www.ichim.org

Les institutions culturelles et le numérique
Cultural institutions and digital technology

École du Louvre
8 - 12 septembre 2003

CULTURAL HERITAGE FORUM DESIGN

Eoin Kilfeather, John McAuley
Digital Media Centre, Dublin Institute of Technology, Ireland

Anthony Corns, Oli McHugh
Discovery Programme, Dublin, Ireland

« Acte publié avec le soutien de la Mission de la Recherche et de la Technologie du Ministère de la Culture et de la Communication »

Abstract

The CIPHER project was set up in April 2002 as a thirty month project supported by the European Union. The project aims to give the public new ways of accessing cultural heritage information from around Europe using new technologies. These technologies are intended to help organise knowledge and cultural heritage narratives. The technology helps to present these narratives on the internet and provides editing, communication and discussion tools. Visitors are able to customise and add to these narratives using tools provided by the partners.

One of the principle activities of the project is the establishment of Cultural Heritage (CH) Forums. A cultural heritage forum is a portal website, in some cases associated with an actual place and concerned with some aspect of shared cultural heritage. A CH forum allows all those interested in its content to explore, learn, interact with and add information to the site. The CIPHER partners have developed technology, tools and methods for making successful and sustainable forums. These tools and technologies are showcased using a number of specific forums.

This paper is concerned with the development of the Irish Cultural Heritage Forum – explorer. This forum was developed by The Dublin Institute of Technology and The Discovery Programme and gives online access to a database of Irish archaeology held by the Discovery Programme. This includes the complete National Monuments Record (NMR). The NMR is a list of categorised and detailed monument data for Ireland and Northern Ireland showing, for example, a monument's geographical position. Importantly the forum provides access to a large set of 'lessons' explaining and putting into context these monuments and artifacts. These lessons are primarily short texts which explain some aspect of Irish archaeology.

The CH forum uses knowledge management techniques developed by the CIPHER partners including a specially adapted ontology of archaeological objects and concepts. The prototype application allows authors and ordinary users of the site to relate their own lessons to this ontology and in this way to construct a broad ranging but structured description of the domain of Irish archaeology.

Keywords: Cultural Heritage Forum, Ontology, Narrative.

1. Introduction

The Irish cultural heritage forum explorer.ie is one of four cultural heritage forums being developed as part of the CIPHER project. Initially the project set out some approaches which sought to differentiate the CIPHER cultural heritage forums from other cultural portal approaches. Amongst these approaches was the use of customized annotation tools to support the construction of personal or shared spaces and the use of those spaces to share narratives about cultural heritage. The explorer.ie website is an implementation of a content management system which incorporates ontological tools to help in narrative construction and discovery. The software allows the user to annotate existing resources (primarily catalogue data) with short narratives and to reuse pre-written elements in their own narratives.

2. Forum Users

The explorer.ie forum was firstly designed for domain experts to begin the process of seeding the forum with archaeological narratives. In this model the forum is seen as having a number of evolutionary changes in its mode of use. As with most portal models there is a strong requirement for a body of well designed initial content in order to encourage wider user participation. The forum can be seen to have the following steps in it evolution;

1. The forum is designed and the tools tested by the developers.
2. The forum is seeded with narratives by domain experts
3. The forum is opened to the successively wider user groups
4. The process repeats and the tools and methods used are refined.

As the CIPHER project is part of the EU “Heritage for All” programme and much of the information held was being directly, or indirectly funded by the general public, it was decided that the Explorer CH Forum should be aimed at as large as possible a spectrum of user groups. Three major user groups were identified as potential users of Explorer, with a combination of their previous knowledge and age group being the determining factor to which group a user would belong. A brief description of each user group and possible examples is summarized in Table 1.

User Category	Description of Users	Example of Users
Basic	Users have no to little previous knowledge of cultural heritage domain Users are 12 years old or younger Non computer experts	Primary school children Users with learning difficulties
Intermediate	Users have no to little previous knowledge of cultural heritage domain Users are over the age of 12 Non computer experts	Secondary school children Adult users with little or no experience of cultural heritage General members of the public
Expert	Users have differing levels of cultural heritage domain knowledge Previous computer knowledge	University and college students Members of special interest groups Professional users e.g. cultural researchers, national organizations

Table 1. Summary of potential Explorer user groups

The initial tools needed to develop the forum and allow domain experts to compose content were not limited to a simple content creation or management system. There are many technologies and methods available for managing web content and these technologies are fast maturing [1].

