has_gender Gender

**Has_genre**

$\exists$ Has_genre Thing $\sqsubseteq$ Record
$\exists$ Has_genre Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ Has_genre Genre

**Has_year**

$\exists$ Has_year Thing $\sqsubseteq$ Record
$\exists$ Has_year Thing $\sqsubseteq$ Chart
$\top \sqsubseteq \forall$ Has_year Year

**NoOneSong**

$\exists$ NoOneSong Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ NoOneSong Year

**Performs**

$\exists$ Performs Thing $\sqsubseteq$ Artist
$\top \sqsubseteq \forall$ Performs Song

**Produce**

$\exists$ Produce Thing $\sqsubseteq$ Group
$\exists$ Produce Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Produce Record

Group 5 - Initial Ontology Structure

## Classes

**Award**

**Gender**

**Group**

**Number_One_Record**

Number_One_Record $\sqsubseteq$ Record

**Number_One_Song**

Number_One_Song $\sqsubseteq$ Song

**Person**

**Price**

**Producer**

Producer $\sqsubseteq$ Person

**Record**

**Record_Label**

**Recording_Artist**

Recording_Artist $\sqsubseteq$ Person

**Song**

**Store**

**TimePeriod**

**Writer**

Writer $\sqsubseteq$ Person

**Year**

Year $\sqsubseteq$ TimePeriod

## Object properties

### Belongs_to

$\exists$ Belongs_to Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Belongs_to Group

### Has_gender

$\exists$ Has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Has_gender Gender

### Is_on

$\exists$ Is_on Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ Is_on Record

### Owns

$\exists$ Owns Thing $\sqsubseteq$ Record_Label
$\top \sqsubseteq \forall$ Owns Record

### Produce

$\exists$ Produce Thing $\sqsubseteq$ Producer
$\top \sqsubseteq \forall$ Produce Song

### Publishes

$\exists$ Publishes Thing $\sqsubseteq$ Person
$\exists$ Publishes Thing $\sqsubseteq$ Record_Label
$\top \sqsubseteq \forall$ Publishes Record
$\top \sqsubseteq \forall$ Publishes Song

### Sings

$\exists$ Sings Thing $\sqsubseteq$ Recording_Artist
$\top \sqsubseteq \forall$ Sings Song

### Won_award

$\sqsubseteq$ topObjectProperty
$\exists$ Won_award Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Won_award Award

### Writes

$\exists$ Writes Thing $\sqsubseteq$ Writer
$\top \sqsubseteq \forall$ Writes Song

### has_Price

$\exists$ has_Price Thing $\sqsubseteq$ Song
$\exists$ has_Price Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ has_Price Price

### topObjectProperty

## Data properties

### hasDownloadFunction

$\exists$ hasDownloadFunction Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Store
$\top \sqsubseteq \forall$ hasDownloadFunction Datatypehttp://www.w3.org/2001/XMLSchema#boolean

### hasSearchFunction

$\exists$ hasSearchFunction Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Store
$\top \sqsubseteq \forall$ hasSearchFunction Datatypehttp://www.w3.org/2001/XMLSchema#boolean

Group 6 - Initial Ontology Structure

## Classes

**Album**

**Artist**

Artist $\sqsubseteq$ Person

**Award**

**BestAlbum**

BestAlbum $\sqsubseteq$ Award

**BestGroup**

BestGroup $\sqsubseteq$ Award

**BestRecord**

BestRecord $\sqsubseteq$ Award

**BestSingle**

BestSingle $\sqsubseteq$ Award

**Composer**

Composer $\sqsubseteq$ Person

**EP**

EP $\sqsubseteq$ Record

**Genre**

**Label**

Label $\sqsubseteq$ Person

**Offer**

Offer $\sqsubseteq$ Store

**Person**

**Producer**

Producer $\sqsubseteq$ Person

**Record**

**Song**

**Store**

## Object properties

**Compose**

$\exists$ Compose Thing $\sqsubseteq$ Composer
$\top \sqsubseteq \forall$ Compose Song

**Has**

$\exists$ Has Thing $\sqsubseteq$ Artist
$\exists$ Has Thing $\sqsubseteq$ Record
$\exists$ Has Thing $\sqsubseteq$ Album
$\top \sqsubseteq \forall$ Has Song
$\top \sqsubseteq \forall$ Has Genre

**Has_released**

$\exists$ Has_released Thing $\sqsubseteq$ Artist
$\top \sqsubseteq \forall$ Has_released Album

**Included**

$\exists$ Included Thing $\sqsubseteq$ Award
$\top \sqsubseteq \forall$ Included BestAlbum
$\top \sqsubseteq \forall$ Included BestGroup
$\top \sqsubseteq \forall$ Included BestSingle
$\top \sqsubseteq \forall$ Included BestRecord

**Labelof**

$\exists$ Labelof Thing $\sqsubseteq$ Label
$\top \sqsubseteq \forall$ Labelof Artist

**Produce**

$\exists$ Produce Thing $\sqsubseteq$ Producer
$\top \sqsubseteq \forall$ Produce Song

**Provide**

$\exists$ Provide Thing $\sqsubseteq$ Store
$\top \sqsubseteq \forall$ Provide Offer

**Won_award**

$\exists$ Won_award Thing $\sqsubseteq$ Album
$\exists$ Won_award Thing $\sqsubseteq$ Record
$\exists$ Won_award Thing $\sqsubseteq$ Song
$\exists$ Won_award Thing $\sqsubseteq$ Artist
$\top \sqsubseteq \forall$ Won_award Award

Group 8 - Initial Ontology Structure

## Classes

**Award**

**Country**

**Covers**

Covers $\sqsubseteq$ Song

**First**

First $\sqsubseteq$ No._1

**Gender**

Gender $\sqsubseteq$ Person

**Genre**

**Group**

Group $\sqsubseteq$ Person

**No._1**

No._1 $\sqsubseteq$ Song

**Person**

**Price**

**Record**

Record $\sqsubseteq$ Song

**Sales**

**Song**

**Year**

## Object properties

**Achieved**

$\exists$ Achieved Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Achieved No._1

**Best_known**

$\exists$ Best_known Thing $\sqsubseteq$ Group
$\top \sqsubseteq \forall$ Best_known Genre

**Has_gender**

$\exists$ Has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Has_gender Gender

**Has_sold**

$\exists$ Has_sold Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Has_sold Sales

**Sings_song**

<http://MoKi_light#Sings_song> $\equiv$ <http://MoKi_light#Sung_by>$^{-}$

**Sold_in**

$\exists$ Sold_in Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ Sold_in Country

**Sung_by**

<http://MoKi_light#Sings_song> $\equiv$ <http://MoKi_light#Sung_by>$^{-}$
$\exists$ Sung_by Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ Sung_by Person

**Won_award**

$\exists$ Won_award Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ Won_award Award

**has_price**

$\exists$ has_price Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ has_price Price

**publish_year**

$\exists$ publish_year Thing $\sqsubseteq$ Song

## A.2 Mediated Ontologies from Experiment 1

Group 3 - Mediated Ontology Structure

### Classes

**Gender**

**Genre**

**Group**

**Person**

**Record**

**Song**

**artist**

artist $\sqsubseteq$ Person

**award**

**composer**

composer $\sqsubseteq$ artist

**country**

**price**

**store**

**year**

### Object properties

**belongs_to_country_of**

$\exists$ belongs_to_country_of Thing $\sqsubseteq$ Group
$\top \sqsubseteq \forall$ belongs_to_country_of country

**best_known_for**

$\exists$ best_known_for Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ best_known_for Record
$\top \sqsubseteq \forall$ best_known_for Song

**has_gender**

$\exists$ has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ has_gender Gender

**has_genre**

$\exists$ has_genre Thing $\sqsubseteq$ Record
$\exists$ has_genre Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ has_genre Genre

**has_price**

$\exists$ has_price Thing $\sqsubseteq$ Record
$\exists$ has_price Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ has_price price

**has_record**

$\exists$ has_record Thing $\sqsubseteq$ Group
$\top \sqsubseteq \forall$ has_record Record

**has_song**

$\exists$ has_song Thing $\sqsubseteq$ Group
$\top \sqsubseteq \forall$ has_song Song

**number_one**

$\exists$ number_one Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ number_one year

**sings_song**

<http://MoKi_light#sung_by> $\equiv$ <http://MoKi_light#sings_song>$^{-}$
$\exists$ sings_song Thing $\sqsubseteq$ Group
$\top \sqsubseteq \forall$ sings_song Song

