NAVIGATING CONTEXT, PATHWAYS AND RELATIONSHIPS IN MUSEUM COLLECTIONS USING FORMAL CONCEPT ANALYSIS

RICHARD J. COLE, FRITHJOF DAU, JON DUCROU, PETER W. EKLUND, TIM WRAY

ABSTRACT | The digital medium allows visitors, curators and art historians to gain new insights into their collections through data analysis and rich, interactive visualizations. Motivated by the rise of large-scale cultural heritage collections that have emerged on the Web, we argue that Formal Concept Analysis can be used to highlight the relationships between objects and their features within digital art collections and provide a means for visitors to explore these collections via interactive, narrated pathways. Our work presents four research projects that span 10 years from 2005 - 2015 – ImageSleuth, The Virtual Museum of the Pacific, A Place for Art and a scalability study of Formal Concept Analysis as applied to a data-set from the Brooklyn Museum. Our approach is based on the idea that much of the meaning that can be interpreted from museum collections lies – at least in part – in the way that objects are related to one another. Our work examines how Formal Concept Analysis can drive explorative, narrative-based visitor experiences and reveal new insights into cultural heritage collections.

KEYWORDS | data visualization, collections, interactive, cultural heritage, museum, formal concept analysis

Introduction

Collections are considered to be the heart of the museum. Within collections, artefacts can be interpreted in terms of their intrinsic qualities, properties and context – especially in comparison with other objects (Pearce 1994; Cairns 2013; Skov and Ingwersen 2008; Styliani et al. 2009). Over the last decades, museums have undertaken a massive effort in digitizing their collections and making them available on the Web (Cameron 2001) – for example, The Rijksmuseum provides access to over 640,000 high resolution photographs of their collection[1], and initiatives such as Artsy[2] and the Flickr Commons[3] unite and display massive aggregations of cultural heritage content from multiple institutions. These online collections provide visitors the means to explore, interpret and appropriate cultural heritage objects (Gorgels 2013; Boyd et al. 2006; Colquhoun 2013; Engberg 2017) and are arguably – much like physical museums and galleries – seen as places where visitors can interact, explore and experience cultural heritage (Cameron and Kenderdine 2007).

Over the past 10 years, there has been a growing interest in the use of interactive digital visualizations to support the exploration of these collections (Windhager et al. 2018; Lev Manovich 2011; L. Manovich 2012; Smite, Manovich, and Raitis 2015). As early as the 1990s, emerging perspectives examined how the digital medium can be used to provide new ways of presenting and contextualizing cultural heritage objects (Schweibenz 1998; Hoptman 1992). More recent work has looked at how it can support the needs of expert and casual users in browsing and exploring these collections (Skov 2009; Goodale et al. 2012; Hall et al. 2012). The rise of cultural heritage data on the Web can be used to reveal new views and insights into collections through the use of interactive data visualizations (Windhager et al. 2018; Dörk, Carpendale, and Williamson 2011; Dörk, Carpendale, and Williamson 2012). Rather than being seen as mere repositories of information, the visual affordances of the digital medium – and in particular Web-based interfaces – can be also be used to highlight the scale, complexity and richness of cultural heritage collections (Whitelaw and Hinton 2010; Ennis-Butler, Hinton, and Whitelaw 2011; Whitelaw 2015). Emerging perspectives also incorporate the importance of play (Huizinga 1971; Gaver 2002) and serendipity (Ramsay 2010; Thudt, Hinrichs, and Carpendale 2012; McCay-Peet and

Toms 2013) as a means of interacting with cultural heritage content with the view that such interfaces should afford casual, curious and creative information seeking behavior (Dörk, Carpendale, and Williamson 2011).

In our research, we explore one particular aspect of information seeking within digitized museum collections – the notion of the curated pathway (Goodale et al. 2012). Taking on the view that the digital medium offers the ability to explore rich "landscapes" of content (Murray 2012; Rudolph Wille 1999; Goldfarb et al. 2011; Wray, Eklund, and Kautz 2013; Clough, Stevenson, and Ford 2009) and the motivation for a story-telling based approach of presenting cultural heritage (Segel and Heer 2010), we argue that the digital medium offers the opportunity to showcase collections in the form of linked narratives – linkages that traverse within and across collections that are driven by the visitor's own exploration and curiosity. Based on over 10 years of research, our work consists of the application of a data science technique called Formal Concept Analysis to create these pathways, and a series of design experiments and associated findings that demonstrate exploratory pathway-based navigation of digital cultural heritage collections.