For the Irish CH forum what was needed were tools which would facilitate a meaningful and coherent story or narrative structure to be developed from existing or newly created content. Newly created content had to be facilitated through an authoring system while existing content had to be searchable using semantically aware methods. The overall aim

was to allow authors to establish semantic relations between narrative resources and allow the author to select and meld together those narrative resources. By allowing authors to transparently browse and link these resources the overall domain description (of Irish archaeology, history and folklore) is reinforced.

3. Domain content

In the case of the explorer forum there already existed a body of catalogued archaeological data which could be built upon and a narrative layer formed to help develop a descriptive explanation of that data. Some of this catalogue data forms the collective digital description of the monuments and finds which have been noted and studied over the past century in Ireland. These catalogues are still in the process of being updated and a significant proportion of the data is poorly or loosely structured. Some monuments and artefacts are unclassified and exist in the catalogue only as a reference. Nevertheless this data was available to the project and formed the backbone of the resources upon which the narrative layer was built. Amongst the catalogues used in the pilot project were;

3.1. The National Monument Record (NMR)

The NMR (National Monuments Record) is a complete list of categorised monument data of archaeological or historical interest for Ireland and Northern Ireland showing geographical positions. Each monument (123,504 in total) has a classcode which identifies it as a specific monument type. In total there are 786 classcodes in the Irish classification system. These classes constitute an evolved vocabulary which has developed from a number of domains including, architecture, archaeology and folklore. As such it does not have a hierarchical basis and the classes themselves have no formal relationships.

3.2. North Munster Project Database

The North Munster project is a ten year study carried out by the Discovery Programme into Bronze Age settlement in the north of Munster province. The study produced a large quantity of data and a significant number of finds of archaeological objects. These objects

were catalogued in a database but the catalogue did not use a controlled vocabulary. Nevertheless the objects were well described and each had a unique identifier.

3.3. Media database

In addition to the NMR and North Munster databases the Discovery Programme holds a large number of images of archaeological finds, monuments. These images together with 3D data from field surveys and video files of archaeological reconstructions constituted the media database for the forum. The content management system had to be able to receive media files in a variety of formats and to automatically adopt the correct presentation template for a given media type. Each of these databases are held using a standard open-source database management system (MySQL).

4. Standardising the domain vocabulary

One of the key difficulties facing the designers when in opening up remote access to the Irish NMR databases and linking resource discovery was the lack of a properly structured thesaurus of terms for the Irish monuments and archaeological objects. In discussions with the domain experts it was decided to use the forum in a pilot project to explore a manageable section of the Irish CH database and to map these data to a domain ontology. The next problem was to decide whether to design a new ontology of Irish archaeology or to adapt an existing ontology.

The CIPHER project had from the outset committed to adopting and promoting cultural heritage standards. For the standpoint of the explorer forum there had been few standardization efforts in Ireland itself. However there were a number of CH standards which had been devised and adopted by bodies in the UK. The Discovery Programme has been a member of a standards body called the Forum for Information Standards in Heritage (FISH).

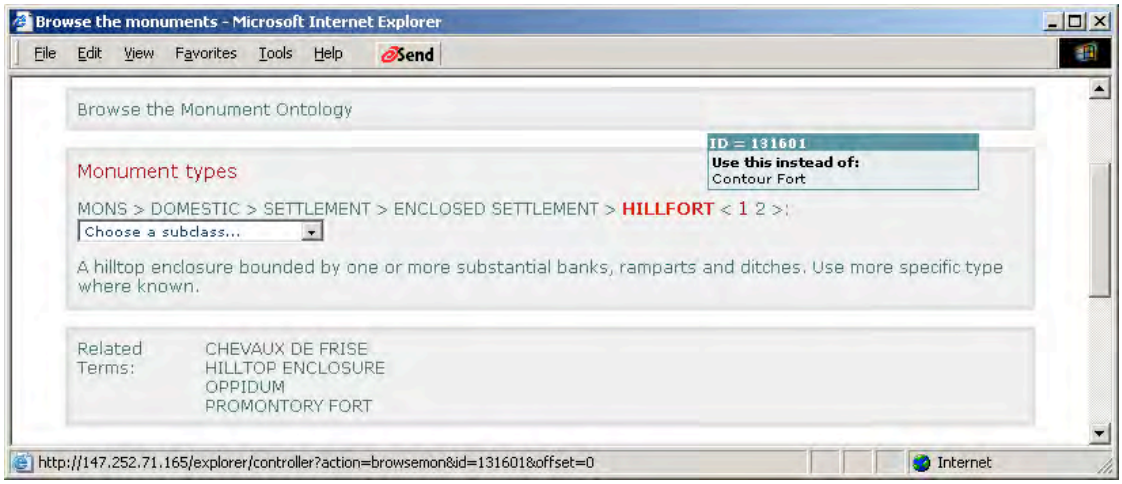
The FISH group has cooperated in the effort to produce guidelines for catalogue standards in the heritage for the UK and Ireland [3]. Amongst these efforts is the MIDAS 'Monument Inventory Data Standard'. Specifically the MIDAS guidelines provided the

