The Antiquarian Librarian & the Pedantic Semantic Web Programmer: Trust, logic, knowledge and inference

Cate Dowd [1]

Abstract. The logic and thinking for the Semantic Web environment has a philosophical base associated with rules and knowledge that draw on traditional concepts, such as 'loci' and 'indexes of place' for the location of items and information. These concepts and other characteristics that led to early library systems and were evident in print publishing, still have some links and relevance for the Semantic Web. For example, location and retrieval principles, once characterised by the 'loci' as a place for an idea, and the headings under which an argument could be found continue to inform the processes of mark-up, indexing and labelling for the Semantic Web.

Web programmers and librarians share important common tools and frameworks. In particular they are guided by sets of rules and logic that are used for the generation of information and knowledge. To generate specific knowledge, in particular for online learning, for example via an online epistemic game, a web developer must carefully design how, when and if information is accessed, which can be achieved via the use of algorithms. To design algorithms for game-play requires an understanding of various types of logic and the construction of values that are assigned to objects that need to translate into meaningful tools for a community of users. The values and decisions reached in the software design process will be informed by the appropriate identification of potential objects; the construction of classes; appropriate assignment of values; choice and selection for distributed data; the construction of ordered lists; labelling and so on.

Early design decisions for potential interaction can also be guided by a thorough understanding of the unique language and meanings that come from the communities of practice who will ultimately exchange information and generate knowledge as they use a system. Design approaches for complex applications, such as a serious online game requires an understanding of the need for multiple rules from which actions and interactions can occur. These rules and outcomes will be based on inference, which can draw on various kinds of logic.

Design features for serious interactive online environments also require structured approaches for understanding the lexical and hermeneutic base of likely end user participant communities. By examining the unique meaning(s) and *semantic* associations of a community's language and communication approaches, and by identifying and classifying potential game-play entities into a representational model, a programmer may gain rich insights to assist the articulation of entity relationships. From there it may be possible to build a formal Ontology which can become a base for an epistemic serious game in the web environment.

Various examples of mark-up languages over time, in particular for online news sites, provide a direct demonstration of the links between constructed language and communities of practice. A light analysis of various mark-up languages and some historic and philosophic perspectives on the construction of know ledge systems and knowledge management may well assist understanding of early software design issues for the Semantic Web. This is also a ripe time to at least raise a discussion about the nature of knowledge, which might emerge from a collective folksonomy, perhaps generated via online news tags for links with social sites, or might be distinct doctrinal or organisational knowledge generated via a formal Ontology, such as that needed for a serious learning environment.

1 INTRODUCTION

Many new forms of 'mark up' developed in the late 20th century based on the Extensible Mark up Language (XML), which changed dramatically the way people began to do business; generate news; and play, especially online and across continents. These applications generally addressed the needs and thinking of various communities of users.

Contemporary mark up languages did not begin with the advent of computing systems; rather they evolved from various non-digital systems, such as library systems which are characterised by indexes, labels and catalogues to assist the location and retrieval of information objects, such as books. The earliest kind of mark-up was a set of instructions within handwritten manuscripts to convey style and layout for the printer. The process of 'annotation was called marking up and the instructions themselves were often referred to as markup"[1]. The annotations related to 'presentation and structure', which are base characteristics of HTML documents, albeit for presentation and display output to a browser, rather than a printer.

Beyond the presentation and structural elements of web documents using HTML, the Semantic Web enables associative meaning to language through syntax, logic and XML server technology. The potential functionality of the Semantic Web depends on logic and information structures that can be understood through analysis of the activities and thinking of the librarian, antiquarian and/or contemporary.

To determine where a book might reside in a library system a librarian refers to a book index and works with a mark-up process to assist that decision. The title alone is too limited. The librarian, as an indexer, will decide the eventual location of the book based on a 'best fit' into a select subject area that matches

[1] Charles Sturt University, Australia, email: cdowd@csu.edu.au

the particular library classification system in use. The alpha-numeric details assigned to a book will make it a unique item within a whole system with potential for cross-referencing using a range of key terms.

It is the relationships and thinking at the cross-referencing level, seen in the library world that opens up the first-order logic associations that area also used in Semantic web applications. At the simplest level some item belongs to some other item. However, thinking and logic for the Semantic web extends to 'reasoning using probability and causality... and the relationships between statements, concepts, or propositions, and rules of inference'[2] .