The paper is structured as follows: first, we present the motivation of our work that describes the state-of-the-art in interactive data visualization for cultural heritage collections that supports the motivation our chosen data science technique – Formal Concept Analysis. Next, an overview of the technique is provided, along with a series of designs that demonstrate its application to cultural heritage collections through Web-based and gesture-driven visualization systems, and an assessment of how our approach could scale to larger cultural heritage collections. Brief results of user evaluations of each design are presented, along with design insights on how the approach of Formal Concept Analysis can be used to provide serendipitous pathway-based interactions within digital cultural heritage collections.

Narrative-based Navigation within Cultural Heritage Collections

There is a growing body of work that looks at the way the digital medium can be used to support the exploration and discovery of cultural heritage collections (Thudt, Hinrichs, and Carpendale 2012; E. Ruffaldi et al. 2008; Whitelaw 2015; Clough, Stevenson, and Ford 2009; Dork, Carpendale, and Williamson 2011). In a survey that examined over 70 visualization systems of cultural heritage content (Windhager et al. 2018), most supported either one or more of the following interactions: the ability to provide a high-level orientation or overview of the collection; the ability to explore objects across a specific dimension and the ability to navigate across the collection while providing the means to 'zoom in' and view a single object or category. However, relatively few systems – only around 20% – support the means of navigating a collection in the form of curated pathways, which is an approach that allows the visitor to visualize and explore the collection as a narrative of objects arranged in a specific sequence or theme. Rather than offering fixed, curatorially defined views of the collection, such "virtual museums" have the potential to offer visitors the ability to follow their own interests and perspectives on the works (Davis 1994; Hooper-Greenhill 1994). Existing research has examined how systems can offer the means to explore and diverge within a cultural collection via the use of algorithmically derived recommendations (Toms and McCay-Peet 2009; McCay-Peet and Toms 2011; McCay-Peet and Toms 2013) while studies based on the information seeking needs of users have suggested the notion of the pathway as means of structuring content and affording creative exploration within cultural heritage collections (Goodale et al. 2012; Hall et al. 2012).

A common theme within visualization research centers around the idea of spatial, exploratory navigation – a concept that was first posited by O'Day and Jeffries (1993) in the way library professionals use multiple, interconnected searches to explore topics in an undirected fashion. This closely ties with the commonly associated information seeking notion of exploratory search that recognizes discovery, learning and serendipity as a significant part of the search process (Marchionini 2006). Similarly, the 'information landscape' metaphor is a concept that's used to describe the spatial navigation of digital spaces (Murray 2012), creative and curious information seeking based on the notion of the city flaneur (Dörk, Carpendale, and Williamson 2011) and as a visual representation for 3-dimensional interactive visualizations (Andrews, Pichler, and Wolf 1996) as applied to cultural heritage content (Emanuele Ruffaldi et al. 2008). In the same way, Formal Concept Analysis can be used a means of linking objects and creating structured knowledge hierarchies – 'landscapes of knowledge' (Rudolph Wille 1999; Jon Ducrou 2007) that can be traversed and explored by the end user. Further, the application of data science and media visualization techniques to cultural heritage collections makes it possible to view familiar materials, objects and archival data in new ways (L. Manovich 2012) – it allows the ability to "think with data", as we can use its techniques to define and explore a feature space (Lev Manovich 2015).

In our work, we argue that Formal Concept Analysis allows for this kind of "thinking" as it is heavily based on the philosophical logic of human thought (R Wille 2005). The technique formalizes the way we perceive data through the

	PAINTING	ABSTRACT	COARSE BRUSH STROKES	GEOMETRIC PATTERNS	PRINT	SCULPTURE
Waiting, Port Kembla	×		×			
Bush Rocks After the Rain	×	×	×			
Port Kembla Landscape	×	×		×		
Large Jug		×		×	×	
Gateway to Mt. Keira		×		×		×
Solar Boat	×	×	×	×		

Table 1. A formal context containing information about artworks and their attributes.

construction of concepts – individual units of thought – and their relationship to one another. We argue that such relationships can be visualized in a meaningful way, allowing users to interactively explore and drive their own narrative through a cultural heritage collection.