**sung_by**

<http://MoKi_light#sung_by> $\equiv$ <http://MoKi_light#sings_song>$^{-}$
$\exists$ sung_by Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ sung_by Group

**won_award**

$\exists$ won_award Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ won_award award

**writes**

$\exists$ writes Thing $\sqsubseteq$ composer
$\top \sqsubseteq \forall$ writes Song

## Data properties

## Individuals

## Datatypes

Group 4 - Mediated Ontology Structure

## Classes

**Gender**

**Genre**

**Group**

**Person**

**Record**

**Song**

**Thing**

**artist**

artist $\sqsubseteq$ Person

**award**

**month**

month $\sqsubseteq$ Thing

**store**

**week**

week $\sqsubseteq$ Thing

**year**

year $\sqsubseteq$ Thing

## Object properties

**belongs_to**

$\exists$ belongs_to Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ belongs_to Group

**has_gender**

$\exists$ has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ has_gender Gender

**has_genre**

$\exists$ has_genre Thing $\sqsubseteq$ Record
$\exists$ has_genre Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ has_genre Genre

**has_year**

$\exists$ has_year Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ has_year year

**produce**

$\exists$ produce Thing $\sqsubseteq$ Group
$\exists$ produce Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ produce Record

**won_award**

$\exists$ won_award Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ won_award award

## Data properties

**Has_downloadFunction**

**hasSearchFunction**

$\exists$ hasSearchFunction Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ store
$\top \sqsubseteq \forall$ hasSearchFunction Datatypehttp://www.w3.org/2001/XMLSchema#boolean

## Individuals

## Datatypes

**boolean**

Group 5 - Mediated Ontology Structure

# Classes

**Gender**

**Genre**

**Group**

**Person**

**Record**

**Song**

**artist**

artist $\sqsubseteq$ Person

**award**

**month**

**producer**

producer $\sqsubseteq$ Person

**store**

**week**

**year**

# Object properties

**belongs_to**

$\exists$ belongs_to Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ belongs_to Group

**has_gender**

$\exists$ has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ has_gender Gender

**is_on**

$\exists$ is_on Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ is_on Record

**produce**

$\exists$ produce Thing $\sqsubseteq$ producer
$\top \sqsubseteq \forall$ produce Song

**won_award**

$\exists$ won_award Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ won_award award

## Data properties

**hasDownloadFunction**

$\exists$ hasDownloadFunction Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ store
$\top \sqsubseteq \forall$ hasDownloadFunction Datatypehttp://www.w3.org/2001/XMLSchema#boolean

**hasSearchFunction**

$\exists$ hasSearchFunction Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ store
$\top \sqsubseteq \forall$ hasSearchFunction Datatypehttp://www.w3.org/2001/XMLSchema#boolean

**price**

$\exists$ price Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Record
$\exists$ price Datatypehttp://www.w3.org/2000/01/rdf-schema#Literal $\sqsubseteq$ Song

Group 6 - Mediated Ontology Structure

## Classes

**Gender**

**Genre**

**Group**

**Person**

**Record**

**Song**

**Thing**

**artist**

artist $\sqsubseteq$ Person

**award**

**composer**

composer $\sqsubseteq$ Person

**month**

month $\sqsubseteq$ Thing

**price**

**producer**

producer $\sqsubseteq$ Person

**store**

**week**

week $\sqsubseteq$ Thing

**year**

year $\sqsubseteq$ Thing

## Object properties

**best_known_for**

$\exists$ best_known_for Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ best_known_for Genre

**compose**

$\exists$ compose Thing $\sqsubseteq$ composer
$\top \sqsubseteq \forall$ compose Song

**has**

$\exists$ has Thing $\sqsubseteq$ Record
$\exists$ has Thing $\sqsubseteq$ artist
$\top \sqsubseteq \forall$ has Song
$\top \sqsubseteq \forall$ has Genre

**has_gender**

$\exists$ has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ has_gender Gender

**has_price**

$\exists$ has_price Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ has_price price

**produce**

$\exists$ produce Thing $\sqsubseteq$ producer
$\top \sqsubseteq \forall$ produce Song

**won_award**

$\exists$ won_award Thing $\sqsubseteq$ artist
$\exists$ won_award Thing $\sqsubseteq$ Song
$\exists$ won_award Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ won_award award

<u>Group 8 - Mediated Ontology Structure</u>

## Classes

### Gender

Gender $\sqsubseteq$ Thing

### Genre

### Group

Group $\sqsubseteq$ Thing

### Person

### Record

Record $\sqsubseteq$ Song

### Song

### Thing

### artist

artist $\sqsubseteq$ Person

### award

### country

### month

### store

### week

### year

## Object properties

### achieved

$\exists$ achieved Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ achieved award

**best_known**

won_award $\exists$ best_known Thing $\sqsubseteq$ Group
$\top \sqsubseteq \forall$ best_known Genre

**has_gender**

$\exists$ has_gender Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ has_gender Gender

**sings_song**

<http://MoKi_light#sings_song> $\equiv$ <http://MoKi_light#sung_by>$^-$
$\exists$ sings_song Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ sings_song Song

**sold_in**

$\exists$ sold_in Thing $\sqsubseteq$ Record
$\top \sqsubseteq \forall$ sold_in country

**sung_by**

<http://MoKi_light#sings_song> $\equiv$ <http://MoKi_light#sung_by>$^-$
$\exists$ sung_by Thing $\sqsubseteq$ Song
$\top \sqsubseteq \forall$ sung_by Person

**won_award**

$\exists$ won_award Thing $\sqsubseteq$ Person
$\top \sqsubseteq \forall$ won_award award

