

PATHS: personalising access to cultural heritage spaces

Kate Fernie, Jillian Griffiths,
MDR Partners, London, UK

Phil Archer, Konstantinos
Chandrinos,
iSieve Technologies, Greece

Andrea de Polo
Alinari 24 ORA, Florence, Italy

Mark Stevenson, Paul Clough,
Paula Goodale, Mark Hall,
University of Sheffield, UK

Eneko Agirre, Oier Lopez de
Lacalle,
University of the Basque Country

Runar Bergheim
Asplan Viak Internet, Norway

Abstract— Digitisation of the cultural heritage means that a significant amount of material is now available through online digital library portals. However, the vast quantity of cultural heritage material can also be overwhelming for many users who lack knowledge of the collections, subject knowledge and the specialist language used to describe this content. Search portals often provide little or no guidance on how to find and interpret this information. The situation is very different in museums and galleries where collections are organized in exhibitions which offer themes and stories that visitors can explore. The PATHS project, which is funded under the European Commission's FP7 programme, is developing a system that explores the familiar metaphor of a trail (pathway) to enhance the discovery and use of the content made available in digital libraries. This paper will report on the findings of the user requirements analysis and the specifications for the first prototype of the PATHS system which is based on contents from Europeana and the Alinari Archives.

Keywords- cultural heritage, digital libraries, pathways, storytelling, Europeana

1. INTRODUCTION

Digitisation of the cultural heritage in recent years means that a significant amount of material is now available through online digital library portals such as Europeana [1], the European aggregator for museums, archives, libraries, and galleries. However, the quantity of material that is available can be overwhelming for many users who lack knowledge of the collections and can have difficulty in finding out what items are available. The standard search paradigm [2] based on simple keyword searching can seem like a closed door to end users who lack the specialized vocabulary used to index content in such collections. The situation is very different in museums and galleries where collections are organized thematically in exhibitions and around stories which help guide visitors through the content.

The PATHS project [3], which is funded under the European Commission's FP7 programme, is exploring alternatives to standard search portals as the means of item discovery and

exploration [4, 5]. We recognise that the users of cultural heritage portals have a range of information needs and that they have very individual styles of seeking that information with some preferring to follow narratives, some appreciating receiving recommendations while others are happy to forage through collections [18].

PATHS is exploring the use of "paths" or trails as a way of enabling users to navigate their way through large and complex collections and the use of a range of visualization techniques to enhance their experience of using digital libraries. The PATHS system is built using a set of web-services that access a shared data storage and provide access to user profiles, items of content, and the functionality needed to visualize content and to follow, and also create, paths. The system is based on content which has been provided by Europeana and Alinari 24 ORA and then processed and enriched to enhance the connections between items. It is this enrichment which forms the base for a system which acts as an interactive guide through digital library collections, offering suggestions about items to look at and links to contextual information.

This paper is organized as follows. We begin with a description of some existing approaches which have used trails to provide access to online information. We then report on the findings of the survey of Paths user requirements, the development of case studies and then report on the first prototype of the PATHS system. The paper concludes by summarizing the role of user evaluation and testing in the future development of the system.

2. BACKGROUND: USING TRAILS TO EXPLORE LARGE DOCUMENT COLLECTIONS

The use of trails to let users explore large, unknown document collections is not new. Typically trails through a

collection are implemented to introduce users to an unfamiliar subject as a step-by-step guided tour where a sequence of content is presented with explanatory text. Walden's Paths [6, 7] pioneered the idea of providing users with manually created paths through a collection (in this case the collection is the web). Webwatcher is another example, a system for automatically generating routes through web content [8].

Trailmeme [9] is a 'content curation' system developed by the Xerox Corporation as 'a way to tell stories with Web content'. Users create trails by collecting a series of web pages, creating a node for each page with added descriptions and tags which are then linked together to form the trail. Users 'walk' trails using the trail map which is relatively straightforward for simple linear trails but more complex when nodes have multiple options. This complexity is in the hands of the trail creator - it requires some skill to create a map style trail where the user can follow all of the nodes easily. There are currently more than 750 trails in Tralimeme, ranging from history and science lessons, to news, business, technology, travel guides, health and popular culture.

Typically such trails are linear in form, with the selected objects or hyperlinks being viewed sequentially to present a narrative created by subject 'experts' for consumption by others. With only a limited degree of interactivity, online trails represent a relatively passive experience for the end user.