Formal Concept Analysis

Formal Concept Analysis was developed in the early 1980s as a mathematization of the human cognitive constructs of concepts and concept hierarchies (Rudolph Wille 1999). As an unsupervised clustering mechanism, the technique is commonly used as a means of data analysis, knowledge representation and visualization (Godin, Pichet, and Gecsei 1989; Carpineto and Romano 2004). In our work with cultural heritage collections, we also use it as a means of creating a structure that can be navigated or explored by an end user. Although this section presents the mathematical foundations of Formal Concept Analysis and their relevance to data science and digital art history, readers can consult Davey and Priestly (2002) for a gentle introduction to Formal Concept Analysis, Capineto & Romano (2004) for a more depth analysis for the computer science audience and Wille & Ganter (1999) for complete coverage of the mathematical foundations and principles.

Formal Concept Analysis is based on the philosophy of human thought and communication. Beginning with an understanding of its cognitive constructs, concepts can be understood as basic units of thought created by observations of existing phenomena "formed in dynamic processes within social and cultural environments" (R Wille 2005, 2). According to its definition (Wille and Ganter 1999), a concept consists of a set of objects as its extension, and all attributes, properties and meanings that apply to those objects as its intension. As an example, if one considers the descriptor of "abstract paintings with geometric patterns" (its intension) and the actual 7 paintings that fit that description (its extension), then a concept is defined as the simultaneous instantiation of that intension and extension – i.e., the qualities of those paintings as attributes (or formal attributes) and the actual paintings as objects (or formal objects) defined by those attributes.

In Formal Concept Analysis, concepts are mathematized as formal concepts defined as a pair (A, B) with a set of objects A (its extension) and a set of attributes B that describe those objects (its intension). In context of a museum collection, a formal concept (A, B) can be used to circumscribe a set of objects that possess attributes or meta-data . For example, the following is a formal concept (A, B) that describes a set of works from the University of Wollongong Art Collection[4], where represents a set of titles of 3 works from the collection, and represents a set of attributes that describe those works:

A={Bush Rocks After the Rain, Solar Boat, Port Kembla Landscape }
B={ abstract, painting }

Formal Concept Analysis typically works within a formal context that contains a fixed set of objects or attributes that is used to define the scope of analysis. When working with cultural heritage data, the formal context could contain the entire data-set or a subset of those objects, such as a group of objects belonging to a particular artist, medium or time period. 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 has attribute " and is often written as gIm. A formal context can be represented as a cross-table where the rows represent G, the columns represent M and the incidence relation is represented by a series of crosses.

An example of a formal context is shown in Table 1: the object set G contains the 6 artworks from the University of Wollongong Art Collection, the attribute set M contains attributes of those artworks that indicate their type, medium and characteristics, and a '×' at the intersection of object-row / attribute-column indicates that the object possesses the attribute.

In Formal Concept Analysis, the process of deriving formal concepts begins by clustering related sets of objects based on their common attributes. This is done through the use of derivation operators. The derivation operator A' as applied to a set of objects A is used to retrieve all attributes of a given object, whereas the derivation operator B' as applied to a set of attributes B is used to retrieve all objects of a given attribute. For example, in the above formal context shown in Table 1, if

A={ Bush Rocks After the Rain }

then

A'={ painting, abstract, coarse brush strokes }

dually, if

B={ coarse brush strokes }

then

B'={ Waiting,Port Kembla,Bush Rocks After the Rain,Solar Boat }.

A formal concept can be computed by taking an attribute and collecting all objects that describe it, and then again collecting all common attributes shared by those objects. For example, starting with the attribute set *B*, the set of objects B' is { Waiting,Port Kembla,Bush Rocks After the Rain,Solar Boat }. The set of all attributes common to those artworks as defined by B'' is { coarse brush stroke,painting } and together these two sets represents the following formal concept, which can be described in natural language as "paintings with coarse brush strokes":

({ Waiting,Port Kembla,Bush Rocks After the Rain,Solar Boat },
{ coarse brush stroke,painting })

The process can also work with objects: a formal concept can also be computed by taking an object (or a set of objects) and collecting all attributes that describe it, and then again collecting all common objects shared by those attributes. Formally, is a formal concept of formal context if:

$$A\subseteq G,\ B\subseteq M,\ A'=B, \text{ and } B'=A.$$

The set A is called the extent and B the intent of the formal concept (A, B). Given this definition, A' represents the intent of the concept (A, B) written as (A, A'). Furthermore, A'' is the smallest extent containing A. Consequently, A $\subseteq$ G is an extent if A'' =A. Similarly, B $\subseteq$ M is an intent if B''=B. Within the formal context, a formal concept represents a maximal rectangle and the set of all formal concepts of (G,M,I) is P (G, M, I) or $\mathrm{B}(\mathrm{K})$.