Via the use of meta names or metadata tags, associations between objects can be made. The descriptive containers, whether for a library system or online news site become the base for building associations between objects that will be ordered according to imposed values and will be processed by levels of logic. An extension of simple associations between objects will build up to form a complex and structured algorithm and indeed artificial intelligence (AI) will emerge as sets of rules lead to knowledge-based outcomes.

In the early 21st century mark up languages express artificial forms of thinking and logic by defining sets of knowledge through entity relationships and corresponding rules. A Semantic Web application will be dependent on many logical features, but above all the design will require an understanding of the language and likely meaning(s) from the communities who will use it, whether the purpose is simply for meaningful search capabilities or more precise vocabularies for shared learning outcomes.

Online games and online news sites depend on precise vocabularies and classifications for objects, as well as rules, so that they can be manipulated by players and readers, or so that information can be shared via syndication sites. The Semantic Web and the exchange of document types now enables new associations and references for multimedia objects which has changed location and retrieval concepts. Indeed these concepts have now shifted towards the fast manipulation of objects, once only seen in stand alone applications.

For many communities the potential of logic via the Semantic web is only just emerging and it coincides with rises in computational speeds and powerful dedicated XML servers which can enable change for organisational process. However, the success of building Semantic Web applications and generating new knowledge requires careful discernments about the complexity of language in the construction of 'knowledge'. For example, multidisciplinary design teams need to come to grips with double hermeneutic situations that can arise from language ambiguity firstly within their own domains and then from a particular community who may use a Semantic Web application.

This paper shows how some linked characteristics of language, logic and thinking have evolved and transformed communication tools at different historic moments into potential knowledge. At the centre of this process is the art and science of indexing, labelling and mark up, and associated characteristics, some of which have remained the same over time. In a trajectory of technological change over time, significant attitudes, methods and tools associated with the location, retrieval and manipulation of objects have contributed to new knowledge and change for individuals and organisations. This paper identifies significant changes in thinking and the emergence of new ideas and structures for communication tools and systems over time, which still hold relevance for the design of knowledge management systems, whether for online news or online distributed serious game play environments.

2 DISTRUST AND THAT PLACE IN THE MIND (LOCI): PRINT TO SEMANTIC WEB

The early transitions from hand written scripts to printed documents during the 15th and 16th were marked by new society attitudes, in particular to the way writing and texts should be approached and used. These approaches contribute to new forms of indexing, mark up and labelling. Ironically, the changes that emerged came out of a mediaeval mindset based on a *mistrust* of the written word. The attitudes and responses to the printed word at the time have some traits that can still be identified in contemporary attitudes towards recent applications, such as the Internet:

> …to perceive writing as more than a reproduction of proximate speech required a leap of faith …they had to convince themselves that texts could be self-contained objects in their own right…and so they started to invent content devices for texts. These devices included signatures, seals, dates, locales, tables of contents, indices, and abstracts. The importance of these devices was to organise ideas according to the logic of the text (and the needs of the reader of the text) [3].

A simple comparison of early attitudes with contemporary approaches to the Semantic Web is most obvious in the areas of 'trust' and 'authentication'. However, the traditional devices also have numerous abstract concepts that still have significance for the Semantic web, such as logical associations between documents, location of documents, signatures and secure systems for transactions, to mention a few. Indeed *trust* and *signatures* not only carry on into the Semantic Web environment with similarities from earlier systems and times but are positioned on the 'top layer of the Semantic Web model'[4] whilst core technologies remain at a lower level.

Computing technologies use 'pointers' of various kinds, often for the location of a particular resource, or for the assignment of some value to an object. In the early print world a pointer may have been to a particular locale, a set of ideas within a written text or to a subject heading. In computing a pointer will point to a resource via a pathway to a particular resource, indeed even by typing a URL (a uniform resource locator) into a web browser.

3 INDEXES, LOCATION & COMMON PLACES

The concepts of location and retrieval have always been closely linked to indexing. To understand the diverse attributes of an index is more complex than first appears. For example, amongst the varies types of indexes used in libraries, even the earliest characteristics of the alphabetic index provide insight into the origins of 'relational' thinking that is still necessary for

the Semantic Web. Important characteristics also stem from the art of rhetoric:

> The alphabetic index is actually a crossroads between auditory and visualist cultures. 'Index' is a shortened form of the original *index locorum* or *index locorum communium,* 'index of places' or 'index of common places'. Rhetoric has provided the various loci or 'places; - headings, we would style them – under which various 'arguments' could be found, headings such as cause, effect, related things, unlike things, and so on.[5]

The idea of the 'argument' from oral cultures is no doubt less precise than most indexing in the 21st century. However, the notion of a 'loci' even if originally "thought of as, vaguely, 'places' in the mind where ideas were stored"[5] translates to an abstract location, even with empty containers, which are well reflected in meta name tags or a class of objects.