Unlike Walden's Paths and Trailmeme which are based on the Web, PATHS is designed to work with digital library collections, with built-in search, browse and explore functionality that will enable the user to find high quality content that is normally hidden from search engine results.

PATHS aims to offer users a more interactive experience using the pathway as a starting point for exploration of the items in a collection. To overcome the difficulty that users face in exploring the unknown contents of these very large collections the project began with the proposition of using state-of-the-art techniques in information access, language processing and personalization within a user centred design approach.

3. FINDING OUT ABOUT PATHS USERS

It was important to the project to begin by defining the target user groups for the PATHS system and then find out about their requirements [10]. We decided to focus on expert and non-expert users in the cultural heritage, education and general leisure domains (e.g. tourism and local history domains) and then set out to develop profiles of their information behaviour characteristics and to build up a detailed understanding of the context in which they might create and use paths, and the processes they would use to do so.

A mixture of methods was used to gather information about potential users of the PATHS system and their requirements including:

- Desk research
- An online questionnaire which was distributed to cultural heritage, educational, professional users and general users
- Face-to-face interviews with expert users in the UK, Italy and Spain
- Workshop involving students and general users
- Workshop with professional users
- A remote path creation exercise involving project partners

The desk research confirmed that the 'pathway' metaphor is not new to the cultural heritage but also confirmed that most existing examples of paths have typically been created by experts for others to use, often in an educational setting. The users who completed our online survey reported a high level of confidence in their ability to search the internet and use of a broad range of sources. Overall the results of our survey suggest the potential for greater numbers of users to make use of path creation tools for a range of purposes and to share these with others for a range of purposes such as storytelling.

Detailed analysis of the results of the user requirements gathering allowed us to develop a conceptual model of how users interact with paths in five core behaviours from: developing a *Concept*; *Collecting* relevant materials; *Creating* a path; *Communicating* about the path and its content; and *Consuming* paths created by other users.

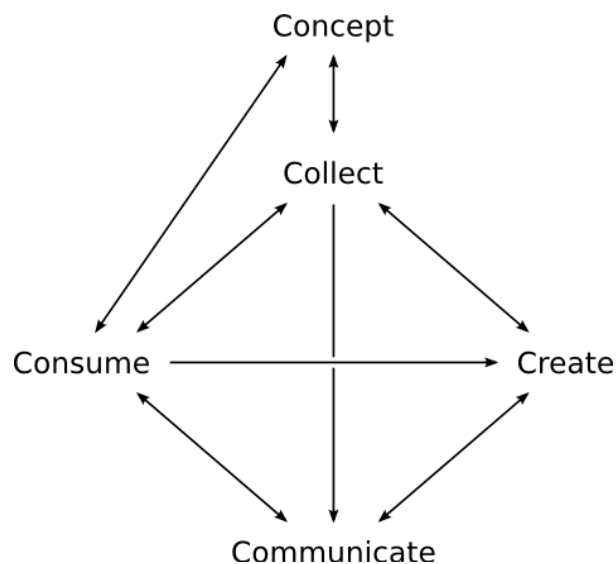


Figure 1 Conceptual Model of User Interactions with Paths

Four generic user profiles were identified from the interview data:

- Path creator (expert)
- Path creator (non-expert)
- Path facilitator
- Path consumer

These generic profiles were used to consider how different users interact with the conceptual model and with their requirements from the PATHS system. For example, an expert path creator is likely to begin with a concept, move on to collecting materials and creating the path while a non-expert path creator might begin by consuming an existing path and then move to collecting materials.

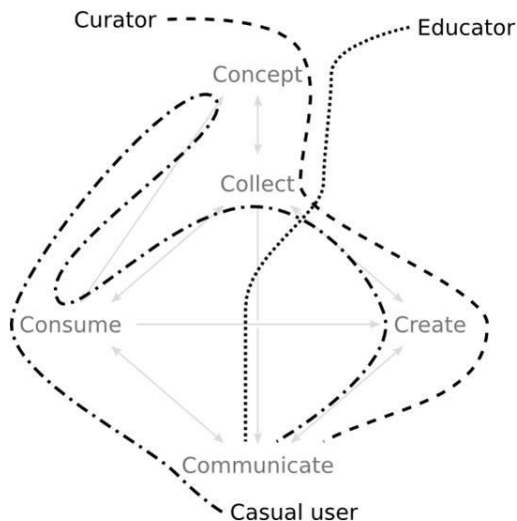


Figure 2: User interactions around the conceptual Model

4. CASE STUDIES

Having modeled different types of possible interaction with PATHS we created a series of concrete use cases. For example, one use case envisaged a curator from the Bletchley Park museum in the UK with excellent knowledge of the history of the site during the Second World War and the collections setting out to create a new online exhibition.