basis for a mapping of the information in the Irish NMR to the conceptual model of the much broader English Heritage CH thesaurus. This effort had successfully been undertaken in by other national and regional heritage bodies, such as Scottish Heritage [4] The adapted and completed ontologies were again exported to a standard RDBMS environment (MySQL) to allow for easier data management and Internet implementation.

4.1. English Heritage National Monument Thesauri

The EH thesaurus provided a structured wordlist and a standardised terminology [5]. It includes the concept of allowed terms, in other words a controlled vocabulary. As was mentioned above some of the Irish catalogue data had not adhered to a controlled vocabulary for description of (especially) archaeological objects. The thesaurus also relates monuments and archaeological objects by similarity. The terms are grouped into hierarchies (a hierarchical relationship) and cross-referenced to other groups of terms which may be relevant to the domain (an associative relationship). Finally the EH thesauri also provide single preferred terms where a synonym occurs (an equivalence relationship). The organization of the thesaurus allows these terms to be selected at a general or specific level, depending on context. Finally the thesaurus was chosen as it is a dynamic tool, which was from the outset intended to be developed by the addition, amendment and deletion of terms, relationships or hierarchies as was required for the adaptation to the Irish NMR.

To adapt the Irish NMR a domain expert (Oli McHugh) examined each of the classcode terms in the Irish catalogue. There are a total of 786 classcodes in the Irish classification system. In order to use the English Heritage (EH) thesaurus of monument types, a matching process was undertaken to relate the Irish monument types to terms within the EH ontology. The Irish classification system represents a resource of terminology which [2] describes the built heritage of Ireland. Matching these terms with the EH ontology thesaurus provided an interrelated tool for use on explorer.ie In each case a determination was made as to the type of relationship which the class had with the EH thesaurus.



A screenshot showing the monument type browser.

4.2. The Equivalence relationship

A term can be ‘preferred’ or ‘non-preferred’; however, it is the preferred term that will be used in the hierarchies (as a visible class) and used for indexing. There were a number of culturally problematic decisions to be taken in deciding the correct approach in determining whether a term in the Irish vocabulary was to be used in ’preference’ or not. Clearly where domain experts in a given country predominantly use one term the ‘preference’ term should be the ‘Regional Term’ in the context of a borrowed ontology. This reversal of preference was one of the steps taken in the term mapping process done for explorer.ie. This was the case for example where the Irish term Fulacht fiadh took preference in the equivalence relationship to the English Heritage term BURNT MOUND

4.3. The Hierarchical relationship

The second stage was to group the terms into hierarchies. They were first gathered into conceptual groups, for example, all monuments that are settlements. Then within each conceptual group the terms are further divided into levels going from the broadest type of term to the narrowest and most specific type of term.

For example:

- SETTLEMENT Conceptual group
- ENCLOSED SETTLEMENT Broadest level or BROADER TERM
- HILLFORT Narrowest level or NARROWER TERM

The EH thesaurus is also poly -hierarchical; that is to say, the broad term may appear in more than one hierarchy and under more than one class. For example: HILLFORT occurs under both the DOMESTIC and DEFENCE Conceptual groups.