## A.3 Unprocessed Ontologies acquired from Experiment 2

<u>Group 1 - Initial Ontology Structure</u>

## Classes

**Area**

**Building**

Building $\sqsubseteq$ Structure

**Department**

Department $\sqsubseteq$ University

**Drinking**

Drinking $\sqsubseteq$ Shopping

**Eating**

Eating $\sqsubseteq$ Shopping

**Recreation**

Recreation $\sqsubseteq$ Area

**Shopping**

Shopping $\sqsubseteq$ Area

**Structure**

**University**

University $\sqsubseteq$ Area

## Object properties

## Data properties

## Individuals

**Ashton_Building**

Ashton_Building : Building

**Augustus_John**

Augustus_John : Drinking

**Computer_Science_Department**

Computer_Science_Department : Department

**Costa_Coffee**

Costa_Coffee : Eating

**Electrical_Engineering_Department**

Electrical_Engineering_Department : Building

**Engineering_Department**

Engineering_Department : Department

**Foundation_Building_Parade**

Foundation_Building_Parade : Shopping

**Harold_Cohen_Library**

Harold_Cohen_Library : Building

**Holt_Building**

Holt_Building : Building

**Liverpool_John_Moores_University**

Liverpool_John_Moores_University : University

**Physics_Department**

Physics_Department : Department

**Quadrangle**

Quadrangle : Recreation

**Student_Union**

Student_Union : Department

**Subway**

Subway : Eating

**University_of_Liverpool**

University_of_Liverpool : University

**Victoria_Building**

Victoria_Building : Building

Group 2 - Initial Ontology Structure

## Classes

**Area**

**Building**

Building $\sqsubseteq$ Structure

**Campus**

Campus $\sqsubseteq$ Area

**City**

City $\sqsubseteq$ Region

**Department**

Department $\sqsubseteq$ Campus

**Faculty_Building**

Faculty_Building $\sqsubseteq$ University_Building

**Open_Space**

Open_Space $\sqsubseteq$ Area

**Other**

Other $\sqsubseteq$ Structure

**Region**

Region $\sqsubseteq$ Area

**Shop**

Shop $\sqsubseteq$ Other

**Structure**

Structure $\sqsubseteq$ Thing

**Thing**

**University_Building**

University_Building $\sqsubseteq$ Building

## Object properties

## Data properties

## Individuals

**Ashton_Building**

Ashton_Building : Faculty_Building

**Computer_Science_Department**

Computer_Science_Department : Department

**Costa_Coffee**

Costa_Coffee : Shop

**Electrical_Engineering_Department**

Electrical_Engineering_Department : Department

**Foundation_Building**

Foundation_Building : University_Building

**Greggs**

Greggs : Shop

**Harold_Cohen_Library**

Harold_Cohen_Library : University_Building

**Holt_Building**

Holt_Building : Faculty_Building

**Liverpool**

Liverpool : City

**Merseyside**

Merseyside : Region

**Quadrangle**

Quadrangle : Open_Space

**Sherrington_Building**

Sherrington_Building : Faculty_Building

**Student_Union**

Student_Union : University_Building

**Subway**

Subway : Shop

**Tesco**

Tesco : Shop

**University_of_Liverpool**

University_of_Liverpool : Campus

**Victoria_Building**

Victoria_Building : Faculty_Building

Group 3 - Initial Ontology Structure

## Classes

**Academic**

Academic $\sqsubseteq$ Area

**Academic_Building**

Academic_Building $\sqsubseteq$ Building

**Administration**

Administration $\sqsubseteq$ Area

**Administration_Building**

Administration_Building $\sqsubseteq$ Academic_Building

**Area**

**Borough**

Borough $\sqsubseteq$ Administration

**Building**

Building $\sqsubseteq$ Structure

**Campus**

Campus $\sqsubseteq$ Academic

**City**

City $\sqsubseteq$ Administration

**Commercial_Building**

Commercial_Building $\sqsubseteq$ Building

**Department**

Department $\sqsubseteq$ Academic

**Department_Building**

Department_Building $\sqsubseteq$ Academic_Building

**Enterprise_Zone**

Enterprise_Zone $\sqsubseteq$ Administration

**Food**

Food $\sqsubseteq$ Retail

**Park**

Park $\sqsubseteq$ Administration

**Retail**

Retail $\sqsubseteq$ Commercial_Building

**Structure**

**Support_Building**

Support_Building $\sqsubseteq$ Academic_Building

## Object properties

## Data properties

## Individuals

**Ashton_Building**

Ashton_Building : Department_Building

**Central**

Central : Borough

**Computer_Science_Department**

Computer_Science_Department : Department

**Costa_Coffee**

Costa_Coffee : Food

**Electrical_Engineering_Department**

Electrical_Engineering_Department : Department

**Foundation_Building**

Foundation_Building : Administration_Building

**Greggs**

Greggs : Food

**Harold_Cohen_Library**

Harold_Cohen_Library : Support_Building

**Holt_Building**

Holt_Building : Department_Building

**Liverpool**

Liverpool : City

**Liverpool_John_Moores_University**

Liverpool_John_Moores_University : Campus

**Liverpool_Science_Park**

Liverpool_Science_Park : Enterprise_Zone

**Material_Science_Deapartment**

Material_Science_Department : Department

**Quadrangle**

Quadrangle : Park

**Tescos**

Tescos : Retail

**University_Square**

University_Square : Park

**University_of_Liverpool**

University_of_Liverpool : Campus

**Victoria_Building**

Victoria_Building : Department_Building

Group 4 - Initial Ontology Structure

## Classes

### Area

### Building

Building $\sqsubseteq$ Structure

### Library

Library $\sqsubseteq$ Building

### Museum

Museum $\sqsubseteq$ Building

### Shop

Shop $\sqsubseteq$ Building

### Structure

## Object properties

## Data properties

## Individuals

### Ashton_Building

Ashton_Building : Building

### Costa_Coffee

Costa_Coffee : Shop

### Greggs

Greggs : Shop

### Harold_Cohen_Library

Harold_Cohen_Library : Library

**Holt_Building**

Holt_Building : Building

**University_of_Liverpool**

University_of_Liverpool : Area

**Victoria_Building**

Victoria_Building : Museum

Group 5 - Initial Ontology Structure

## Classes

### Academic

Academic $\sqsubseteq$ Area

### Academic_building

Academic_building $\sqsubseteq$ Building

### Area

### Building

Building $\sqsubseteq$ Structure

### Bus_stop

Bus_stop $\sqsubseteq$ Structure

### Campus

Campus $\sqsubseteq$ University

### Department

Department $\sqsubseteq$ University

### Museum

Museum $\sqsubseteq$ Building

### Office

Office $\sqsubseteq$ Building

### Pub

Pub $\sqsubseteq$ Building

### Region

Region $\sqsubseteq$ Area

**School**

School $\sqsubseteq$ University

**Shop**

Shop $\sqsubseteq$ Building

**Structure**

**University**

University $\sqsubseteq$ Area

## Object properties

## Data properties

## Individuals

**Ashton_Building**

Ashton_Building : Building

**Augustus_John**

**Computer_Science_Department**

Computer_Science_Department : Department

**Costa_Coffee**

**Electrical_Engineering_Department**

Electrical_Engineering_Department : Building

**Engineering_Department**

Engineering_Department : Department

**Foundation_Building_Parade**

**Harold_Cohen_Library**

Harold_Cohen_Library : Building

**Holt_Building**

Holt_Building : Building

**Liverpool_John_Moores_University**

Liverpool_John_Moores_University : University

**Physics_Department**

Physics_Department : Department

**Quadrangle**

**Student_Union**

Student_Union : Department

**Subway**

**University_of_Liverpool**

University_of_Liverpool : University

**Victoria_Building**

Victoria_Building : Building

## A.4 Mediated Ontologies from Experiment 2

Group 1 - Mediated Ontology Structure

# Classes

**Academic_building**

Academic_building $\sqsubseteq$ Building

**Administration**

Administration $\sqsubseteq$ Area

**Administration_building**

Administration_building $\sqsubseteq$ Building

**Area**

**Building**

Building $\sqsubseteq$ Structure

**Bus_stop**

Bus_stop $\sqsubseteq$ Structure

**Department**

Department $\sqsubseteq$ University

**Library**

Library $\sqsubseteq$ Building

**Museum**

Museum $\sqsubseteq$ Building

**Pub**

Pub $\sqsubseteq$ Building

**Recreation**

Recreation $\sqsubseteq$ Area

#### Region

Region $\sqsubseteq$ Area

#### Shop

Shop $\sqsubseteq$ Building

#### Shopping

Shopping $\sqsubseteq$ Area

#### Structure

#### University

University $\sqsubseteq$ Area

## Object properties

## Data properties

## Individuals

#### Ashton_Building

Ashton_Building : Building

#### Augustus_John

#### Computer_Science_Department

Computer_Science_Department : Department

#### Costa_Coffee

#### Electrical_Engineering_Department

Electrical_Engineering_Department : Building

#### Engineering_Department

Engineering_Department : Department

#### Foundation_Building_Parade

#### Harold_Cohen_Library

Harold_Cohen_Library : Building

**Holt_Building**

Holt_Building : Building

**Liverpool_John_Moores_University**

Liverpool_John_Moores_University : University

**Physics_Department**

Physics_Department : Department

**Quadrangle**

**Student_Union**

Student_Union : Department

**Subway**

**University_of_Liverpool**

University_of_Liverpool : University

**Victoria_Building**

Victoria_Building : Building

Group 2 - Mediated Ontology Structure

## Classes

**Academic_building**

Academic_building $\sqsubseteq$ Building

**Administration**

Administration $\sqsubseteq$ Area

**Administration_building**

Administration_building $\sqsubseteq$ Building

**Area**

**Building**

Building $\sqsubseteq$ Structure

**Campus**

Campus $\sqsubseteq$ Area

**City**

City $\sqsubseteq$ Region

**Department**

Department $\sqsubseteq$ Campus

**Library**

Library $\sqsubseteq$ Building

**Museum**

Museum $\sqsubseteq$ Building

**Pub**

Pub $\sqsubseteq$ Building

### Recreation

Recreation $\sqsubseteq$ Area

### Region

Region $\sqsubseteq$ Area

### Shop

Shop $\sqsubseteq$ Building

### Structure

Structure $\sqsubseteq$ Thing

### Thing

## Object properties

## Data properties

## Individuals

### Ashton_Building

### Computer_Science_Department

Computer_Science_Department : Department

### Costa_Coffee

Costa_Coffee : Shop

### Electrical_Engineering_Department

Electrical_Engineering_Department : Department

### Foundation_Building

### Greggs

Greggs : Shop

**Harold_Cohen_Library**

**Holt_Building**

**Liverpool**

Liverpool : City

**Merseyside**

Merseyside : Region