- The curator starts with a concept or theme, in this case an aspect of the Bletchley Park story;
- Carries out background research on the theme, consulting for example published articles, collection catalogues and other reference materials.
- Identifies sub-themes (related concepts), for example code-breaking, wartime life, early computing etc.
- Collects items from the collections to illustrate these concepts, for example photographs, documents, recordings, videos, maps etc.
- Identifies relevant items from external collections
- Writes or edits descriptions of items in the collection
- Prepare materials, for example introductions to the sub-themes
- Creates pathways and designs end-user interactions, for example encouraging users to:
 - Stray off the path and explore concepts further
 - Add comments

- Upload personal content, e.g. photographs or stories about themselves when they worked at Bletchley Park
- Communicate about the existence of the pathway to end users, for example by:
 - Adding a start point to the pathway on the home page of the organization’s website
 - Publishing a news story and sharing the news through social networking sites
- End-users consume the pathway, for example by reading stories, following the links, sharing comments, exchanging news etc.

The Bletchley Park Case study is an example of a path being created by an expert for non-expert end-users to use (see figure 2). Another case study envisaged a pathway being created by a non-expert user with an interest and enthusiasm in the local history of their area. In this case study the enthusiast begins by finding interesting content, reads about it and then begins to discover interesting connections or themes which are later organized into a Pathway. In this case study the pathway is shared with the members of a local history society and fellow students who are invited to comment – and add to – the Pathway that has been created.

We envisaged another case study of a Pathway used in a mobile application by a tourist to explore Roman London, and taking in a visit to the Museum of London.

The use cases that were envisaged flowed directly from the research done into user profiles and cognitive styles during the project, which included group Path-creation exercises. For example, one group focused on the use case of a primary teacher preparing to pathway for a class to explore the theme of war-time Britain. This group focused on children’s experiences of the war-time and identified a set of items that could be used to build a narrative including music, ration books, posters and other documents

