2010

Virtual museums and web-based digital ecosystems

Peter Eklund

University of Wollongong, peklund@uow.edu.au

Peter Goodall

pgoodall@uow.edu.au

Timothy Wray

University of Wollongong, twray@uow.edu.au

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>

Part of the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Eklund, Peter; Goodall, Peter; and Wray, Timothy: Virtual museums and web-based digital ecosystems
2010.

<https://ro.uow.edu.au/infopapers/3520>

Virtual museums and web-based digital ecosystems

Abstract

The Virtual Museum of the Pacific was developed by the authors and launched by the Australian Museum in November 2009. A digital ecosystem implemented as Web 2.0 application, the VMP is an experimental platform with information and knowledge acquisition for the Australian Museum's Pacific collection. Importantly, the VMP facilitates a number of social media interfaces that enable content to be added and tagged, the control vocabulary to be extended, user perspectives to be defined and narratives added via wiki. It is therefore extensible. This paper concentrates on technical issues involving several search methods: attribute search based on a control vocabulary, query refinement and query-by-example. In doing so, the paper illustrates the issues involved in abstracting a generic collection management framework that can be re-used for developing other semantic Web digital ecosystems for other collection content.

Disciplines

Physical Sciences and Mathematics

Publication Details

Eklund, P. W., Goodall, P. & Wray, T. (2010). Virtual museums and web-based digital ecosystems. 4th IEEE International Conference on Digital Ecosystems and Technologies, DEST 2010 (pp. 141-146). Piscataway, New Jersey, USA: IEEE.

Virtual Museums and Web-based Digital Ecosystems

P. Eklund, P. Goodall, T. Wray
School of Information Systems and Technology
University of Wollongong, Australia
pjgoodall@gmail.com
peklund,twray@uow.edu.au

Abstract—The Virtual Museum of the Pacific was developed by the authors and launched by the Australian Museum in November 2009. A digital ecosystem implemented as Web 2.0 application, the VMP is an experimental platform with information and knowledge acquisition for the Australian Museum’s Pacific collection. Importantly, the VMP facilitates a number of social media interfaces that enable content to be added and tagged, the control vocabulary to be extended, user perspectives to be defined and narratives added via wiki. It is therefore extensible.

This paper concentrates on technical issues involving several search methods: attribute search based on a control vocabulary, query refinement and query-by-example. In doing so, the paper illustrates the issues involved in abstracting a generic collection management framework that can be re-used for developing other semantic Web digital ecosystems for other collection content.

I. INTRODUCTION

In last year’s proceedings, we reported a digital ecosystem based on the representation of an information space derived from concept lattices for tagging, loading and annotating digital objects from a museum collection [12]. The program, a Web 2.0 application called the Virtual Museum of the Pacific (VMP), has now been released. Since then we have re-used the software framework to built other simple Web-based digital ecosystems. This paper reflects on our framework design. COLLECTIONWEB provides an extensible environment for the creation of attribute lists and user-defined perspectives that can be used to flexibly navigate a collection of digital objects based on any user-defined semantic theme. The program also contains a wiki component to encourage its user communities to add narratives to the collection.

The paper is structured as follows. First, we introduce Formal Concept Analysis so that the paper is self-contained. From the theory of Formal Concept Analysis we can then present the digital ecosystems development framework called COLLECTIONWEB. COLLECTIONWEB is a the platform for the creation of a VMP as well as several other Web-based Digital Ecosystems that have or are being developed, including a system called ACE. Next, we turn to the design of the program surveying the literature leading to the current system: the COLLECTIONWEB is the culmination of more than 10 years of research and practice in the use of concept lattices for information browsing [4], [5], [9], [7]. We also examine the feature set of the COLLECTIONWEB to provide a snapshot of the current functionality of the system. We explain and present

its interfaces with reference to a technical audience with an interest in knowledge modeling and developing Semantic Web applications.