**Quadrangle**

**Sherrington_Building**

**Student_Union**

**Subway**

Subway : Shop

**Tesco**

Tesco : Shop

**University_of_Liverpool**

University_of_Liverpool : Campus

**Victoria_Building**

Group 3 - Mediated Ontology Structure

## Classes

### Academic

Academic $\sqsubseteq$ Area

### Academic_Building

Academic_Building $\sqsubseteq$ Building

### Administration

Administration $\sqsubseteq$ Area

### Administration_building

Administration_building $\sqsubseteq$ Building

### Area

### Building

Building $\sqsubseteq$ Structure

### Bus_stop

Bus_stop $\sqsubseteq$ Structure

### Campus

Campus $\sqsubseteq$ Academic

### City

City $\sqsubseteq$ Administration

### Department

Department $\sqsubseteq$ Academic

### Library

Library $\sqsubseteq$ Building

### Museum

Museum $\sqsubseteq$ Building

### Pub

Pub $\sqsubseteq$ Building

### Recreation

Recreation $\sqsubseteq$ Area

### Region

Region $\sqsubseteq$ Area

### Shop

Shop $\sqsubseteq$ Building

### Structure

## Object properties

## Data properties

## Individuals

### Ashton_Building

### Central

### Computer_Science_Department

Computer_Science_Department : Department

### Costa_Coffee

### Electrical_Engineering_Department

Electrical_Engineering_Department : Department

**Foundation_Building**

**Greggs**

**Harold_Cohen_Library**

**Holt_Building**

**Liverpool**

Liverpool : City

**Liverpool_John_Moores_University**

Liverpool_John_Moores_University : Campus

**Liverpool_Science_Park**

**Material_Science_Deapartment**

Material_Science_Deapartment : Department

**Quadrangle**

**Tescos**

**University_Square**

**University_of_Liverpool**

University_of_Liverpool : Campus

**Victoria_Building**

Group 4 - Mediated Ontology Structure

## Classes

**Academic**

Academic $\sqsubseteq$ Area

**Academic_building**

Academic_building $\sqsubseteq$ Building

**Administration**

Administration $\sqsubseteq$ Area

**Administration_building**

Administration_building $\sqsubseteq$ Building

**Area**

**Building**

Building $\sqsubseteq$ Structure

**Library**

Library $\sqsubseteq$ Building

**Museum**

Museum $\sqsubseteq$ Building

**Pub**

Pub $\sqsubseteq$ Building

**Recreation**

Recreation $\sqsubseteq$ Area

**Region**

Region $\sqsubseteq$ Area

**Shop**

Shop $\sqsubseteq$ Building

**Shopping**

Shopping $\sqsubseteq$ Area

**Structure**

**bus_stop**

bus_stop $\sqsubseteq$ Structure

## Object properties

## Data properties

## Individuals

**Ashton_Building**

Ashton_Building : Building

**Costa_Coffee**

Costa_Coffee : Shop

**Greggs**

Greggs : Shop

**Harold_Cohen_Library**

**Holt_Building**

Holt_Building : Building

**University_of_Liverpool**

University_of_Liverpool : Area

**Victoria_Building**

Victoria_Building : Museum

Group 5 - Mediated Ontology Structure

## Classes

**Academic**

Academic $\sqsubseteq$ Area

**Academic_building**

Academic_building $\sqsubseteq$ Building

**Administration**

Administration $\sqsubseteq$ Area

**Administration_building**

Administration_building $\sqsubseteq$ Building

**Area**

**Building**

Building $\sqsubseteq$ Structure

**Bus_Stop**

Bus_Stop $\sqsubseteq$ Structure

**Campus**

Campus $\sqsubseteq$ University

**Department**

Department $\sqsubseteq$ University

**Museum**

Museum $\sqsubseteq$ Building

**Recreation**

Recreation $\sqsubseteq$ Area

**Region**

Region $\sqsubseteq$ Area

**Shop**

Shop $\sqsubseteq$ Building

**Structure**

**University**

University $\sqsubseteq$ Area

## Object properties

## Data properties

## Individuals

**Ashton_Building**

Ashton_Building : Building

**Augustus_John**

**Computer_Science_Department**

Computer_Science_Department : Department

**Costa_Coffee**

**Electrical_Engineering_Department**

Electrical_Engineering_Department : Building

**Engineering_Department**

Engineering_Department : Department

**Foundation_Building_Parade**

**Harold_Cohen_Library**

Harold_Cohen_Library : Building

**Holt_Building**

Holt_Building : Building

**Liverpool_John_Moores_University**

Liverpool_John_Moores_University : University

**Physics_Department**

Physics_Department : Department

**Quadrangle**

**Student_Union**

Student_Union : Department

**Subway**

**University_of_Liverpool**

University_of_Liverpool : University

**Victoria_Building**

Victoria_Building : Building

# Appendix B

# Lab Instruction for Students (Experiment 1)

## B.1 Experiment 1

### Lab 2 - Domain Concept Modelling

A music store has asked you to produce a domain concept model that will help customers find the music they want to buy. In order to produce this model you will have to complete the following tasks:

1. Define five competency questions that could be asked of your domain concept model.
2. List all the terms that will be needed for your domain concept model.
3. Order your concepts and properties into hierarchical structures. You should specify the domain and range of your properties as precisely as you can.
4. Use *MoKi* to build a digital version of your model.

The following instructions will guide you through these tasks. You should describe your own conceptual model, therefore there is no 'correct' solution to this exercise. You are encouraged to use the paper form provided to make note of the reasons for the decisions you make during this exercise.

#### Task 1: Competency Questions

We will start by considering what competency questions we can use to help determine the scope and structure of our domain concept model. A good competency question will give you some idea of the structure that you require and will also determine what useful information can be retrieved from your model. In this case you should think about what information a customer in a music store might require or what information a retailer might want to communicate to the customer. Identify **five** competency questions that could be asked of a domain concept model for the music store. Write down your competency questions on the paper form provided.

### Task 2: Identify Terms

Using a pen and paper, list all the terms that you think should be included in your conceptual model of the music store domain. Your competency questions should help you determine what terms should be included. Include candidate terms for both concepts and properties. You should also include the following terms as your basic concepts: **Genre**, **Group**, **Person**, **Record** and **Song**.

### Task 3: Concept and Property Hierarchies

Your domain concept model will consist of **Concepts** and **Properties**. Divide the list of terms you have created into those that can be viewed as concepts and those that can be viewed as properties. Order your concepts into a hierarchy with more general terms towards the top of the tree and more qualified terms toward the bottom. Also, you may want to order your properties into a hierarchical structure, remembering that sub properties define a more specific relationship to that of their super property. After you have completed both hierarchies, annotate each property with its range and domain. Remember that the domain and range should use the most specific concepts that the property can be applied to, for example $Writer(D) \rightarrow authorOf \rightarrow Book(R)$ is more correct than $Person(D) \rightarrow authorOf \rightarrow Book(R)$.

## Task 4: Using MoKi to create a digital version of your Domain Concept Model

MoKi is a browser-based modelling tool for building models, specifically it is designed to enable collaborative concept model building. We will be using its basic functionality to create a digital version of your concept model. To use *MoKi* open Mozilla Firefox (MoKi is optimized for Firefox and will not work properly in other browsers) and do the following steps:

1. Go to `https://dkmtools.fbk.eu/moki/liverpool_session/groupXX/` where 'XX' is your group number. Click on the 'log in' link in the top-right corner of this page and enter your username as 'User1' and password 'lvpuser1' (up to five users can access each domain model at once, however for this exercise we recommend only one person should do the modelling).
2. You will now have access to *MoKi's* model building features. On the left-hand side of the page is the main menu. Your first task will be to add your concept tree to the model, to do this click on the 'IsA[1] Browser' feature in the main menu.
3. You should now see the a basic tree structure containing the five predefined concepts from task 2. To extend this structure select the concept to be extended and click the 'Add' button. Enter the concept name (should be a unique name) and click 'OK'. The concept tree will now be extended to include your new concept. Repeat this process until you have completed your structure, it should now resemble the concepts hierarchy from task 3. If you need to delete a concept, select it in the browser and use the 'Delete' button.
4. Once you have completed your concept tree you will now want to define your properties. There is no equivalent browser for properties so you will want to create each property using the 'Add/Edit a Property' button in the main main menu. Add the property name and a entity description and then click the 'Save Page' button.
5. You have now created a new property and will be directed to the property page which by default will only contain the entity description. To add the domain, range, subproperty and superproperty relationships click 'Lightly Structured' link at the top of the page. You should now be able to add these relationships using the buttons provided.
6. Repeat 4 & 5 until you have entered all the properties from your properties hierarchy from task 3. When you have finished this process you may want click the 'List All Properties' button to check for any errors or missing information.

[1] *IsA* is a way of describing superclass and subclass relationships between concepts that distinguishes from the relationships between instances (individuals) and classes, for example a Lion is a *InstanceOf* Mammal and a Mammal *IsA* Animal.