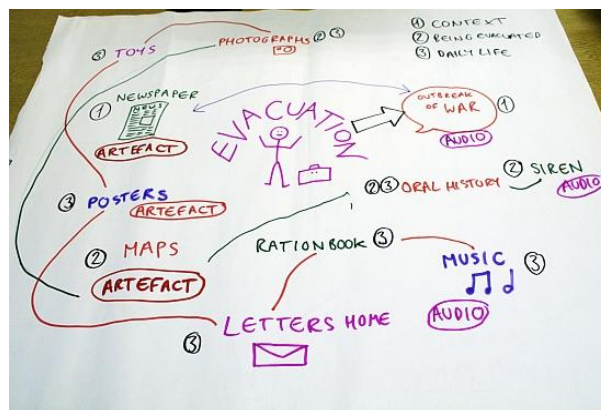


Figure 3: Designing narrative pathways

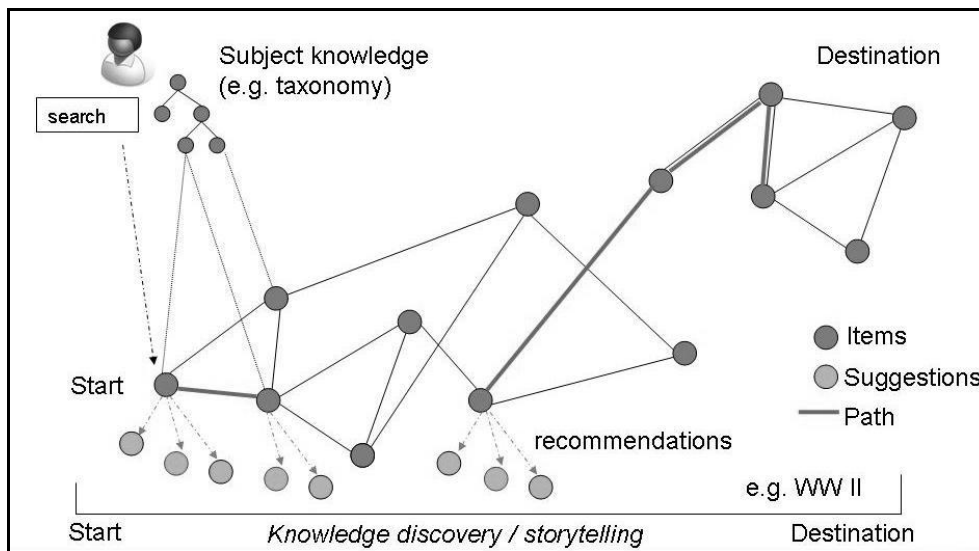


Figure 4: Pathways through a collection

The Case Studies and group workshops helped to confirm both requirements that were easily foreseen (such as being able to search a collection) and helped to identify requirements that were either unforeseen or which users placed much greater importance than anticipated. For example, users were very vocal in their support for everyone to be able to contribute their own knowledge even if they were consuming someone else's Path. The desire to share Paths with others – friends, students, colleagues, visitors – was clear and a recognised trend in cultural heritage systems [11]. Users expressed a clear requirement to be able to clone existing Paths and re-use them to create something new.

5. FROM USER REQUIREMENTS TO FUNCTIONAL SPECIFICATIONS

The needs and requirements that were expressed by users through the research were listed in a project deliverable and assigned a priority level in the style of RFC2119 keywords MUST, SHOULD and MAY [10, 12]. However, there are further requirements that need to be met in order to realise the platform as a whole, such as managing user accounts. PATHS is also proposing innovations that users have not necessarily asked for directly. As Henry Ford famously said, “If I'd asked customers what they wanted, they would have said ‘a faster horse.’” Ford offered them the motor car.

All of the requirements that were identified during the course of our user research were analyzed in order to prepare the functional specification for the first prototype of the PATHS platform was defined. In defining the specification we took into consideration the need to:

1. Provide a core system which could be evaluated by users to inform the design of a second prototype;
2. Build a system architecture that can be extended in the second prototype and in future applications;

3. Integrate the sample data collections provided to the project by Europeana and Alinari 24 ORA;

To address the known problem that users face when looking at unknown collections for the first time, and to accommodate users' different cognitive styles and preferences in seeking information, the PATHS prototype includes a variety of visualisation techniques as well as the Paths creating and following functionality [13]. The visualisation techniques are designed to provide the user with an overview over the collection and to improve access, for example by grouping like items together [14, 15, 16].

6. THE PATHS PROTOTYPE SYSTEM

The first PATHS prototype system is based on two collections of data, one provided by Alinari 24 ORA and the other provided by Europeana. This data has been processed and enriched using language processing techniques to extend the indexing provided in the original metadata, generate links between items in the collection based on similarity, and also to create links between items and external contextual materials such as pages in Wikipedia [17]. The enriched data produced by these techniques forms the basis for the PATHS prototype and increases the potential for navigation and visualization of the collection.

In the first PATHS prototype, our aim has been to demonstrate a collection discovery environment that combines the best aspects of following a narrative Pathway and visualisation techniques in an integrated system. Our idea is to offer users both options and ways to assist non-expert users or other novices in making the transition from being a passive consumer of existing Paths, to exploring and creating Paths for others to enjoy, and to cater for different user preferences in terms of the way in which users like to consult digital libraries.

The user interface has 3 main areas:

- **SEARCH** – Searching is one of the core functions of the PATHS system and is designed to support discovery of items with the collection and Paths. Users can start a search by a simple keyword or by using the facets provided.
- **EXPLORE** – this area is designed to enable users to explore the document space using strategies intended to suit different cognitive styles. For example, a thesaurus is offered for users with a preference for text; illustrated topics for users who prefer images, and topic maps for users who like to have an overview of the whole documents space.
- **PATHS** – this area provides access to existing Pathways (or narratives) that have been published by users and also provides access to Path creation tools. Users can follow a Path from the beginning or jump to a node and discover ancillary materials along the way by following recommendations and links.

6.1 Creating and Using Paths

Figure 5 below shows an example of a Path based on content from the Alinari collection. The Path consists of a linear sequence of items taken from the collection used by the author to create a narrative (in this case about mountains). Each item in the Path is displayed in a ‘node component’, which includes the original metadata plus annotations and recommendations – links to similar items, contextual materials and subject keywords. User can follow a Path from the beginning in a linear fashion (in classic storytelling mode).