II. FORMAL CONCEPT ANALYSIS

Formal Concept Analysis (FCA) [18] is a well-known technique for data analysis that involves the synthesis of a formal concept as a collection of objects that exhibit a common set of attributes. FCA offers a formalization of concepts understood in the philosophical tradition, namely where a concept is considered a unit of thought constituted by its extension and its intension. The theory of FCA argues [17] that by directly representing the context in which formalized concepts are defined, it is possible to provide enough references to allow the interpretation of the formalized concepts in human communication and augmentation. A *formal context* $\mathbb{K} := (G, M, I)$ is a triple where G is a set of formal *objects*, M is a set of *attributes* and I is an *incidence* relation between the objects and the attributes. $I \subseteq G \times M$ is a binary relation where $(g, m) \in I$ is read “object g has attribute m ” and written gIm . A formal context can be represented as a cross-table where the rows represent objects (G), the columns attributes (M) and the incidence relation I by a series of crosses as shown in Tab. I.

A formal concept can be derived by taking an object and collecting all attributes describing it, then collecting all objects described by those attributes. For instance, taking the object $e002509$, let B be its attributes, $B = \{\text{body accessories, fan, melanesia, fiji}\}$. Now, let A be the objects that have all attributes in B . $B = \{e002509, e090525\}$ so the formal concept (A, B) is $\{\{e090525, e002509\}, \{\text{body accessories, fan, melanesia, fiji}\}\}$.

Concepts can be arranged in a specialization hierarchy. A concept (A, B) is a sub-concept of another concept (C, D) if $A \subseteq C$ (or equivalently $B \supseteq D$). Via this definition, more specific concepts have fewer objects and more attributes, e.g., $(A, B) < (C, D)$ where $(A, B) = \{e002509, e090525\}, \{\text{body accessories, fan, melanesia, fiji}\}$ and $(C, D) = \{e002509, e050525, e058551-004\}, \{\text{body accessories, melanesia, fiji}\}$.

The set of all formal concepts of a formal context, together with the specialization relation, forms an algebraic structure called a concept lattice [18] shown in Fig. 1 and

Fig. 1. The corresponding line diagram $\mathfrak{B}(\mathbb{K})$ of the cross-table (context) presented in table I.

TABLE I

A FORMAL CONTEXT $\mathbb{K} := (G, M, I)$ CONTAINING INFORMATION ABOUT THE OBJECTS (G) AND THEIR ATTRIBUTES (M) FOR THE SUB-CONTEXT RESULTING FROM APPLYING “BODY ACCESSORY” AND “ORIGIN” SCALES FROM THE AUSTRALIAN MUSEUM’S PACIFIC COLLECTION.

$\mathbb{K}$	body accessories	fan	head ornament	ankle ornament	fly whisk	neck ornament	melanesia	polynesia	fiji	papua new guinea	samoa
e002509	×	×					×		×		
e090525	×	×					×		×		
e058551-004	×				×		×		×		
e091567	×					×		×	×		
e091570	×					×	×		×		
e002415	×	×						×			×
e002416	×	×						×			×
e058169	×		×					×			×
e058169	×				×			×			×
e011543	×			×			×			×	

denoted $\mathfrak{B}(\mathbb{K})$, where $\mathbb{K}$ is the formal context. The concept lattice provides a mechanism to visualize information contained within a formal context. Evident from the diagram is the specialization hierarchy over the concepts and also the implicational logic between objects and attributes that are implicit in the context, so for example all $\{\text{head ornament}\} \rightarrow \{\text{body accessories, polynesia, samoa}\}$.

Further, let (G, M, W, I) be a many valued context and for each $m \in M$ let $S_m := (G_m, M_m, I_m)$ be a scale for m . The *derived context* of (G, M, W, I) with respect to *plain scaling* with the scales $(S_m | m \in M)$ is then (G, N, J) where:

$$N := \prod_{m \in M} \{m\} \times M_m \quad \text{and}$$

$$(g, (m, n)) \in J \Leftrightarrow \exists w \in W (g, m, w) \in I \text{ and } (w, n) \in I_m.$$

This definition of a many-valued context allows for the creation of contexts that include cells that contain something other than a boolean ($\times$) value and permits the presentation

of attributes and their values in Formal Concept Analysis. Conceptual scales [13] are a powerful tool that can be used to store views that partition the data being analysed. For example, table I is actually a sub-context combining two conceptual scales for “body accessory” and “origin”, this can be seen by the fact that “body accessory” an attribute common among all the objects in the context and that the mutually exclusive attributes “polynesia” and “melanesia” are present for each object in the context. Within a FCA system multiple views can be stored and combined to effectively query the data, “body accessory” and “origin” (shown in table I) are 2 of 15 predefined conceptual scales used in the VMP which include other scales, such as “general keywords”, “indigenous names” or “body accessories”. Other scales (which we call “perspectives” in the user interface but can also be considered as facets in an information retrieval sense) are extensible and can be defined by the stakeholders depending on their interests.