7. If you need to delete a property click the 'Add/Edit Property', type the name of the property you want to delete, press return and go to the entity page. Once there, use the delete link at the top of the page to remove the concept or property from the model.

## B.2 Experiment 2

## Lab 2 - Domain Concept Modelling

A music store has asked you to produce a domain concept model that will help customers find the music they want to buy. In order to produce this model you will have to complete the following tasks:

1. Define five competency questions that could be asked of your domain concept model.
2. List all the terms that will be needed for your domain concept model.
3. Order your concepts and properties into hierarchical structures. You should specify the domain and range of your properties as precisely as you can.
4. Use *MoKi* to build a digital version of your model.

The following instructions will guide you through these tasks. You should describe your own conceptual model, therefore there is no 'correct' solution to this exercise. You are encouraged to use the paper form provided to make note of the reasons for the decisions you make during this exercise.

### Task 1: Competency Questions

We will start by considering what competency questions we can use to help determine the scope and structure of our domain concept model. A good competency question will give you some idea of the structure that you require and will also determine what useful information can be retrieved from your model. In this case you should think about what information a customer in a music store might require or what information a retailer might want to communicate to the customer. Identify **five** competency questions that could be asked of a domain concept model for the music store. Write down your competency questions on the paper form provided.

### Task 2: Identify Terms

Using a pen and paper, list all the terms that you think should be included in your conceptual model of the music store domain. Your competency questions should help you determine what terms should be included. Include candidate terms for both concepts and properties. You should use some higher level terms such as **Genre**, **Group**, **Person**, **Record** or **Song**.

### Task 3: Concept and Property Hierarchies

Your domain concept model will consist of **Concepts** and **Properties**. Divide the list of terms you have created into those that can be viewed as concepts and those that can be viewed as properties. Order your concepts into a hierarchy with more general terms towards the top of the tree and more qualified terms toward the bottom. Also, you may want to order your properties into a hierarchical structure, remembering that sub properties define a more specific relationship to that of their super property. After you have completed both hierarchies, annotate each property with its range and domain. Remember that the domain and range should use the most specific concepts that the property can be applied to, for example $Writer(D) \rightarrow authorOf \rightarrow Book(R)$ is more correct than $Person(D) \rightarrow authorOf \rightarrow Book(R)$.

### Task 4: Use Protégé to to create a digital version of your Domain Concept Model

You can now use Protégé to build a digital version of your domain concept model. Please refer the notes from lab 2 and 3 if you are not sure how to do this.

## Examples

The examples provided here use a movie domain concept model.

### Competency Questions

Q: Which female directors have won an academy award?

Q: What film is [Actor] best known for?

Q: Which films are remakes of a foreign language film?

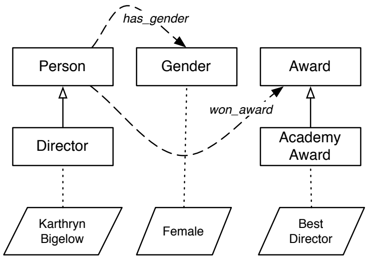


FIGURE B.1: Here is an example domain concept model fragment that would be able to answer the first competency question example. The individuals (instances) have been included here, but the important element is how the concepts and properties interact. Here a *Director* is a *Person* and a *Person* can have a *won_award* relationship with *Award* and a *has_gender* relationship with *Geder*

.

### Property Hierarchies

Below is an example of how a subproperty can be defined.

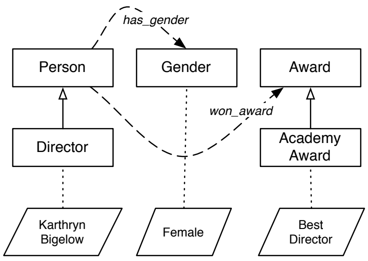


FIGURE B.2: Here is the same model fragment as obove with the addition of the property *won_academy_award*. *won_academy_award* is subproperty of *won_award* and is valid here because all winners of an academy award can also be said to have been winners of an award. This subproperty gives more precise information about the entities contained within the model.

# Appendix C

# CampusMap instructions

## Instructions

### General Notes

This map allows you to place map objects that correspond to places that would be of interest to new students and visitors. You are able to place map objects anywhere within the campus area, however it would be most helpful if you could concentrate on the area surrounding the Victoria Building Museum and Gallery.

### Login

To login to your campus map, use the control panel on the right hand side of the screen and click the 'Login' button.

If you have agreed to participate in this project, use the login provided in the acknowledgement email and then click 'OK'.If you have not registered, please contact *a.r.minnion@liv.ac.uk* and a login can be provided.

### Modes

When you first login you are in the neutral **Browse** mode.
In the right hand control panel you can choose to toggle on/off **Input Mode**,**Properties Mode** or **Edit Mode**

### View Options

Use the view options panel to view you existing map objects on the map.

### Input Mode

After entering input mode the border around the screen should turn green and the cursor should change to a cross-hair.

Now you need to plot your object on the map, to do this click on the screen in the desired location, and then continue clicking until an outline of your map object is made. Once you have placed 3 points on the screen, the 'Complete Shape'button appears, clicking this button will automatically join the last point to the first point specified to complete the shape and the user will move on to the next screen.

Now you will need to provide a name and brief description of your map object.

- Try to use the most common name for the object being represented
- Include any information such as alternative names in the description

The next step is to categorise your map object. Once you have completed the name and description click 'Continue'. Your map object will appear in the current object panel at the bottom left of the panel. Now you will be presented with a set of folders in which to place your map objects. To begin with you will have two folder Area and Structure. You can select these folders by clicking on them (they will be highlighted in blue). When you select a folder it appears as the target folder in the bottom-right panel.

You can add folders that will be created inside the selected folder. To do this click on the 'New Folder' button, specify a name and click 'OK'

Folders should represent groupings for your map objects, for example if you have various map objects that represent different student housing, then a folder named 'Student Housing' might be appropriate. If you would like to have an even greater level of differention, then you might also provide folders inside Student Housing for 'Halls of Residence' so that you can distinguish between purpose built student housing and normal housing which happens to be used for student accommodation.

***** You are strongly encouraged to add folders when you can *****

The way you organise your map objects will help make *CampusMap* more useful to new students and visitors.

Once you have created a map object and added a new folder (if necessary), then you can confirm the object by dragging it from the current object panel (bottom-left) to the selected folder panel (bottom-right). After a short wait for the database to be updated, a dialog box will confirm that your map object has been added.

### Edit Mode

Clicking on the 'Edit Mode' button, a dialogue box will appear where you will be able to move your map objects to different folders and delete any existing map object or folder that you have created. You can delete any folders you have created as long as they are empty.

### Properties Mode

This mode is not currently available.

# Bibliography

[1] R. Abbasi and S. Staab, *Organizing resources on tagging systems using t-org*, In Proc. of Bridging the Gap between Semantic Web and Web 2.0, workshop at ESWC 2007, 2007, pp. 97–110.

[2] D.M. Allen and T. Schneider, *The role of the community in a technical support community: A case study.*, Lecture Notes in Computer Science, vol. 8029 LNCS, (1)Technical Services, Colorado Springs, Cisco Systems, Inc., 2013.

[3] J. Angele, D. Fensel, D. Landes, and R. Studer, *Developing knowledge-based systems with mike*, Automated Software Engineering **5** (1998), no. 4, 389–418.

[4] S. Angeletou, *Semantic enrichment of folksonomy tagspaces.*, International Semantic Web Conference (Amit P. Sheth, Steffen Staab, Mike Dean, Massimo Paolucci, Diana Maynard, Timothy W. Finin, and Krishnaprasad Thirunarayan, eds.), Lecture Notes in Computer Science, vol. 5318, Springer, 2008, pp. 889–894.

[5] S. Auer, S. Dietzold, , and T. Riechert, *Ontowiki - a tool for social, semantic collaboration*, The Semantic Web - ISWC 2006, 5th International Semantic Web Conference, ISWC 2006, Springer, 2006, pp. 736–749.

[6] S. Bechhofer, Iain Buchan, D. De Roure, P. Missier, J. Ainsworth, J. Bhagat, P. Couch, D. Cruickshank, M. Delderfield, I. Dunlop, Matthew Gamble, Danius Michaelides, Stuart Owen, David Newman, Shoaib Sufi, and Carole Goble, *Why linked data is not enough for scientists.*, Future Generation Computer Systems **29** (2013), no. 2, 599 – 611.