In the printed book, in mediaeval times "the vague psychic 'places' became quite physically and visibly localised."[5] The organised spaces and place for categorised texts within books were perhaps the first signs of refined notions for categories, indexical thinking and labelling. They offered structure for the location and retrieval of information and content within a book, which no doubt was more refined than simply pointing to a stack of books.

4 FIRST-ORDER PREDICATE LOGIC IN LIBRARY SYTEMS AND THE RISE OF FORMAL LOGIC

In the English speaking world a uniform approach to cataloguing and indexing emerged in the 19th century with the Decimal Dewey System. The classification system was, and still is, based on decimal numbers and subject areas with additional information about the time of publishing and other details. It was designed for ease of sorting, labelling and retrieving books and documents, and the potential for cross referencing within the system marked the beginning of a structured way in which to define simple sets of knowledge within a system, towards an information system.

The importance of relationships between resources, including categories for subject, date and type of document, brought about an entirely new system that by the late 20th century would be reflected in various digital environments, including the web. In the digital library environment resources would be labelled and sorted by the support of the Resource Description Framework[1], commonly called the RDF, which was integrated other library conventions, such as the Dublin Core, which defines the core descriptive terms for a library resource.

The evolution of logic to assist the location of objects or information artefacts using technological systems are characterised by the ongoing relevance of first-order logic, which is a form of predicate logic that indicates that a 'resource has a *property* and a value for that object'[2]. These features form the base of the RDF and can be interpreted within the library system as 'the book [subject] has the title [predicate] with a specific value [object]'. This formal type of logic has some similarities with other types of logic that have been developed within computing such as the 'Hoare triple, consisting of an assertion, a precondition and postcondition'[2]. The various forms of logic and the formal use of language have increasing relevance as the complexity of a system is revealed. For example a serious online game environment is characterised by many 'pre-conditions and post-conditions' that are essential, simply for game-play.

[1] The Resource Description Framework provides guidelines for specific and universal terms or tags in which to describe a type of content or attribute, such as geographic location of a resource.

5 FROM DISTRUST AND SETTING STANDARDS TO ORDER, LOGIC AND A.I.

By the early 20th century industrialisation had shaped printing, publishing and libraries and generated new attitudes and organisational processes. In particular, 'standardisation' influenced attitudes and process in areas such as 'standardised typefaces and typesetting conventions'[3]. By the 21st century 'the evolutionary appropriation of printing technology led to the construction of a communication artefact with the ... features of standardisation and mass reproducibility – an artefact whose widespread adoption has been associated with such major transformations as the coming of the nation state and the rise of modern science'[6].

Just as mass communication processes and artefacts have been linked to major societal change, the new logic and thinking associated with the Semantic Web has begun to transform our understanding of knowledge and knowledge management, and changed the ways in which we exchange information, interact and learn. However, the traditional tools and methods associated with indexing, labelling, mark-up and logic have increasing importance and continue to be refined in the 21st century, alongside ongoing principles for location and retrieval of objects, but the speed of location and retrieval is mostly invisible.

In addition more sophisticated forms of logic have emerged to achieve deeper levels of intelligence, including artificial intelligence, characterised by 'Higher Order Logic and inference rules'[2]. The relationship between artificial intelligence and knowledge is a large topic, but at a fundamental level it opens up an important topic for early software design in so far as *Tacit* knowledge from a community of potential users of a system must be transformed into *explicit* knowledge before any rules or inferences can be applied, and this can only be achieved once all manner of 'things' have the right labels.

6 [MARKUP] SHIPPING NEWS TO THE WORLD WIDE WEB

The early process of annotation for manuscripts was called *marking up* and the instructions themselves were often referred to as *markup*"[1]. Two early types of mark up for printers included "procedural markup for a typesetter with instructions about how to lay out text, for example, insertion of bold and italic type and different sizes, and descriptive mark-up which would indicate the type of content – for example, emphasis and chapter heading.."[7].