III. SEMANTIC CONCEPTS AND ONTOLOGIES

The digital ecosystem of the Virtual Museum of the Pacific at its simplest level revolves around digital representations of real-world museum artifacts in the virtual collection as individuals. These individuals are encircled by attribute clouds that relate to different feature sets: material, geographic, anthropological, scholarly descriptions, impromptu or informal narratives etc. This architecture is illustrated in Fig. 2.

A key understanding for constructing a generic platform for Web-based Digital Ecosystems is to understand how the digital ecosystem metaphor can be made generic to any collection content. To learn how COLLECTIONWEB can be viewed as a digital ecosystem, we consider its semantic concepts and the ontology of the Digital Ecosystem [1] to situate our work according the standard references and refer to Fig. 2

Individuation – although abstractions from real world objects, the digital objects in a virtual museum can be considered as individuals, most commonly, an individual is thought of by representation of one or more high-resolution image but more precisely it is represented as a unique enterprise key with multiple facets represented as metadata, some of which might be rich media including images.

The metadata itself can be multi-faceted, a single real-world object can be represented by several digital images (representing various views of the object) and may contain various metadata tags. For example a given digital object might include the attributes ‘pine’, ‘stone’ and ‘shell’ if from the the Australian Museum’s Pacific Collection or ‘aboriginal’, ‘female’, ‘Sydney’, ‘canvas’ if it describes an art work in ACE.

Attributes are themselves aggregated, so that ‘pine’ is a subclass of ‘wood’ – this builds a containment hierarchy. Likewise, an attribute as a location has a natural containment hierarchy, so the ‘Sydney’ is part of ‘New South Wales’ which in turn is part of ‘Australia’. Attributes can therefore be individuals when they are leaf nodes in the attribute hierarchy. Likewise, users of the COLLECTIONWEB as individuals.

Fig. 2. The Digital Ecosystem of the Virtual Museum.

Packet Level – the act of tagging an attribute to an object is a packet, this itself can be an act derived from a control vocabulary of attributes via the free definition of an attribute to be used as a tag by a user. Perspectives too are packets, for example the ‘raw material’ perspective might include the attributes ‘wood’, ‘stone’, ‘shell’, ‘metal’ and ‘charcoal’ that help describe what an object is made of. Likewise, the creation of a story or narrative via the wiki describes a packet as does associating one digital object with another or with rich multimedia. A user is associated with a group – this is also a packet – e.g. curator, director, dealer, collector, researcher, public or indigenous community member.

Context – the arrangement of digital objects within a formal concept indicates all relevant objects within that context. Context then has several meanings in the COLLECTIONWEB, the context of the collection, say the 427 objects selected from over 60,000. Also, the search, browsing or navigation context, take for instance the Web services call *getConcept*(‘polynesia’, ‘samoa’, ‘wood’). This might return five objects that satisfy the attributes ‘polynesia’, ‘samoa’ and ‘wood’. The induced formal concept is a context. So too, a ‘perspective’ is a context because it represents a many-valued context in the theory of formal concept analysis [18].

Ecosystem – the information space formed by the object attributes themselves, 427 objects in the case of the VMP and 60 in the case of ACE, but more more ambitiously, multiple, linked museums sharing the same control vocabulary form the infrastructure of a more extensive digital ecosystem. Similarly, the stakeholders all represent ecosystems of related interest. Exactly an extension of the VMP work is proposed by the Australian Museum to link its partner museums and pacific

content.