[7] S. Bechhofer, A. Gangemi, N. Guarino, F. van Harmelen, Horrocks. I., M. Klein, C. Masolo, D. Oberle, S. Staab, H. Stuckenschmidt, and R. Volz, *Tackling the ontology acquisition bottleneck: An experiment in ontology re-engineering*, 2003.

[8] T. Berners-Lee, J. Hendler, and O. Lassila, *The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities.*, Scientific American **284** (2001), no. 5, 34 – 43.

[9] C. Bizer, T. Heath, and T. Berners-Lee, *Linked Data - The Story So Far*, International Journal on Semantic Web and Information Systems (IJSWIS) **5** (2009), no. 3, 1–22.

[10] S. L. Bryant, A. Forte, and Bruckman A., *Becoming wikipedian: Transformation of participation in a collaborative online encyclopedia*, Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work (New York), Association for Computing Machinery, 2005, pp. 1–10.

[11] T. Bylander and B. Chandrasekaran, *Generic tasks for knowledge-based reasoning: the "right" level of abstraction for knowledge acquisition*, International Journal of Man-Machine Studies **26** (1987), no. 2, 231 – 243.

[12] S. A. Caraballo, *Automatic construction of a hypernym-labeled noun hierarchy from text*, Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics (Stroudsburg, PA, USA), ACL '99, Association for Computational Linguistics, 1999, pp. 120–126.

[13] C. Caracciolo, J. Heguiabehere, A. Gangemi, C. Baldassarre, J Keizer, and M. Taconet, *Knowledge management at fao: A case study on network of ontologies in fisheries*, Ontology Engineering in a Networked World (Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta, and Aldo Gangemi, eds.), Springer Berlin Heidelberg, 2012, pp. 383–405 (English).

[14] I. Celino, S. Contessa, M. Corubolo, D. Dell'Aglio, E. D. Valle, S. Fumeo, and T. Krüger, *Urbanmatch - linking and improving smart cities data.*, LDOW (C. Bizer, T. Heath, T. Berners-Lee, and M. Hausenblas, eds.), CEUR Workshop Proceedings, vol. 937, CEUR-WS.org, 2012.

[15] P. Cimiano, *Ontology learning and population from text [electronic book] : algorithms, evaluation and applications / by philipp cimiano.*, Online access with purchase: Springer, Boston, MA : Springer Science+Business Media, LLC, 2006., 2006.

[16] P. Cimiano, A. Hotho, and S. Staab, *Learning concept hierarchies from text corpora using formal concept analysis*, Journal of Artificial Intelligence research **24** (2005), 305–339.

[17] W. J. Clancey, *Heuristic classification*, Artificial Intelligence **27** (1985), no. 3, 289–350.

[18] N. J. Cooke, *Varieties of knowledge elicitation techniques*, International Journal of Human-Computer Studies **41** (1994), no. 6, 801 – 849.

[19] M. d'Aquin, A. Gangemi, A. Passant, A. Isaac, and P. Laublet, *Combining linked data and knowledge engineering best practices to design a lightweight role ontology.*, Applied Ontology **6** (2011), no. 3, 223 – 246.

[20] Joseph G. Davis, *From crowdsourcing to crowdservicing.*, IEEE Internet Computing **15** (2011), no. 3, 92 – 94.

[21] F. de Haan, *Building a semantic map: Top-down versus bottom-up approaches.*, Linguistic Discovery **8** (2010), no. 1, 102 – 117.

[22] Klaas Dellschaft and Steffen Staab, *On how to perform a gold standard based evaluation of ontology learning*, In Proceedings of the 5th International Semantic Web Conference (ISWC'06, Springer, 2006, pp. 228–241.

[23] Robert Denaux, Catherine Dolbear, Glen Hart, Vania Dimitrova, and Anthony G Cohn, *Supporting domain experts to construct conceptual ontologies: A holistic approach*, Web Semantics: Science, Services and Agents on the World Wide Web **9** (2011), no. 2, 113 – 127, ¡ce:title¿Provenance in the Semantic Web¡/ce:title¿.

[24] Ronald Denaux, Ian Holt, Ilaria Corda, Vania Dimitrova, Catherine Dolbear, and Anthony G Cohn, *Roo: A tool to assist domain experts with ontology construction*, Proceedings of the 5th European Semantic Web Conference, 2008.

[25] Chiara Di Francescomarino, Chiara Ghidini, and Marco Rospocher, *Evaluating wiki-enhanced ontology authoring*, Proceedings of the 18th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2012), Lecture Notes in Computer Science, vol. 7603, Springer Berlin Heidelberg, 2012, pp. 292–301.

[26] A. Doan, R. Ramakrishnan, and A. Y. Halevy, *Crowdsourcing systems on the world-wide web.*, Communications of the ACM **54** (2011), no. 4, 86 – 96.

[27] Fefie Dotsika, *Uniting formal and informal descriptive power: Reconciling ontologies with folksonomies.*, International Journal of Information Management **29** (2009), no. 5, 407 – 415.

[28] F. Echarte, J. J. Astrain, A. Cordoba, and J. Villadangos, *Self-adaptation of ontologies to folksonomies in semantic web.*, Proceedings of World Academy of Science: Engineering & Technology **45** (2008), 335 – 341.

[29] O. Etzioni, M. Banko, and M. J. Cafarella, *Machine reading*, AAAI Spring Symposium: Machine Reading, AAAI, 2007, pp. 1–5.

[30] J. Fang and J. Evermann, *Evaluating ontologies: Towards a cognitive measure of quality*, EDOCW, IEEE Computer Society, 2007, pp. 109–116.

[31] E. A. Feigenbaum, *Knowledge engineering: : The applied side of artificial intelligence*, Annals of the New York Academy of Sciences **426** (1984), no. 1, 91–107.

[32] C. Fellbaum, *Wordnet: an electronic lexical database*, MIT Press, Cambridge, Mass, 1998.

[33] D. Fensel, *The knowledge acquisition and representation language karl*, Kluwer, 1995.

[34] Enrico Francesconi, Simonetta Montemagni, Wim Peters, and Daniela Tiscornia, *Integrating a bottom-up and top-down methodology for building semantic resources for the multilingual legal domain*, Semantic Processing of Legal Texts (Enrico Francesconi, Simonetta Montemagni, Wim Peters, and Daniela Tiscornia, eds.), Lecture Notes in Computer Science, vol. 6036, Springer Berlin / Heidelberg, 2010, pp. 95–121.

[35] S. Golder and B. A. Huberman, *The Structure of Collaborative Tagging Systems*, (2005).

[36] Asunciòn Gòmez-Pèrez, Mariano Fernàndez-Lòpez, and Oscar Corcho-Garcia, *Ontological engineering with examples from the areas of knowledge management, e-commerce and the semantic web*, Springer-Verlag New York, Inc., 2003.

[37] J. M. Gómez-Pérez, V. Méndez, J. Candini, and J. C. Muñoz, *Electronic invoice management in the pharmaceutical sector: The pharmainnova case*, Ontology Engineering in a Networked World (Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta, and Aldo Gangemi, eds.), Springer Berlin Heidelberg, 2012, pp. 407–422 (English).

[38] B. M. Good and M. D. Wilkinson, *Ontology engineering using volunteer labor*, WWW '07: Proceedings of the 16th international conference on World Wide Web (New York, NY, USA), ACM, 2007, pp. 1243–1244.

[39] P. F. Green and M. Rosemann, *An ontological analysis of integrated process modelling*, CAiSE (Matthias Jarke and Andreas Oberweis, eds.), Lecture Notes in Computer Science, vol. 1626, Springer, 1999, pp. 225–240.

[40] T. R. Gruber, *Toward principles for the design of ontologies used for knowledge sharing?*, International Journal of Human-Computer Studies **43** (1995), no. 5–6, 907 – 928.

[41] M. Grüninger and M. Fox, *Methodology for the Design and Evaluation of Ontologies*, IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, April 13, 1995, 1995.

[42] Glen Hart, Catherine Dolbear, and John Goodwin, *Lege feliciter: Using structured english to represent a topographic hydrology ontology.*, OWLED (Christine Golbreich, Aditya Kalyanpur, and Bijan Parsia, eds.), CEUR Workshop Proceedings, vol. 258, CEUR-WS.org, 2007.

[43] Glen Hart, Martina Johnson, and Catherine Dolbear, *Rabbit: Developing a control natural language for authoring ontologies*, The Semantic Web: Research and Applications (Sean Bechhofer, Manfred Hauswirth, Jörg Hoffmann, and Manolis Koubarakis, eds.), Lecture Notes in Computer Science, vol. 5021, Springer Berlin Heidelberg, 2008, pp. 348–360 (English).