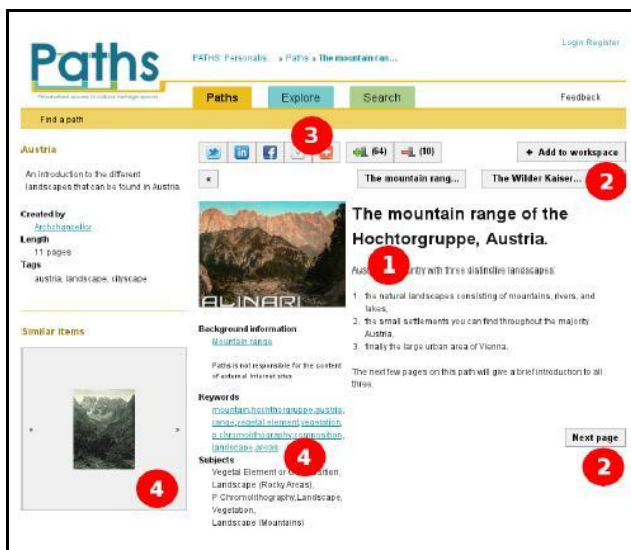


Figure 5: The path-following interface showing of an item with narrative (1), navigation tools (2), social features (3), and starting points for exploration (4).

The feedback that we received during the requirements gathering exercise indicated that users would like flexibility. For this reason the PATHS system is designed to allow users to join a Path in the middle (each node is separately

discoverable in the search interface), to navigate forwards or backwards and diverge from the Path to look at similar items and explore a subject in more depth.

The Path creation functionality provided in the system is designed to enable authors to create virtual exhibitions or to narrate stories about items in the collections. Any kind of story or narrative can be created using the interface providing that the content is available in the collection and the inspiration of the author. For the moment the prototype system is based on contents provided by Europeana and Alinari 24 ORA, these include thumbnail images but it would be possible to embed other types of preview (such as sound files or video clips) in the node component in the future.

6.2 Exploring the collection

The Paths component of the system is complemented by functionality which is designed to encourage exploration (see figure 6) such as standard vocabularies, visual representations of hierarchical topic structures and keyword clouds.

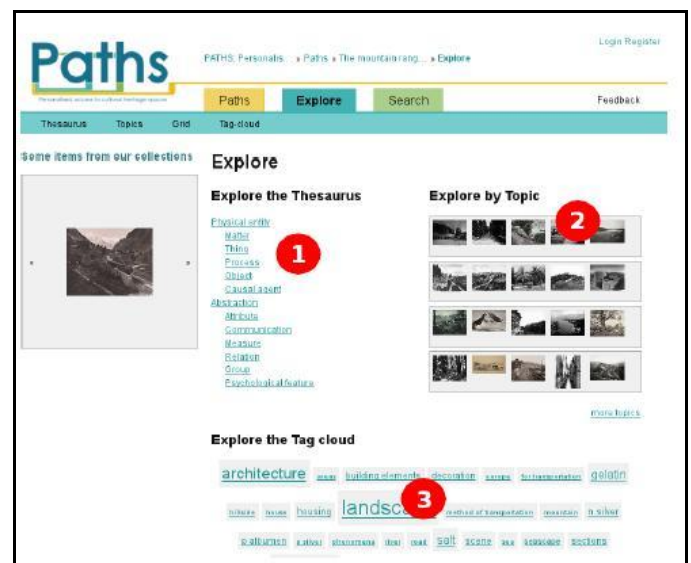


Figure 6: Exploration interface showing standard vocabularies (1), visual topics (2), keyword clouds (3).

The different methods which are offered to enable users to explore the collections are designed to appeal to users with different information seeking preferences – with some taking images as the starting point for exploration, others preferring text. Each of the starting point leads to search results which can be mined in a variety of ways.

Users have a workspace and, as they follow Pathways or explore the collection they can save items, which can later be used to form a Path. Once a Path has been published it can be shared with friends, colleagues or with everybody. PATHS is following the trend in cultural heritage systems to encourage interaction [11] by allowing users to share, tag and comment on content and Paths as they use the collection.

One of the principles on which this prototype is based is to offer users a seamless experience (one which hides the transitions between following Paths and exploring the collections). Users should not notice as they move from one to another, they should always feel in control of what they are doing, which we hope will lead to a satisfying experience.

7. EVALUATION AND TESTING

In developing the first prototype of the PATHS system we have analysed and interpreted the user requirements which were gathered early in the project. Our aim has been to implement the core part of an innovative, flexible and adaptable system which enables users to make full use of the digital library collections which have built up in recent years.

The next step for the project is to demonstrate and test the system in order to gather feedback from end-users from each of the four generic profiles we identified (expert and non-expert Path creators, facilitators and Path consumers). This will enable us to assess how well the first prototype system has met user requirements and refine our ideas for the implementation of the second PATHS prototype.