Cross-Ecosystem Relationships – any inferences that are formed as a result of user communities tagging and annotating objects (or writing wiki entries) in their own group, represent relationships across ecosystems. For example, historians, anthropologists and curators are each members of their own ‘group’. If Anthropologists are a group within their own ‘ecosystem’ and the historians belong to another, a cross-ecosystem relationship is formed from their respective collaborative inputs into the COLLECTIONWEB.

A. Navigation and Conceptual Neighborhoods

Kim and Compton [15] developed a document navigation interface using Formal Concept Analysis [18] and a conceptual neighborhood display. Their program, *KANavigator* uses

Fig. 3. ImageSleuth: the interface presents only the extent of the current concept as thumbnails and generalizations/specializations by removal/addition of attributes (tags) to reach the upper and lower neighbors (shown to the top/bottom of the thumbnails). Pre-defined scales (called perspectives) are displayed on the left. The thumbnail images shown in the screenshot are from the SIMS2.

annotated documents that can be browsed by keyword. The program displays the direct neighborhood (in particular the lower neighbors) as its interface. At the time, this emphasis on the use of textual labels as representations of single formal concepts (as opposed to a line diagram of the concept lattice) broke with many of the discipline's traditions. However, subsequent usability studies have shown that an interface based on a conceptual neighborhood has significant usability merit [10], [8], simplifying the interaction and enabling non-expert users to intuitively interact with a concept lattice as a representation of an information space.

COLLECTIONWEB interfaces follows from experimentation with three different applications — ImageSleuth 2.0 [9] is the progenitor COLLECTIONWEB application shown in Fig. 3. It employs the conceptual neighbourhood paradigm for browsing and display purposes. The COLLECTIONWEB is however supplemented with the ability to tag and annotate objects (images and their metadata) using an extensible control vocabulary and a wiki. Another COLLECTIONWEB system called ACE (was developed by Wray and Eklund) and usability testing was conducted on that program involving a trial with 25 test subjects. Interface design flaws were then discovered with ACE usability testing as the focus and these were then applied to COLLECTIONWEB and downstream to upgrade the VMP interface. Within a COLLECTIONWEB application, users can create new perspectives (concept scales [18]). The user can edit and add new attributes (tags) to the control vocabulary that can be used to create/edit the perspectives. New objects can be added to the system allowing the virtual collection content to grow. The program therefore represents an extensible museum content management system with a flexible mechanism for adding, exploring and tagging the set of objects in the collection. This can be done through the provision of Web Services, allowing the development of any Web or desktop application on any platform to query and extend the collection.

B. Design and Architecture

COLLECTIONWEB is no substitute for a physical museum experience, it is intended as an entirely new experience. This apparent conflict needs careful stage management in terms of

Fig. 4. The VMP tagging and browsing interface

the diffusion of the COLLECTIONWEB, incremental, bottom-up and relatively small scale initially. Our first efforts are with 427 in the VMP and 60 objects with ACE for this reason.

Further, the project adopts an agile development approach, with whole of system development, unit and iterative usability testing as key aspects of the development life-cycle.

Fig. 5. COLLECTIONWEB: museum management interface for adding a new image to the collection. Navigation to other management interfaces is via the hyperlinks across the top-frame.

Fig. 6. Adding a new attribute into the COLLECTIONWEB prototype system. This allows the vocabulary of tags for describing artifacts to be extended by curators, researchers or indigenous user communities. Existing attributes (tags) are listed in the left frame, a new attribute can be added via the textbox.

IV. COLLECTIONWEB: INTERFACES AND OPERATION

Within a COLLECTIONWEB application, users can upload new content in the form of objects and tags. It is important that a rigid view of the collections shall not be imposed - for this reason, any meta-data associated with an object is converted into a form that can be represented by tags. Subsequently, the meta-data and description of any object within the collection can be described by tags within a control vocabulary as shown in Fig. 7. User communities can expand this vocabulary by adding new tags and definitions as shown in Fig. 6. Additionally, Fig. 5 demonstrates the ability to upload images and media associated with each object. Every object must have at least one image representation, so as to provide a meaningful and visual experience while browsing the collection. COLLECTIONWEB uses most of its interface to show thumbnails of images in the extent of the chosen concept (see Fig. 9). The lattice structure around the current concept is represented through the list of upper and lower neighbours which allow the user to move to super- or sub-concepts.