The set of all formal concepts derived from Table 1, and rendered as a concept lattice in Figure 1, is shown as follows:

1. ({ [all objects] }, { })
2. ({ Waiting, Port Kembla, Bush Rocks After the Rain, Solar Boat, Port Kembla Landscape }, { painting })
3. ({ Bush Rocks After the Rain, Solar Boat, Port Kembla Landscape, Large Jug, Gateway to Mt. Keira }, { abstract })
4. ({ Waiting, Port Kembla, Bush Rocks After the Rain, Solar Boat }, { painting, coarse brush strokes })
5. ({ Bush Rocks After the Rain, Solar Boat, Port Kembla Landscape }, { painting, abstract })
6. ({ Port Kembla Landscape, Large Jug, Gateway to Mt. Keira }, { abstract, geometric patterns })
7. ({ Bush Rocks After the Rain, Solar Boat }, { painting, abstract, coarse brush strokes })
8. ({ Solar Boat, Port Kembla Landscape }, { painting, abstract, geometric patterns })
9. ({ Large Jug }, { abstract, geometric patterns, print })
10. ({ Gateway to Mt. Keira }, { abstract, geometric patterns, sculpture })
11. ({ Solar Boat }, { painting, abstract, coarse brush strokes, geometric prints })
12. ({ }, { [all attributes] })

Through the process of generating formal concepts, one can already gain insights about the collection as objects that share the same features are clustered. For example, entry #8 indicates that artworks titled 'Solar Boat' and 'Port Kembla Landscape' are similar because they are both abstract paintings with geometric patterns. In entry #9, the artwork that has coarse brush strokes is also co-associated with the attribute 'painting'. These two entries provide different views, or different ways one would "think" about the collection of artworks. In addition, concepts can also be "more broad" or "more specific" than one another – this is described via the subconcept-superconcept relation.

The subconcept-superconcept relation is defined by the following:

$$(A_1,B_1)\leq(A_2,B_2)\iff A_1\subseteq A_2\wedge B_2\subseteq B_1$$

As above, (A_1, B_1) is called a subconcept of (A_2, B_2) and conversely, (A_2, B_2) is a superconcept of (A_1, B_1). Subconcepts are said to less general (≤) than their superconcepts and the superconcepts more general (≥) than their subconcepts. For the set of concepts $\mathrm{B}(\mathrm{K})$ there is always a greatest subconcept and a smallest superconcept. $\mathrm{B}(\mathrm{K})$ together with the order relation (≤) forms a complete lattice $\BB(\KK)$. $\BB(\KK)$ is called the concept lattice of K.