Mark-up for printed material in the 19th century was often very limited or virtually absent, especially for printed newspapers. For example, in Australia, any notion of style for presentation and layout in newspapers was extremely limited and 'until the

1860's…news simply consisted of paragraphs with no attempt at display'[8]. The attitudes and approaches at that time reflected the mindset and culture of the time. They were colonial times with very practical needs to communicate real world information whether for shipping news or parliamentary reports. They were not news stories as we understand them today, and the community of users would have had very different expectations compared to today's newsreaders.

The introduction of pictorial elements into newspapers and the use of mark-up for that end is a story of its own, but the development of mark-up shares some organisational roots that are noteworthy. Online newspapers in the late 20th century are guided by various international standards and protocols set by the telecommunications sector and the W3C[2] organisation, in particular for mark up and conformance, firstly to enable interoperability and to standardise the presentation and layout of documents. Presentation and layout standards in the web environment continue to share core concepts for mark-up with newspaper layout, typically for columns, tables and images.

These characteristics from newspapers, shared in the web environment have cyclical roots in so far as web standards are linked to telecommunications and the "the first newspaper in America…was published by the postmaster"[6]. The expert layout and mark-up knowledge may not have come full swing, but the link with postal and telecommunications concepts continue in the 21st century.

Today, the layout and *presentation* of online digital news formats requires strict vocabularies and rules that allow for far more *flexibility* and *style* than what could have been achieved with earlier printed documents. However, the core principles of layout, indexing, mark-up and labelling continue into the 21st century. Perhaps a small community of users are still interested in shipping news, but even they are most likely to seek information from a different kind of dock, and in real-time. Most of us are no longer waiting for our ship to come in - and we generally don't wait months to read about it, even if we are waiting.

7 CLASSIFICATIONS, METADATA & THE LABELLING OF OBJECTS

In the mid 1990's electronic library cataloguing systems in many countries migrated to web based interfaces. The importance of access to online information about printed resources led to a new set of standards to define the existing data related to collections, subject areas, location and other details for books and journals.

A new standard for 'metadata' called The Dublin Core was introduced as an ongoing set of initiatives from a working group for improved descriptions in the online environment using specific terms, labels and attributes for identification and location of library resources.[9]. The standards have gradually been integrated with other web standards for multimedia objects, for instance a video may be "assigned a DOI (Digital Object Identifier) number so that it can always be found, for intellectual property reasons or if its location on the web changes".[7].

The DOI for a chunk of video is just one example of a label that can be used in the digital environment. From a librarian's perspective the video might also have some 'faceted' associations that are determined by a controlled vocabulary [7], i.e. the object has a relationship to other related objects, which are defined by the metadata name.

The idea of objects having abstract associations is not new as is evident in Ong's discussion of how print was first regarded as an object that soon invited a label:

> Once print had been fairly well interiorised, a book was sensed as a kind of object which 'contained' information, scientific, fictional or other, rather than, as earlier, a recorded utterance…Each individual book in a printed edition was physically the same as another, an identical object, as manuscripts were not….with print, two copies of a given work did not merely say the same thing, they were duplicates of one another as objects.[5]

Steinberg, cited in Ong adds that "The situation invited the use of labels, and the printed book, being a lettered object, naturally took a lettered label, the title page"[5].

Alphabetic classifications or catalogues are visibly evident in the physical library shelves at Trinity College library, Dublin (See figure 1) where the letters of the alphabet can be seen on each shelf level. This simple alphabetic catalogue system represents just one type of cataloguing prior to the introduction of the Dewey Decimal System.

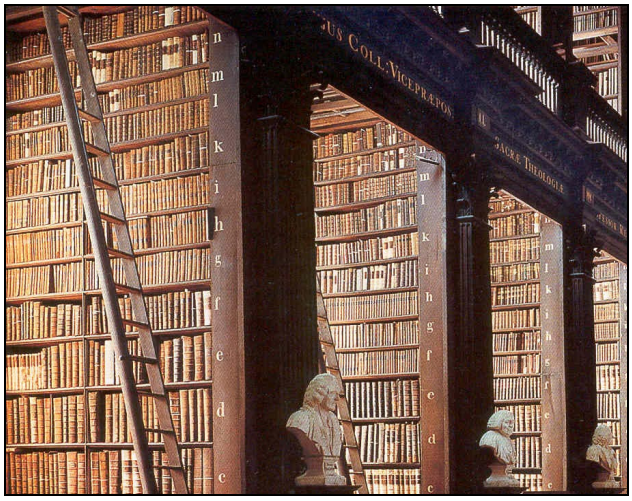

Figure 1. The book shelves of Trinity College Library, Dublin, with obvious signs of an alphabetic classification system.