4.4. The Associative relationship

Terms can be associated with each other but not necessarily connected by a hierarchy. This means that two terms similar in concept can be associated even though they come under different broad terms. These are referred to as 'related terms'. Such terms are used as an aid to help user find terms similar to the initial term but which are not always immediately obvious. In the initial pilot project the associative terms were left unchanged from the EH ontology but it is anticipated that as the system is used by Irish domain experts terms will present themselves for association.

4.5. Scope Notes

The scope note indicates exactly how the term is to be used in the context of the ontology. That is, it provides a definition and any point that should be taken into consideration for the use of the term. As was mentioned above in many cases the Irish term used in the NMR was ambiguous and no formal scope note existed. In many cases this has required examining specific monument instances to determine the correct mapping.
For example: HILLFORT SN A hilltop enclosure bounded by one or more substantial banks, ramparts and ditches. From the definition it is clear that this is a monument type that can be used to index hillforts of all periods, not just bronze-age hillforts.

5. Modifying the Ontology Structure

In order to use the English Heritage (EH) thesaurus of monument types, a matching process had to be undertaken for each Irish monument types and archaeological object type to find where possible a matching terms within the EH ontology. The Irish classification system itself represents a resource of terminology which describes the built heritage unique to Ireland.

The mapping process therefore produced the following results. Of the 786 classes in the Irish vocabulary 472 (60%) had a direct match to a term in the EH thesaurus. In this case the term was directly mapped. However 224 (28%) terms were closely related to terms in the EH ontology but were not linguistically similar enough to provide a direct mapping. In these cases the terms were mapped as preference terms, for example 'Ringfort'. A further 101 terms (13%) are terms used principally in the context of Irish folklore or archaeology and did not have a match with any English heritage terms. These terms are viewed as candidates for new classes within the ontology. Examples of these terms are 'Fian-bhoth' and 'Baulk'.

Other domain specific differences occurred where a terms in the EH and Irish thesauri were very similar but semantically dissimilar such as the term Sheila-na-gig "A small carved figure, usually female in appearance, probably representing fertility charms, found on Romanesque churches in the West of England". Is not of course the same as the Irish term 'Sheela-na-gig'. These differences were flagged in the knowledge base.

6. Building a narrative layer

In building a tool to help domain experts to illustrate and explorer the Irish archaeological record a number of guiding criteria were introduced. The general aim was to facilitate in as simple a manner as possible the use of the system by non-computer experts. The first principle was that the system should be entirely web driven with no proprietary software required so that even at prototype stage the system could be used by as wide a user base as possible.

Secondly the system had to encourage users to make use of the ontological structure of the monument catalogue and to use that structure in navigating the narratives. Thirdly the system should provide as seamless a possible transition between the browsing and authoring modes of use. To facilitate this architecture and to keep the application as flexible as possible it was decided to implement the system as a tiered Java Enterprise web application.

6.1. Narrative components

The concept of dynamic narratives is an evolving one and clearly these systems are at an experimental stage however a number of general principles for supporting dynamic narratives have been suggested [2]. In developing the explorer system a decision was made to introduce a basic narrative unit which was called a 'lesson'. From a user's standpoint the lesson is a piece of text which relates to some aspect of cultural heritage. Lessons are stored in a database for retrieval or later editing.

The broader unit of narrative in the explorer.ie system is called a 'theme'. Conceptually a theme is a linked collection of lessons. These concepts were developed in consultation with domain experts. After initial consultation these experts indicated that they felt a dynamic narrative system which attempted to match lessons solely through an ontological association was too restrictive. The concept of the themed lesson was an attempt to explicitly relate lessons through authorial choice.

6.2. Lesson creation

When a user has authenticated the system will allow him / her to create a lesson. The lesson creation tool is a Java applet with html editing capabilities which has been modified from Open Source. The author enters a title into the form and some text. At this point the author is encouraged to highlight terms in the text which he / she considers to be key terms. Additionally the user may enter terms which do not appear in the text but which the author considers to be relevant to the text. Upon submitting the lesson and key terms to the system a number of steps follow.

Firstly the terms highlighted are processed through a Porter Stemmer algorithm to remove 'stop words'. So for example the term 'hillforts' will produce a match for 'hillfort' (singular). This method helps to improve the chances of finding a match for the keyword in the monument ontology. When the lesson has been submitted the system then attempts to find a match for any of the keywords in the lesson with, firstly, the monuments ontology. So in the above example the term hillfort will find a match within the ontology. These terms are flagged as 'recommended' terms. Clearly this recommendation is based on a textual and not a semantic match but the term is presented to the author together with the term scope note so that he / she can make a determination of the term's suitability.