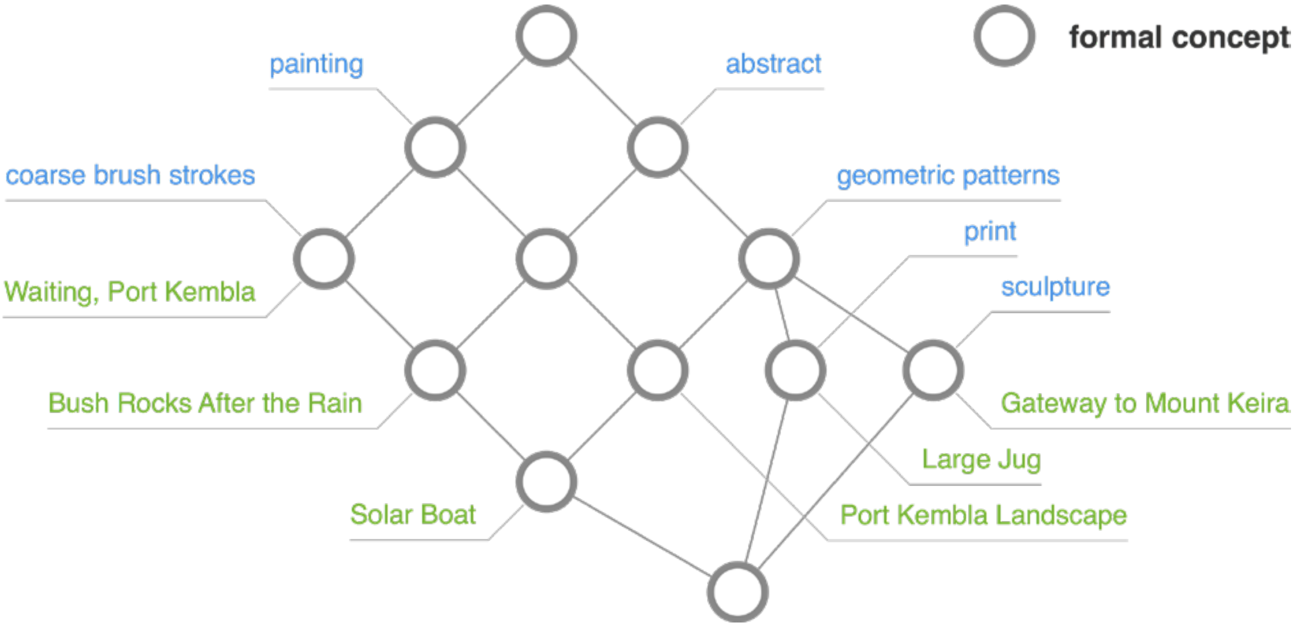

Figure 1. A concept lattice that describes artworks and their attributes as derived from the formal context in Table 1.

The concept lattice line diagram provides a visual representation of the concepts and their concept hierarchies. In Figure 1 for example, one can for example see a clear difference between the geometric, print and sculptural works clustered to the right-hand side of the diagram and the work titled 'Waiting, Port Kembla' to the left, which is distinctively different as it is a painting that has coarse brush strokes. We can also infer that the works title 'Large Jug' and 'Gateway to Mt. Keira' are similar by virtue of their proximity on the diagram and the fact that they both abstract works with geometric patterns. The construction of concept lattices allows the ability to construct categories, hierarchies and infer implicit relationships between objects, even though such relationships are not explicitly specified. Rather than work with externally defined boundaries and categories, Formal Concept Analysis examines the commonalities and differences between objects within a collection and draws implicit associations between such. It also creates a structure, an "information landscape" of categories and concepts (Rudolph Wille 1999) based on cultural heritage content (E. Ruffaldi et al. 2008) that can be driven by the visitor's own interaction and exploration of the content. The following sections demonstrate data visualization systems and prototypes that afford this kind of pathway and narrative-based exploration within cultural heritage collections.

Navigating Images Using Formal Concept Analysis

Image collections exist in many forms: from collections of digital photographs to large-scale collections of museum archives. Irrespective of the type of collection, an established method for browsing images is to display them as thumbnails (Samadani, Lim, and Tretter 2007; Lev Manovich 2011). Thumbnails are usually displayed in a fixed, grid layout. They are a useful means of displaying and aggregating visual-based media as the image itself is considered to be a key component of visualization (L. Manovich 2012).

Early experiments examined how Formal Concept Analysis can not only provide a means of data analysis and visualization, but also provide a way of navigating and exploring image and document collections. The first experiments with concept-lattice based navigation began with email document browsing (Cole, Eklund, and Stumme 2003) that were subsequently applied to images & DVDs covers (J Ducrou 2007). ImageSleuth, shown in Figure 2, was developed as a way navigating collections of annotated images and for content-based retrieval using Formal Concept Analysis, as it combines the data science and unsupervised clustering mechanisms of the technique with the visual