[44] M.A. Hearst, *Automatic acquisition of hyponyms from large text corpora*, Proceedings of the 14th conference on computational linguistics, 1992, cited By (since 1996)1, pp. 539–545.

[45] T. Heath and C. Bizer, *Linked data: evolving the web into a global data space / tom heath, christian bizer.*, Synthesis lectures on the semantic web, theory and technology: 1, San Rafael, Calif. : Morgan & Claypool, 2011., 2011.

[46] J. Hendler, J. Holm, C. Musialek, and G. Thomas, *Us government linked open data: Semantic.data.gov*, Intelligent Systems, IEEE **27** (2012), no. 3, 25–31.

[47] R. Hoekstra, *The knowledge reengineering bottleneck.*, Semantic Web **1** (2010), no. 1-2, 111.

[48] C. W. Holsapple and K. D. Joshi, *A collaborative approach to ontology design*, Commun. ACM **45** (2002), no. 2, 42–47.

[49] M. Horridge, H Knublauch, A Rector, R Stevens, and C Wroe, *A Practical Guide To Building OWL Ontologies Using The Protege-OWL Plugin and CO-ODE Tools Edition 1.3*, (2009).

[50] Bernhard Humm and Alexey Korobov, *Introducing layers of abstraction to semantic web programming*, On the Move to Meaningful Internet Systems: OTM 2011 Workshops - Confederated International Workshops and Posters: EI2N+NSF ICE, ICSP+INBAST, ISDE, ORM, OTMA, SWWS+MONET+SeDeS, and VADER 2011, Hersonissos, Crete, Greece, October 17-21, 2011. Proceedings, 2011, pp. 412–423.

[51] P. Jain, P. Hitzler, P. Z. Yeh, K. Verma, and A. P. Sheth, *A.p.: Linked data is merely more data*, In: AAAI Spring Symposium 'Linked Data Meets Artificial Intelligence', AAAI, Press, 2010, pp. 82–86.

[52] L. Johnson and N. E. Johnson, *Knowledge elicitation involving teachback interviewing*, Knowledge Acquisition for Expert Systems (AlisonL. Kidd, ed.), Springer US, 1987, pp. 91–108 (English).

[53] E. Kapetanios, M. Schaal, E. Kapetanios, and M. Schaal, *A Model and Algebra for Collaborative Semantic Annotation in Digital Libraries.*

[54] A. Kittur and R. E. Kraut, *Harnessing the wisdom of crowds in wikipedia: Quality through coordination*, Cscw: 2008 Acm Conference on Computer Supported Cooperative Work, Conference Proceedings (2008), 37–46 732.

[55] A. Kittur, J. V. Nickerson, M. S. Bernstein, E. Gerber, A. D. Shaw, J. Zimmerman, M. Lease, and J. Horton, *The future of crowd work*, CSCW (Amy Bruckman, Scott Counts, Cliff Lampe, and Loren G. Terveen, eds.), ACM, 2013, pp. 1301–1318.

[56] C.-C. Kiu and E. Tsui, *Taxofolk: A hybrid taxonomy-folksonomy structure for knowledge classification and navigation.*, Expert Systems with Applications **38** (2011), no. 5, 6049–6058.

[57] G. Klyne, J. J. Carroll, and B. McBride, *Resource description framework (rdf): Concepts and abstract syntax*, February 2004.

[58] G. Kobilarov, T. Scott, S. Oliver, C. Sizemore, M. Smethurst, and C. Bizer, *Media meets semantic web - how the bbc uses dbpedia and linked data to make conections*, In European Semantic Web Conference, Semantic Web in Use Track, 2009.

[59] A. Koukourikos, G. Vouros, P. Karampiperis, and V. Karkaletsis, *Using open information extraction and linked open data towards ontology enrichment and alignment.*, Lecture Notes in Business Information Processing, vol. 112 LNBIP, (1)University of Piraeus, Department of Digital Systems, 2012.

[60] Katalin Kovacs, DolCatherin, Glen Hart, John Goodwin, and Hayley and Mizen, *A methodology for building conceptual domain ontologies*, Tech. report, Ordinance Survey Research Labs, 2006.

[61] S. Kuznetsov, *Motivations of contributors to wikipedia*, SIGCAS Comput. Soc. **36** (2006), no. 2.

[62] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morsey, P. van Kleef, S. Auer, and C. Bizer, *Dbpedia - a large-scale, multilingual knowledge base extracted from wikipedia*, Semantic Web Journal (2013), Under review.

[63] D. B. Lenat and R. V. Guha, *Building large knowledge-based systems; representation and inference in the cyc project*, 1st ed., Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.

[64] Sook Lim, *How and why do college students use wikipedia?*, Journal of the American Society for Information Science and Technology **60** (2009), no. 11, 2189–2202.

[65] T. P. Lobo and G. H. Cárcel, *Integrating product information in the pharmaceutical sector*, Ontology Engineering in a Networked World (Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta, and Aldo Gangemi, eds.), Springer Berlin Heidelberg, 2012, pp. 423–434 (English).

[66] M.F. López, A. Gómez-Pérez, and N. Juristo, *Methontology: from ontological art towards ontological engineering*, Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering (Stanford, USA), March 1997, pp. 33–40.

[67] M.F. Lopez, A. Gomez-Perez, J.P. Sierra, and A.P. Sierra, *Building a chemical ontology using methontology and the ontology design environment*, Intelligent Systems and their Applications, IEEE **14** (1999), no. 1, 37–46.

[68] A. Maedche and S. Staab, *Ontology learning for the semantic web*, IEEE Intelligent Systems and Their Applications **16** (2001), no. 2, 72–79, cited By (since 1996)509.

[69] V. Malaisé, P. Zweigenbaum, and B. Bachimont, *Detecting semantic relations between terms in definitions*, COLING 2004 CompuTerm 2004: 3rd International Workshop on Computational Terminology (Geneva, Switzerland) (S. Ananadiou and P. Zweigenbaum, eds.), COLING, August 29 2004, pp. 55–62.

[70] F. Manola and E. Miller, *Rdf primer*, 2004.

[71] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, R. Oltramari, L. Schneider, and I. Horrocks, *Wonderweb deliverable d17. the wonderweb library of foundational ontologies and the dolce ontology*, 2002.

[72] P. Mcnamee, R. Snow, and P. Schone, *Learning named entity hyponyms for question answering*, In Proceedings of the Third International Joint Conference on Natural Language Processing, 2008.

[73] G. A. Miller, *Wordnet: A lexical database for english*, Commun. ACM **38** (1995), no. 11, 39–41.

[74] G. A. Miller and F. Hristea, *Wordnet nouns: Classes and instances*, Computational Linguistics (2006).

[75] S. T. Moturu and H. Liu, *Evaluating the trustworthiness of wikipedia articles through quality and credibility*, WikiSym '09: Proceedings of the 5th International Symposium on Wikis and Open Collaboration (New York, NY, USA), ACM, 2009, pp. 1–2.

[76] S. Neubert, *Model construction in mike (model-based and incremental knowledge engineering).*, Lecture Notes in Computer Science (1993), no. 723, 200.

[77] A. Newell, *The knowledge level*, Artificial Intelligence **18** (1982), no. 1, 87–127.

[78] I. Niles and A. Pease, *Towards a standard upper ontology*, FOIS, 2001, pp. 2–9.

[79] O. Nov, *What motivates wikipedians?*, Commun. ACM **50** (2007), no. 11, 60–64.

[80] N. F. Noy, J. Mortensen, M. A. Musen, and P. R. Alexander, *Mechanical turk as an ontology engineer?: using microtasks as a component of an ontology-engineering workflow*, Proceedings of the 5th Annual ACM Web Science Conference (New York, NY, USA), WebSci '13, ACM, 2013, pp. 262–271.

[81] N. F. Noy and M. A. Musen, *Anchor-prompt: Using non-local context for semantic matching*, In Proceedings of the Workshop On Ontologies and Information Sharing at the International Joint Conference on Artificial Intelligence (Ijcai, 2001, pp. 63–70.

[82] Object Management Group, *Ontology definition metamodel (omg) version 1.0*, Tech. Report formal/2009-05-01, Object Management Group, 5 2009.