A. Adding Narratives to the Collection

Integrating a wiki with COLLECTIONWEB was the first step towards providing user communities with the ability to add and enhance the virtual collection. Just as in other wiki's the content will need to be moderated and users will be asked to register and verify their identity by email. The other management interfaces, such as for adding objects, attributes and perspectives, have more restricted access.

Fig. 7. Editing an object individually in the COLLECTIONWEB. The list of all current attributes (tags) for the top-most object in the left frame are listed under “current”. The remaining attributes that can be selected are listed under “available” in the right frame.

Fig. 8. Adding or editing an object’s attributes in COLLECTIONWEB prototype. In this figure all the objects highlighted (red) have the attribute leaf.

B. Search and Query-By-Example

Browsing is achieved by moving to neighboring concepts. In many cases the user will want to go directly to images having a certain set of attributes. This is offered by the direct search function that enables the use to type attributes into a text field. Another type of search is performed by the query-by-example. Instead of defining a set of attributes, a set of objects is defined as the sample set. The query-by-example function

Fig. 9. ACE: the browsing interface for COLLECTIONWEB.

Fig. 10. COLLECTIONWEB prototype wiki. Text can be added to Objects in the collection via the browsing interface. This provides some early functionality required to collect information from the user communities of the virtual museum.

then computes the common attributes of these images (in the selected sub-context) and returns all other images having these attributes. In this way, query-by-example is the dual of the search function. While the search for images having certain attributes is not affected by the removal or addition of perspectives to the sub-context, query-by-example depends on the selected sub-context. The more attributes taken into consideration, the smaller the set of images that have exactly the same query example.

C. Concept Similarity

The aim of query-by-example is to find objects which are similar to the objects in a given sample set. This is a narrow understanding of similarity, implying equivalence in the considered sub-context; for the query-by-example function two objects are “similar” in a sub-context. If the objects are uniquely described by the attributes in the chosen sub-context then query-by-example seldom yields new information.

A more general approach is to define a similarity measure for pairs of concepts. Similarity of two objects g and h is then described as the similarity of the attribute sets g' and h' . In order to use the grouping of objects provided by the formal concepts, the VMP works with a similarity measure on semi-concepts which allows the return of a ranked list of similar concepts. The similarity measure is derived from the metrics described in [6] and provide a relevance ranking mechanism.

V. RELATED WORK AND FUTURE DEVELOPMENT

In terms of the design and implementation of the COLLECTIONWEB framework, we have previously identified the individuals and their roles within the digital ecosystem along with an outline of an access control model [11]. In spite of this, we have yet to fully test it among a large user base of tagging, annotations or other cross-ecosystem interactions. In addition, we will aim to address the challenge of incorporating user-defined tags and annotations within existing collections. These

Fig. 11. COLLECTIONWEB: A popular way of browsing is to use the current object to find like objects. This is done via by query-by-example, this illustrates another modality for search and discovery.

ideas are currently being explored in the literature concerned with the harvesting of social tagging meta-data in adding value to collections [3][14][2] and a parallel cross-institutional project known as the steve.museum [16].

The scalability of the COLLECTIONWEB framework remains a further, and critical, determinant of its practical use in large collections. In doing so, we plan on making the content within the Virtual Museum of the Pacific and ACE collection-complete, allowing us to assess the performance of its Web Services components and its practical use in augmenting the exploration and social tagging of large collections.

VI. CONCLUSION

In this paper we have described the COLLECTIONWEB framework as a way of building Web-based digital ecosystems. In COLLECTIONWEB systems objects of a digital collection of virtual museum artifacts are derived from facets of the physical artifacts held in original collection. COLLECTIONWEB permits several diverse search methods: attribute search based on a control vocabulary, search via query refinement and query-by-example. The framework provides a number of management interfaces that enable content to be added and tagged, the control vocabulary to be extended, user perspectives to be defined and narratives added via a wiki. We describe COLLECTIONWEB as a Semantic Web application generation framework that uses a Formal Concept Analysis Web service engine. We identify elements of the digital ecosystem by purpose,

function and stakeholder. We describe and offer solutions to digital ecosystems design in terms of the deployment, the technological platform.