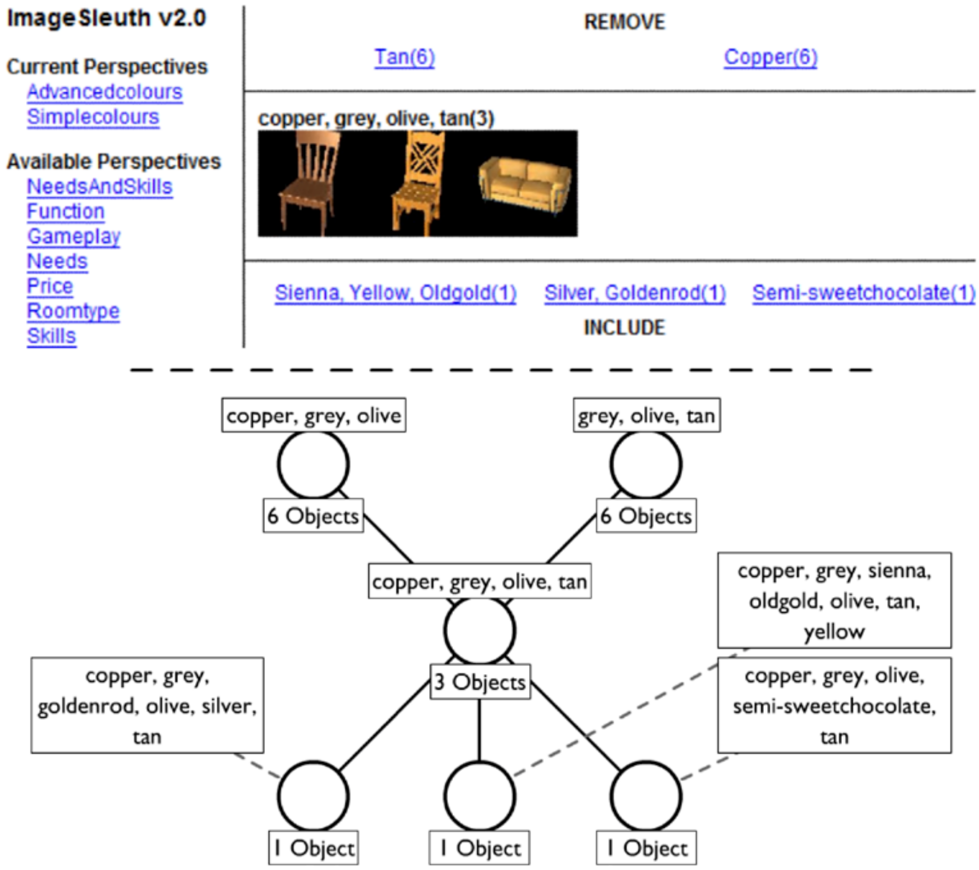

Figure 2. A screenshot of ImageSleuth and the lattice representation of the corresponding neighborhood showing its main navigation functions. Users can narrow or broaden their view on the collection by including or removing attributes from their current view or select the facets of their search via the use of perspectives so that users can navigate by color, function or price (or any combination thereof).

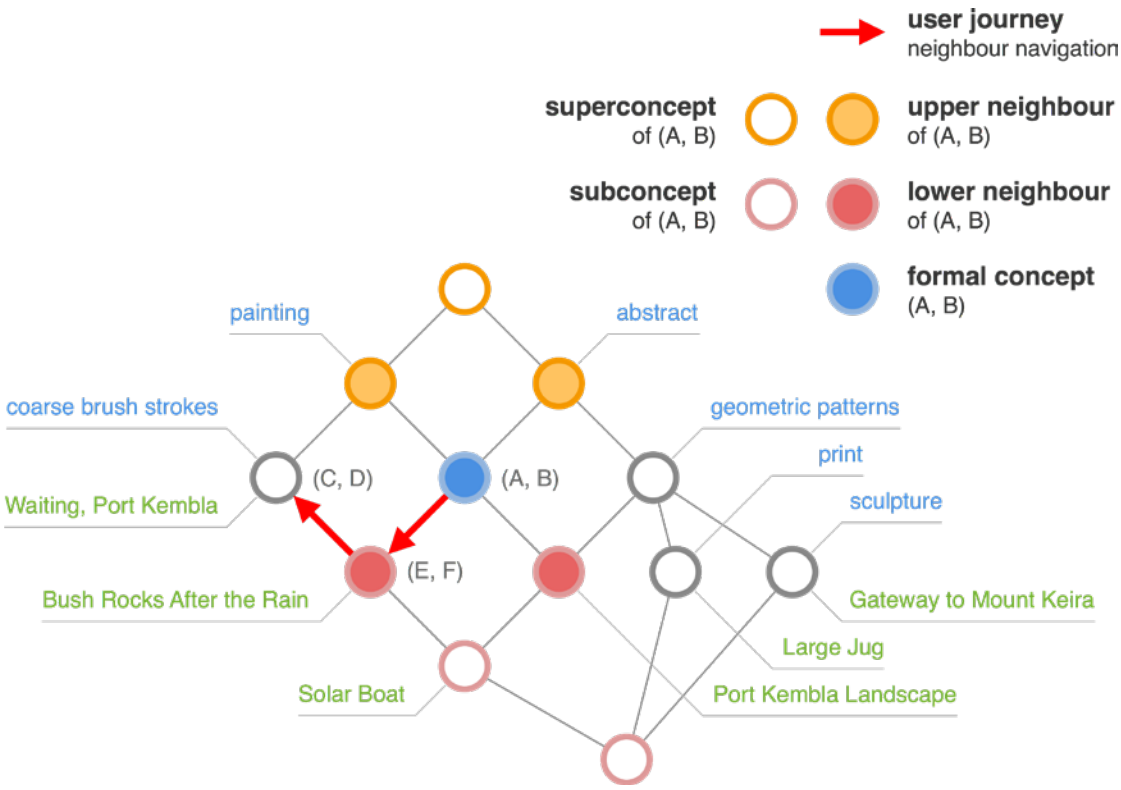

Figure 3. A concept lattice that describes artworks and their attributes as derived from the formal context in Table 1, demonstrating step-wise traversal of neighboring concepts.