Where a term is chosen by the author it is shown within its hierarchy and the author can then either chose to associate the lesson with this node of the hierarchy or chose to browse through the hierarchy looking for other terms. The author may chose to associate his / her lesson with as many monument classes as he / she wishes. These monument associations are held in a database table for retrieval at a later point. Furthermore where the term exists the system will examine the Irish NMR for examples within the catalogue of examples of that monument class. In this way the author can chose an example or examples of specific monuments to associate with his lesson. This was judged to be desirable as there are instances where a lesson may be specific to (to continue the above example) a single hillfort. This process continues until the author has finished associating monument classes and instances with the lesson. The author is again asked to highlight terms but at this time to associate the terms with the archaeological objects thesaurus. This process is identical to the process followed in associating monument types and instances but in this case the archaeological object finds database is referenced.

Finally the author is invited to upload a piece of media to associate with the lesson. Media can be a wide range of types from images to interactive 3D models. Typically however the media is an image and the system firstly creates a thumbnail version of the image and stores it together with the full size image in the system. The author is then prompted whether the image is to be associated ontologically with the lesson's (already) related classes or instances. This stage is to encourage the possibility of media object reuse in another context possibly by another author. The author is given the opportunity to edit their lesson and ontological associations before committing the lesson to the system for storage in the database. The author is then asked if he / she wishes to create another lesson and this process continues until the author has completed. The authored lessons are then available for the creation of a so called themed lesson. Conceptually the themed lesson is a sequence of lessons to be presented as a whole. The author is given a tool which allows the sequence to be rearranged and for lessons to be added or deleted. These sequences are then stored by the system.

6.3. Lesson browsing

Lessons can be browsed by users of the system in a number of ways. The first and most powerful method is through themed lessons. The themed lesson represents a domain

expert's presentation of some area of archaeological heritage. The user is given a structured narrative introduction to the area with little ambiguity. However the themed lesson also allows access to the wider ontological structuring of the domain and the narrative lesson components associated with particular classes. The user is given a tool to browse the thesaurus terms and on the surface level view scope notes which in themselves constitute an explanatory layer. However where a lesson or lessons have been associated with a term the user is presented with the opportunity to view these lessons. In this way even when browsing a themed lesson the user is exposed to alternate lessons touching on the same concepts.

As the level of content increases within the system the number of cross referenced narratives increases. This applies both at the level of content browsing and authoring. As lessons are created they become available for inclusion in other themed narratives. In initial trials it has been seen that authors quickly realise the opportunities this tool presents.

7. Conclusion

In this paper, we have presented a novel method for creating and sharing multimedia cultural heritage narratives. The CIPHER project and the explorer.ie forum seeks to demonstrate ways to represent cultural heritage information in hypermedia environments and to present that information in a coherent manner. In order to fulfill this requirement, a ontologically structured domain model has been adapted from existing standards and implemented in a tiered web application. (A version of the web application software can be found at http://www.explorer.ie/ and the source code is freely available)

References

[1] Rosenfeld, L, Morville, P. (2002) Information Architecture for the World Wide Web, 2nd Edition. O'Reilly

[2] Mullholland, P., Collins, T. (2002) "Using Digital Narratives to Support the Collaborative Learning and Exploration of Cultural Heritage" in International Workshop on Presenting and Exploring Heritage on the Web, DEXA 2002. IEEE CS Press PR01668

[3] FISH Working group (2002) 'MIDAS A Manual and Data Standard for Monument Inventories'. English Heritage Press XB20005

[4] Baines, A.(2001) Scottish Thesaurus of Monument Types: final report. 2001. RCAHMS Press. Edinburgh

[5] Charles, D. (2001) "In The Beginning Was The Word: Making English Heritage Thesauri Available On-line", in Cultivate Interactive, issue 3, 29 January 2001

[6] Corns, A. et alia.(2003) "Internet cultural portals and the development of an Irish Cultural Heritage Forum". IN CAA Enter the Past Conference Proceedings, Vienna April 2003.

[7] Kilfeather, E. et alia. (2003) "An ontological application for archaeological narratives" in DEXA 2003 Proceedings, Prague. IEEE.