In the late 20th century mark up in the digital environment was associated with a document type of definition (DTD) of SGML[3] for the basic display of information in a web browser, rather than for any display in a printed book. The type of document was called a HTML document, which was, and still is, just one type of document. Document Type Definitions are not limited to the

[2] W3C is the acronym for the World Wide Web consortium, consisting of various working groups who develop standards for the web to ensure interoperability, accessibility and levels of conformance.

[3] SGML is the acronym for Standardised Generalised Mark up Language, a Mark up language developed prior to the advent of the Internet.

web, for example book publishing has its own DTD[1], but HTML as a document type was limited in what it could do.

8 THE EXTENSIBLE MARKUP LANGUAGE (XML) AND THE RESOURCE DESCRIPTION FRAMEWORK (RDF)

Real-time sharing of information and real-time transactions via the Internet using the XML began in 1996. XML could describe storage layout and logical structures[10], which when combined with XML servers and new organisational processes, would change the ways that people did business, generated news and played online and across continents. XML took little time to trigger new perceptions of communication and the closing of distances. It also raised many issues associated with knowledge, such as whether information was from an authentic source and could be trusted.

The relationship between library systems and the introduction of XML should not be understated. The Dublin Core, mentioned earlier, as a base digital system for classifications complimented by the Resource Description Framework (RDF) for integration of library resources not only provided new levels of access and exchange of digital data for libraries, but has become the framework for other innovations.

The RDF is a framework for 'resources' and is concerned with logical associations between resources, which might indeed be objects. It has become 'a model of statements...about resources, and associated URIs [Uniform Resource Identifiers] that have a uniform structure of three parts: subject, predicate and object." [2] The logic and structured associations based on syntax and semantics associated with the RDF are the base of the Semantic Web. Whilst the RDF framework might be considered relatively simple it has relevance for the construction of a Semantic Web Ontology[4] which is necessary for the artificial intelligence of a complex online game.

For the modern, rather than the antiquarian librarian, the RDF is also still significant because it has become a model for the "expression of semantic information (meaning) [and can] assist interoperability of data, computer-understandable semantics for metadata, and better precision in resource discovery than can be achieved with full text search."[7].

9 INTEGRATED XML APPLICATIONS

Metadata standards in the early 21st century are made up of many works in progress, partly due to the opportunity for developers to build their own language or tags using the XML standards. An important multimedia XML-based mark-up language that has been in use since the late 1990s is SMIL, or the Synchronised Multimedia Integration Language, which is for 'integrating streaming audio and video with images and text for the web"[7]. The SMIL descriptions for integrated media forms mark an important shift from mark up that simply displayed static objects in a web browser. These descriptions changed the concept of mark-up for *presentation* and display within a web browser to a range of new ideas about the manipulation of data as objects.

Explorations to also describe even higher level detail for video content as new faceted objects have been in progress for several years. In particular via the standards associated with the Moving Pictures Expert Group (MPEG), such as MPEG-21 and MPEG-7, which are concerned with segmentation classifications for video attributes. Some levels of research have also steered towards "multi-level video indexing approaches using the RDF to contain both Dublin Core and MPEG-7 descriptions of the same content"[7]. These more sophisticated levels of integrated research can be applied to interactive video projects and digital television, but they also use levels of indexing and labelling that require serious modelling due to complexity and the need for representation across the design process, and they raise various issues for implementation beyond the scope of this paper.

Another example of integrated web technologies that epitomises the use of mark up, indexing and labelling, even in name, is Annodex. Annodex is a 'file format for annotating (indexing) time–continuous bit streams so that people can use text queries to search for video clips, then hyperlink to other video, audio or web content….it [uses] the Continuous Media Markup Language (CMML)' [7]. The idea of continuous mark-up perhaps captures the essence of XML technology, which seems limitless.

The various examples of mark-up clearly indicate how XML applications have drawn heavily on the core concepts of library systems, such as indexing, and various formal logic. The guidelines for implementing the Dublin Core in XML [11] are openly available to developers and might suggest that the range of mark-up applications will continue to grow and change the knowledge bases of today. The range of XML applications in recent years that are already pervasive include the use of 'SportsML and NewsML, and [perhaps less pervasive] Bookmark Exchange language (XBEL)"[12]. The use of XML in online news is perhaps the most dominant.