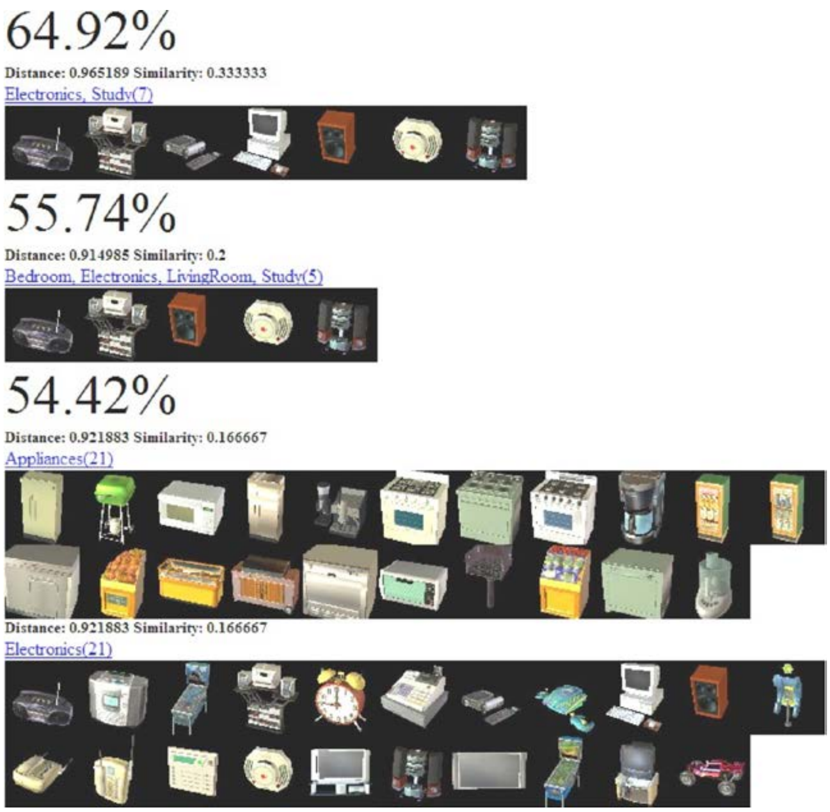

Figure 4. Results of a concept similarity search in ImageSleuth based on the "Appliances, Electronics, Study" concept, showing distance and similarity for matching formal concepts.

affordances of thumbnail displays and image mosaics (L. Manovich 2012). As it is the first program to depart from the use of the concept lattice line diagram (Figure 1) to visualize an information space (J Ducrou, Vormbrock, and Eklund 2006), the program demonstrated a number of interactions that supported the information activities of overview, orientation and step-wise navigation of a collection.

The original prototype of ImageSleuth was built using a data-set from the popular computer game "The Sims 2". It features the ability to explore 412 objects of household furniture and fittings as described by 120 attributes which include in-game properties, suggestions for use and automatically extracted color information. As shown in the bottom half of Figure 2, a user would 'step through' the constructed concept lattice of 7,516 concepts by restricting or broadening their view on the collection and thus 'move' through the information space. For example, in the top half of Figure 2, the program provides a thumbnail view of all objects that are colored copper, grey, olive and tan. At this point, a user could either narrow their focus on the collection and view a subset of these objects by 'including' more attributes at the bottom of the screen or broaden their view on the collection and explore related, but less specific objects by 'removing' attributes as displayed at the top of the screen. This process of 'including' and 'removing' represents a step-wise traversal through the information space generated by the concept lattice.

Figure 3 demonstrates how this step-wise traversal allows the ability to explore related and orthogonal concepts within a concept lattice. In this example, a user may initially start at concept (A, B) which represents all abstract paintings. The user may then "move down" within the lattice to concept (E, F) by including the attribute 'coarse brush strokes', taking them to the more specific concept of 'abstract paintings with coarse brush strokes'. Finally, the user would again "move up" to the broader concept (C, D) by excluding the attribute 'abstract', ending at the final concept of 'paintings with coarse brush strokes'. Throughout the process, the user encounters works titled 'Bush Rocks After the Rain' and 'Waiting, Port Kembla', given that both of these works share similar features. Each step through the concept lattice takes the user at a slightly different point within the information space, allowing them to walk a pathway through a conceptual map of the collection.