ACKNOWLEDGMENT

This work is supported by the Australian Museum and the Australian Research Council linkage grant LP0884075.

REFERENCES

- [1] H. Boley and E. Chang. Digital ecosystems: Principles and semantics. In *Digital EcoSystems and Technologies Conference, 2007. DEST '07*, pages 398–403. IEEE Press, 2007.
- [2] D. Boyd, M. Davis, C. Marlow, and M. Naaman. HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, To Read. In *Proceedings of the ACM Conference on Hypertext and Multimedia*, 2006.
- [3] S. Chan, L. Kelly, A. Russo, and J. Watkins. How will social media affect museum communication? In *Proceedings of the conference of Nordic Digital Excellence in Museums*, 2006.
- [4] R. Cole and P. Eklund. Analyzing an email collection using formal concept analysis. In *European Conf. on Knowledge and Data Discovery, PKDD'99, LNAI1704*, pages 309–315. Springer-Verlag, 1999.
- [5] R. Cole and P. E. G. Stumme. Document retrieval for email search and discovery using formal concept analysis. *Applied Artificial Intelligence*, 17(3):257–280, 2003.
- [6] F. Dau, J. Ducrou, and P. Eklund. Concept Similarity and Related Categories in SearchSleuth. In *Proceedings of the 16th International Conference on Conceptual Structures (ICCS'08)*, LNCS 5113, pages 255–268. Springer-Verlag, 2008.
- [7] J. Ducrou and P. Eklund. An intelligent user interface for browsing and search mpg-7 images using concept lattices. *International Journal of Foundations of Computer Science*, 19(2):359–381, 2008.
- [8] J. Ducrou and P. Eklund. Faceted document navigation using conceptual structures. In P. Hitzler and H. Schärff, editors, *Conceptual Structures in Practice*, pages 251–278. CRC Press, 2009.
- [9] J. Ducrou, B. Vormbrock, and P. Eklund. FCA-based Browsing and Searching of a Collection of Images. In *Proceedings of 14th International Conference on Conceptual Structures, LNAI4068*, pages 203–214. Springer, 2006.
- [10] P. Eklund, J. Ducrou, and P. Brawn. Concept lattices for information visualization: Can novices read line diagrams. In P. Eklund, editor, *Proceedings of the 2nd International Conference on Formal Concept Analysis - ICFCA'04, LNAI2961*. Springer-Verlag, February 2004.
- [11] P. Eklund, P. Goodall, T. Wray, V. Daniel, and M. V. Olffen. Folksonomy with Practical Taxonomy, a Design for Social Metadata of the Virtual Museum of the Pacific. In *Proceedings of the 6th International Conference on Information Technology and Applications*, pages 112 – 117, 2009.
- [12] P. Eklund, T. Wray, P. Goodall, B. Bunt, A. Lawson, L. Christidis, V. Daniels, and M. V. Olffen. Designing the digital ecosystem of the virtual museum of the pacific. In *Proceedings of the 3rd IEEE International Conference on Digital Ecosystems and Technologies*, 2009.
- [13] B. Ganter and R. Wille. Conceptual scaling. In F. Roberts, editor, *Applications of Combinatorics and Graph Theory to the Biological and Social Sciences*, pages 139–167. Springer-Verlag, 1989.
- [14] J. Hunter, I. Khan, and A. Gerber. Harvana: harvesting community tags to enrich collection metadata. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital Libraries*, 2009.
- [15] M. Kim and P. Compton. The perceived utility of standard ontologies in document management for specialized domains. *International Journal of Human-Computer Studies*, 64(1):15–26, 2006.
- [16] J. Trant and B. Wyman. Investigating social tagging and folksonomy in art museums with steve.museum. In *International World Wide Web Conference*, 2006.
- [17] R. Wille. Restructuring lattice theory: an approach based on hierarchies of concepts. In I. Rival, editor, *Ordered Sets*, pages 445–470. Reidel, 1982.
- [18] R. Wille and B. Ganter. *Formal Concept Analysis: Mathematical Foundations*. Springer, Berlin, 1999.