8. CONCLUSIONS

With growing amounts of digital cultural heritage information becoming available, online users require assistance with information discovery than is offered by standard search interfaces. In this paper, we suggest that by using the metaphor of Paths it is possible to create narratives incorporating items in a collection, and to enhance the experience for users by at the same time encouraging exploration. Pathways can be surfaced as different routes users may take through an information space. This helps users to locate and associate information in unfamiliar collections and ultimately to help them to fulfil broader activities, such as exploring a given subject or theme and constructing knowledge.

The PATHS project has implemented a first prototype of the planned system as a means of investigating the responses from users and groups which regularly make use of cultural heritage information. A second prototype system will be developed following user evaluation, with the overall objective of providing a more effective means of discovering the contents of digital libraries. Pathways are offered as tools that can be created to guide and assist individuals and user communities with information discovery and exploration, and to offer both expert and non-expert Path creators the potential to narrate stories and produce exhibitions based on the content of a digital library. Our goal is to enhance the user's information access experience of digital library resources.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 270082. We acknowledge the contribution of all project partners involved in PATHS (<http://www.pathsproject.eu>).

REFERENCES

- [1] Europeana website: [<http://www.europeana.eu/portal/>]
- [2] Sutcliffe, A., and Ennis, M. Towards a cognitive theory of information retrieval. *Interacting with Computers* 10 (1998), 321-351.
- [3] PATHS project website [<http://www.paths-project.eu/>]
- [4] Geser, G. Resource discovery - position paper: Putting the users first. *Resource Discovery Technologies for the Heritage Sector 6* (2004), 7-12.
- [5] Steemson, M. Dicult experts seek out discovery technologies for cultural heritage. *Resource Discovery Technologies for the Heritage Sector 6* (2004), 14-20.
- [6] Walden's Path startup page [<http://walden.cs.tamu.edu/walden/>]
- [7] Shipman III, F., Furuta, R., Brenner, D., Chung, C., and Hsieh, H. Using paths in the classroom: experiences and adaptations. In *Proceedings of the ninth ACM conference on Hypertext and hypermedia: links, objects, time and space structure in hypermedia systems: links, objects, time and space structure in hypermedia systems* (1998), ACM, pp. 267-270.
- [8] Joachims, T., Freitag, D., Mitchell, T., et al. Webwatcher: A tour guide for the world wide web. In *International Joint Conference on Artificial Intelligence* (1997), vol. 15, Lawrence Erlbaum Associates Ltd., pp. 770-777.
- [9] Trilmeme website [<http://trilmeme.com/>]
- [10] Goodale, P., Hall, M., Fernie, K., and Archer, P., Paths D1.1 User Requirements Analysis, download from [<http://www.paths-project.eu/eng/Resources>]
- [11] Bernstein, S. Where do we go from here? continuing with web 2.0 at the Brooklyn museum. In *Museums and the Web 2008: the international conference for culture and heritage on-line* (2008).
- [12] S. Bradner. Key words for use in RFCs to Indicate Requirement Levels. March 1997. Internet RFC 2119. [<http://www.ietf.org/rfc/rfc2119.txt>]
- [13] Butavicius, M. A., and Lee, M. D. An empirical evaluation of four data visualization techniques for displaying short news text similarities. *International Journal of Human-Computer Studies* 65, 11 (2007), 931-944.
- [14] Hornb, K., and Hertzum, M. The notion of overview in information visualization. *International Journal of Human-Computer Studies* 69, 7-8 (2011), 509-525.
- [15] Westerman, S. J., and Cribbin, T. Mapping semantic information in virtual space: dimensions, variance and individual differences. *International Journal of Human-Computer Studies* 53, 5 (2000), 765-787.
- [16] Chen, C., Cribbin, T., Kuljis, J., and Macredie, R. Footprints of information foragers: behaviour semantics of visual exploration. *International Journal of Human-Computer Studies* 57, 2 (2002), 139-163.
- [17] Agirre, E., and Lopez de Lacalle, O., with Soroa, A., Stevenson, M., Fernando, S., Aletras, N., Kukurikos, A., and Fernie, K., PATHS D2.1 Processing and representation of Content for the first prototype, download from [<http://www.paths-project.eu/eng/Resources>]
- [18] Clough, P., Ford, N., and Stevenson, M., Personalising Access to Cultural Heritage Collections Using Pathways, PATCH 2011, pre-print download from [<http://www.paths-project.eu/eng/Resources>]