[83] R.M. Ortega-Mendoza, L. Villaseñor-Pineda, and M. Montes-y Gómez, *Using lexical patterns for extracting hyponyms from the web*, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **4827 LNAI** (2007), 904–911, cited By (since 1996)2.

[84] K. Panciera, A. Halfaker, and L. Terveen, *Wikipedians are born, not made a study of power editors on wikipedia*, Wikipedians Are Born, Not Made A Study of Power Editors on Wikipedia, Group'09 Proceedings, 2009, pp. 51–60.

[85] P. Pantel and D. Lin, *Discovering word senses from text*, 2002, cited By (since 1996)127, pp. 613–619.

[86] Fernando Silva Parreiras, Gerd Gröner, Tobias Walter, and Steffen Staab, *A model-driven approach for using templates in owl ontologies*, Knowledge Engineering and Management by the Masses (Philipp Cimiano and H.Sofia Pinto, eds.), Lecture Notes in Computer Science, vol. 6317, Springer Berlin Heidelberg, 2010, pp. 350–359 (English).

[87] A. Passant, A. Isaac, and P. Laublet, *Combining linked data and knowledge engineering best practices to design a lightweight role ontology.*, Applied Ontology **6** (2011), no. 3, 223–246.

[88] B. S. Poore, *Wall-e and the 'many, many' maps: Toward user-centred ontologies for the national map.*, Cartographica **45** (2010), no. 2, 113 – 120.

[89] J Preece and D Maloney-Krichmar, *Online communities: Design, theory, and practice.*, Journal of Computer-Mediated Communication **10** (2005), no. 4.

[90] R. Priedhorsky, J. Chen, S. K. Lam, K. Panciera, L. Terveen, and J. Riedl, *Creating, destroying, and restoring value in wikipedia*, Creating, Destroying, and Restoring Value in Wikipedia, Group'07: Proceedings of the 2007 International Acm Conference on Supporting Group Work, 2007, pp. 259–268.

[91] D. A. Randell, Z. Cui, and A. G. Cohn, *A spatial logic based on regions and connection*, Proceedings 3rd International Conference On Knowledge Representation And Reasoning, 1992.

[92] J. Rao and X. Su, *A survey of automated web service composition methods*, In Proceedings of the First International Workshop on Semantic Web Services and Web Process Composition, SWSWPC 2004, 2004, pp. 43–54.

[93] H. Rheingold, *The virtual community: Homesteading on the electronic frontier*, MIT Press, Reading, MA, 1993.

[94] B. A. Rios-Alvarado, I. Lopez-Arevalo, and V. J. Sosa-Sosa, *Learning concept hierarchies from textual resources for ontologies construction*, Expert Systems with Applications **40** (2013), no. 15, 5907 – 5915.

[95] A. Ritter, S. Soderland, and O. Etzioni, *What is this, anyway: Automatic hypernym discovery*, AAAI Spring Symposium: Learning by Reading and Learning to Read, AAAI, 2009, pp. 88–93.

[96] Marta Sabou, Chris Wroe, Carole Goble, and Heiner Stuckenschmidt, *Learning domain ontologies for semantic web service descriptions*, Web Semantics: Science, Services and Agents on the World Wide Web **3** (2005), no. 4, 340 – 365.

[97] E. T. K. Sang, *Extracting hypernym pairs from the web*, Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions (Stroudsburg, PA, USA), ACL '07, Association for Computational Linguistics, 2007, pp. 165–168.

[98] C. Sarasua, E. Simperl, and N. F. Noy, *Crowdmap: Crowdsourcing ontology alignment with microtasks*, The Semantic Web – ISWC 2012 (P. Cudré-Mauroux, J. Heflin, E. Sirin, T. Tudorache, J. Euzenat, M. Hauswirth, J. X. Parreira, J. Hendler, G. Schreiber, A. Bernstein, and E. Blomqvist, eds.), Lecture Notes in Computer Science, vol. 7649, Springer Berlin Heidelberg, 2012, pp. 525–541.

[99] G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. Van de Velde, and B. Wielinga, *Knowledge engineering and management: The commonkads methodology*, MIT Press, 2000.

[100] N. Shadbolt and K. O'Hara, *Linked data in government.*, IEEE Internet Computing **17** (2013), no. 4, 72 – 77.

[101] A. Singhal, *Modern information retrieval: a brief overview*, Bulletin of the IEEE Computer Society Technical Committee on Data Engineering **24** (2001), 2001.

[102] K. Siorpaes, *Ontogame: Weaving the semantic web by online gaming*, In M. Hauswirth, M. Koubarakis and S. Bechhofer, Eds., Proceeedings European Semantic Web Conference, 2008.

[103] R. Snow, D. Jurafsky, and A. Y. Ng, *Learning syntactic patterns for automatic hypernym discovery*, Advances in Neural Information Processing Systems (NIPS 2004), November 2004.

[104] M. Strohmaier, S Walk, J. Pöschko, D. Lamprecht, T. Tudorache, C. Nyulas, M. A. Musen, and N. F. Noy, *How ontologies are made: Studying the hidden social dynamics behind collaborative ontology engineering projects*, Web Semantics: Science, Services and Agents on the World Wide Web **20** (2013), no. 0, 18 – 34.

[105] R. Studer, V. R. Benjamins, and D. Fensel, *Knowledge engineering: Principles and methods*, Data Knowledge Engineering **25** (1998), no. 1-2, 161–197.

[106] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, *The neon methodology for ontology engineering*, Ontology Engineering in a Networked World (Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta, and Aldo Gangemi, eds.), Springer Berlin Heidelberg, 2012, pp. 9–34 (English).

[107] Y. Sure, *On-to-knowledge ontology based knowledge management tools and their application*, KI **16** (2002), no. 1, 35–37.

[108] Y. Sure, M. Erdmann, J. Angele, S. Staab, R. Studer, and D. Wenke, *Ontoedit: Collaborative ontology development for the semantic web*, The Semantic Web — ISWC 2002 (I. Horrocks and J. Hendler, eds.), Lecture Notes in Computer Science, vol. 2342, Springer Berlin Heidelberg, 2002, pp. 221–235 (English).

[109] D. Sutton and V. Patkar, *Commonkads analysis and description of a knowledge based system for the assessment of breast cancer.*, Expert Systems with Applications **36** (2009), no. 2 PART 1, 2411–2423.

[110] T. Tudorache, C. Nyulas, N. F. Noy, and M. A. Musen, *Webprotégé: A collaborative ontology editor and knowledge acquisition tool for the web.*, Semantic Web (1570-0844) **4** (2013), no. 1, 89.

[111] M. Uschold and M. Gruninger, *Ontologies: Principles, methods and applications*, Knowledge Engineering Review **11** (1996), 93–136.

[112] M. Uschold and M. King, *Towards a methodology for building ontologies*, In Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95, 1995.

[113] C. Van Damme, M. Hepp, and K. Siorpaes, *Folksontology: An integrated approach for turning folksonomies into ontologies*, In Proceedings of the ESWC Workshop "Bridging the Gap Between Semantic Web and Web 2.0, Springer, 2007.

[114] B. Villazón-Terrazas, J. Ramírez, M. C. Suárez-Figueroa, and A. Gómez-Pérez, *A network of ontology networks for building e-employment advanced systems.*, Expert Systems with Applications **38** (2011), no. 11, 13612 – 13624.

[115] L. von Ahn and L. Dabbish, *Designing games with a purpose.*, Communications of the Acm **51** (2008), no. 8, 58 – 67.

[116] L. von Ahn, B. Maurer, C. McMillen, D. Abraham, and M. Blum, *reCAPTCHA: Human-Based Character Recognition via Web Security Measures*, Science **321** (2008), no. 5895, 1465–1468.

[117] J. Wales, *Internet encyclopaedias go head to head.*, Nature **438** (2005), no. 7070, 900–901.

[118] D. Xavier, F. Morán, R. Fuentes-Fernández, and G. Pajares, *Modelling knowledge strategy for solving the dna sequence annotation problem through commonkads methodology.*, Expert Systems with Applications **40** (2013), no. 10, 3943–3952.

[119] H-L. Yang and C-Y. Lai, *Motivations of wikipedia content contributors*, Computers in Human Behavior **26** (2010), no. 6, 1377 – 1383, ¡ce:title¿Online Interactivity: Role of Technology in Behavior Change¡/ce:title¿.

[120] L. Zhao and R. Ichise, *Mid-ontology learning from linked data.*, Lecture Notes in Computer Science, vol. 7185 LNCS, Principles of Informatics Research Division, National Institute of Informatics, 2012.