10 XML SCHEMAS AND THE INTEGRATION OF THE DUBLIN CORE FOR ONLINE NEWS

The RDF as an XML application makes it possible to define sets of knowledge through entity relationships, using some higher level logic. Contemporary mediated communication forms that have integrated these standards have been able to build new associative entity relationships, using shared lists and various indexes. In particular these can be seen within online news publishing and in syndicated news stories.

Integrated XML applications, such as the integration of the Dublin Core and XML schemas[5], can be easily identified in the source code of an online news site. In particular the links between document types are evident. For example, the source code from the Australian Broadcasting Corporation (ABC) news website[13] (see figure 2) shows a relative link to the Dublin Core tags, represented by the repeated acronym 'DC', followed by a metadata category name for each title and descriptive category. Other specific terms of reference for document types defined by the ABC are also embedded in the code.

[4] A Semantic Web Ontology is more sophisticated than a set of classes and objects because it also has a set of inference rules.

[5] Schemas are essentially different classes of documents with elements and attributes that conform to a document type.

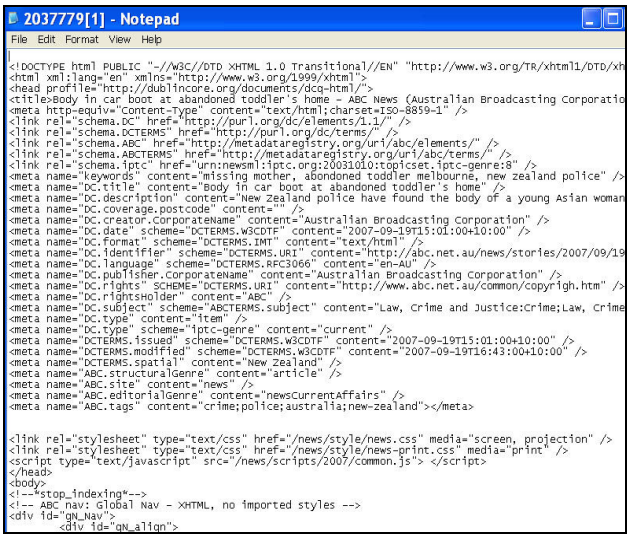

Figure 2. The source code from an ABC web site showing the Dublin Core terms and ABC terms to enable a resource, such as an article in the "news" category to be displayed in a Web browser. Source: ABC.

11 REALLY SIMPLE SYNDICATION (RSS)

The syndication of news stories is achieved using "RSS (Rich Site Summary) technology with metadata tags, which is XML based"[2]. RSS has actually evolved as three different versions and is more commonly referred to as Really Simple Syndication, or even news feeds, because it enables regular 'feeds' into an existing web page if a user chooses to make the appropriate links.

Digital 'tags' created by individuals or a particular community are the 'mark up labels' of the digital environment that have already generated some social change for mediated communications. The creation of shared tags is sometimes referred to as a form of "collective indexing"[7] and has facilitated the growth of Social Sites. Rather than using a formal taxonomy or even an ontology to define the nature of shared documents, social sites simply link to each other by agreement between folks, hence social sites are referred to in terms of "collaborative tagging and folksonomies" [7].

Current examples of social sites include Delicious, Flickr and Digg which provide digital feeds for photographic resources and news stories from a range of news sources, both independent and mainstream.

12 ONLINE EPISTEMIC GAMES: THE NEED FOR RULES, ALGORITHMS AND MODELS

An online epistemic game is generally designed as an immersive world for learning. The participant players do not simply scan web sites in search of higher levels of knowledge or undergoing traditional instructions from an avatar. Rather, they are situated in a constructed and simulated environment in which they learn about a professional role through engagement and immersion. Shaffer refers to this kind of learning or knowledge building as 'epistemic frames; collections of skills, knowledge, identities, values and epistemology that professionals use to think in innovative ways'[14].

The design of an epistemic game in the web environment requires deep knowledge of the participant communities and future players. In the early design phase it is crucial to represent and model the core areas for potential game play, which will slowly be shaped by the specific language and semantics of a community. It is also critical to understand the relationship between game play and rules, but also to make distinctions between doctrinal rules and inference rules which follow the limited logic of potential algorithms that might be developed. They are not the same thing. Nonetheless, in game-play 'playing a role means following some set of rules for behaviour'[14].

Online multiplayer games and simulations use various forms of mark up and scripting languages, for example Second Life as a virtual reality (VR) environment uses the Linden Scrip. This type of script defines many complex entities, including graphics and will process many computations to enable action. It draws on various forms of mark up, including XML and uses multiple forms of indexing, identification schemes and labels to ultimately enable the manipulation of many different objects, and at fast speeds.

A multiplayer environment can be modelled by 'classes' of objects, which are indeed the base of object orientated programming, but XML integrated applications will use schemas for classifying objects. There are many ways to begin to represent and communicate the entities of a future digital game play environment, such as using software tools for the Ontology Web Language (OWL)[6] or lexical based tools to help sort text and document types, like Leximancer[7]. However, the Universal Modelling Language (UML) (see figure 3) still provides an established base to begin building classes of objects and sub-classes that can represent and communicate core entity relationships and other features such as sequences of events.

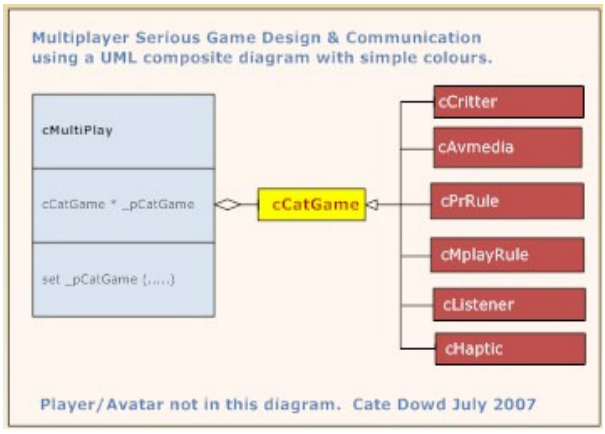

Figure 3. A UML diagram for a Serious Game design indicating potential classes of objects and a composite subclass with method and function.

[6] OWL is a recommended standard for processing content and information based on semantic meaning, that began in 2004 via the W3C http://www.w3.org/TR/owl-features/

[7] Leximancer is a software tool to assist text mining and is located at http://www.leximancer.com/cms/

The coded language, computations and processing in an online game can generate outcomes from player interactions via algorithms that will open up new choices for a player. The construction of multiple algorithms within a knowledge system suggests it is likely to be a complex system. However, for a system to achieve any levels of social intelligence, especially as a simulation, multidisciplinary design teams will need to grasp a large amount of information, which is perhaps best mapped via the use of formal models.

Models for representations of reality need to be simple and be able to communicate ideas during the design process. Object orientated analysis, towards design, aims to help 'represent ….real-world problem [s] in a format which a computer finds easy to deal with'[15]. As knowledge representations are developed via models they will inevitably begin to be embedded with various forms of logic that are necessary for an interactive system capable of actually managing knowledge. Ultimately the applied logic should assist developmental learning for a community of users if the system is designed as a serious game environment.

In addition to carefully planned learning outcomes, an online game-player will often experiences some sense of direct manipulation, which might be called an 'experience of agency'[16] or immersion. This could arise from navigation through a visual space with an Avatar or come from the effects of an explosion. In essence a script can trigger a momentary position and visual-auditory dimensions and associations between objects, which can be referred to as a pre-condition, which with a certain command, can generate a post-condition. The immersive experience is based on the mix of logic, graphic realism, and the rewards embedded in the design. From a design perspective, the game will have imposed values that aim to generate a learning outcome, in particular the design of a serious game. Clearly this will be a different shared experience than what might be generated collectively in a massive-multiplayer online game environment.

The hierarchical structures of the Semantic Web and the use of inference rules can also apply to a network of addresses drilling down to *locate* an individual machine responsible for particular computations in a distributed system, which may indeed be necessary for an online serious game shared by large professional communities.

13 CONCLUSION AND REFLECTIONS

The online game-player experience or the use of XML based RSS tags to allow a news reader access to syndicated online news stories, and even collaborative tagging, depend on multiple systems rich with indexes and labels. Some of the oldest principles of mark up, location and retrieval have clear pathways from print publishing and library systems into computing and information systems. They highlight the relevance of earlier efforts and the need for refinements, in particular for Semantic Web applications. The Resource Description Framework has formed an important base for higher level semantics alongside the use of multiple rules for inference and various forms of logic that will continue to be used for knowledge management purposes.

Information and knowledge management over time is clearly embedded in philosophical issues, in which human beings as creative developers must make absolute decisions about access and the construction of conditional situations for others. These decisions and situations are based on individual and collective value systems which can often be in conflict.

A socially intelligent computing system in the Semantic Web environment will be partly shaped by the applied logic, mark up, indexing and labelling associated with that system. Some of the characteristics of information and knowledge management systems discussed in this paper in various ways have been more visible in earlier information and knowledge systems and are now perhaps lost within the 'abstractions of a complex device such as a computer… [which many people still prefer to see] as a single, comprehensible unit'[17].

Perhaps more than ever the design and development of complex systems must look to deeper understanding and interpretation of content, in particular to the coded language and nuances of future communities of users, simply in order to get the game design right. This could be partly achieved through a formal focus on interpretation and the identification of 'hermeneutic' patterns across domains.

The location and retrieval of objects from within a socially intelligent system, digital or otherwise, is most likely to work well where objects or agents having unambiguous labels; are classified and grouped into appropriate indexes; conform to mark-up; and are subject to appropriate rules and various forms of applied logic.

The design of online news and more complex online systems, such a multiplayer game environment is a pedantic semantic situation for a programmer and a design team! It might be so, for several reasons, but it is no doubt partly due to the ambiguity of terms found across disciplines and domains and the need for very precise vocabularies in the Semantic Web environment. The tags and algorithms of the new artificial intelligence, like the catalogue labels, seals, signatures and stacks of the old, are still entwined in the topics of trust, distrust, logic, knowledge, rules and inference.

ACKNOWLEDGEMENTS This research is supported under the Australian Research Council's *Linkage Projects* funding scheme (project number LP0775418)

REFERENCES

[1] Chapman, N. and J. Chapman, *Digital Multimedia*. 2000: Wiley, New York.

[2] Alesso, P. and C.F. Smith, *Thinking on the Web: Berners-Lee, Gödel, and Turing*. 2006: John Wiley & Sons, Inc, New Jersey.

[3] Kaufer, D. and K. Carley, *Communication at a Distance: The influence of Print on Sociological Organisation and Change*. 1993: Lawrence Erlbaum Associates, Hillsdale, New Jersey.

[4] Antoniou, G. and F.V. Harmelen, *A Semantic Web Primer*. 2004: MIT Press, London.

[5] Ong, W., *Orality and Literacy* 1982: Routledge, London.

[6] Boczkowski, P., *Digitising the News: Innovation in Online Newspapers*. 2004: MIT Press, Massuchesetts.

[7] Brown, G. and J. Jermey, *The Indexing Companion* 2007 Cambridge University Press, Melbourne.

[8] Mayer, H., *The Press in Australia*. 1964: Landsdowne Press Pty Ltd, Melbourne.

[9] DCMI, U.B. *DCMI Metadata Terms*. 2006 [cited 24th September 2007]; Available from:http://dublincore.org

[10] W3C. *Extensible Markup Language (XML) 1.0 (Fourth Edition)*. 2006 [cited 22nd September 2007]; Available from:http://www.w3.org/TR/2006/REC-xml-20060816/#sec-origin-goals.

[11] Powell, A. and P. Johnston. *Guidelines for implementing Dublin Core in XML*. 2003 [cited 24th Sept 2007; Available from: http://dublincore.org/documents/dc-xml-guidelines

[12] XML.org. *Standards, Specifications and Initiatives* 2006 [cited 24th September]; Available from:http://publishing.xml.org/standards

[13] ABC. *ABC Online*. 2007 [cited 19th September 2007]; Available from:http://www.abc.net.au/.

[14] Shaffer, D.W., *How Computer Games Help Children Learn*. 2006: Palgrave Macmillan, New York.

[15] Rucker, R., *Software Engineering and Computer Games*. 2003: Pearson Education Limited, Essex UK.

[16] Murray, J., *From Game-Story to Cyberdrama*, in *First Person: New Media as Story, Performance and Game*, N. Wardrip-Fruin and P. Harrigan, Editors. 2004, Massachusetts Institute of Technology: Massachusetts.

[17] Brookshear, G.J., *Computer Science: An Overview*. 9th ed. 2007: Pearson Education, New York.