ImageSleuth also introduced a number of different features that supported more conventional kinds of information seeking activities, such as faceted and similarity-based search (J Ducrou, Vormbrock, and Eklund 2006; Jon Ducrou 2007). Users could select 'perspectives' (shown on the left in Figure 2) that allowed them to define how they wanted to explore the collection by restricting and combining subsets of attributes within the formal context. For example, by selecting the 'Simplecolours' perspective, a user could navigate the generated hierarchy of concepts based only on their color, or they could combine this with the 'Function' and 'Price' perspectives if they also wanted to explore those facets of the information space. The application also offered a search functionality, so that it can return a set of matching formal concepts based on keywords – for instance, the keywords 'red' and 'lounge' returned concepts that clustered "all red lounges", and a query-by-example feature returned matching formal concepts based on their similarity to a single object. The system also offered a concept similarity measure (Saquer and Deogun 2001), so that it could take a given single concept such as "electronics used to support study" and provide a list of order-ranked concepts that are similar to those group of objects, matching the concepts "electronics used to support study in the bedroom and living room", "appliances" and "electronics" (Figure 4).

Navigating Ethnographic Collections using The Virtual Museum of the Pacific

The Virtual Museum of the Pacific was an interactive visualization (Eklund et al. 2012) that was developed to provide access to 427 ethnographic objects from the Australian Museum's cultural collections. It provided access to high resolution imagery and accompanying interpretive texts of ethnographic objects originating from the Papua New Guinea and Pacific Island nations. The visualization was to develop to assess how it can support exploratory, non-hierarchical and pathway-based navigation of an ethnographic cultural collection.

The Virtual Museum of the Pacific (Figure 5) is based on the design of ImageSleuth, and inherited many of its features – namely, the ability to navigate a pathway through a collection based on a traversal of related concepts that are formed based on a museological view of cultural heritage. It also supported a tagging interface that allowed creator communities to add their own tags to the collection and therefore not only influence the content and interpretation of the objects, but also the way that they are related to one another via the mechanisms of Formal Concept Analysis.

A significant factor that informed the design and presentation of Pacific material culture within the Virtual Museum of the Pacific was based on the process of selecting objects. Working with the Australian Museum, a project anthropologist was commissioned to select 427 objects from the museum's collection of approximately 60,000 Pacific objects. Objects were selected with similar and overlapping qualities, so that the Virtual Museum of the Pacific could convey new and serendipitous connections via its conceptual-based browsing paradigm. This was done by selecting multiple types of objects from different categories (body accessories, ceremonial objects, ornaments etc.) without restricting each category to objects from a particular provenance, or to objects made with a particular material. This ensured that the Virtual Museum of the Pacific could represent a diverse cross-section of Pacific material culture while also providing opportunities for exploration by linking objects across multiple facets and dimensions.

Two user experience evaluations were conducted for the Virtual Museum of the Pacific: one based on 16 staff members from the Australian Museum with backgrounds in collection management and education anthropology and another based on three demographic groups: researchers and librarians, regular users of social media, and participants who identified themselves as Pacific Islander. All participants generally praised the use of prominent visual imagery – a claim supported by prior research (Skov and Ingwersen 2008) – along with the ability to 'group' related objects together. However, participants from both groups also commented that the interface was relatively difficult to learn and added an extra layer of complexity – given the high number of navigation options – with many stating that at times they felt 'lost' within the collection. While the Virtual Museum of the Pacific provided a browsing interface that showcased many aspects of Pacific culture, this led to the development of future prototypes that aimed to provide a simpler interaction paradigm based on the notion of a pathway.

Navigating Pathways with A Place for Art

A Place for Art is an iPad app that publicly displays 77 paintings, works on paper and Australian First Nation works from the University of Wollongong Art Collection. It allows users to explore the collection by using gesture-based interactions to navigate pathways of conceptually related content. As a case study, it examines how an exploratory visualization with a highly responsive and tactile interface can facilitate discovery and serendipity in a collection of artworks that, for most participants and target users of this app, surround them in the built environment of the university campus environment. The app is also a digital counterpart to an exhibition book – also called 'A Place for Art' (Lawson 2012) – that uses the same artworks but exhibits them in a completely non-linear way.

The app's interaction design departs significantly from that of the Virtual Museum of the Pacific by presenting the works as a visible, sliding horizontal pathway (Figure 6). The pathway is

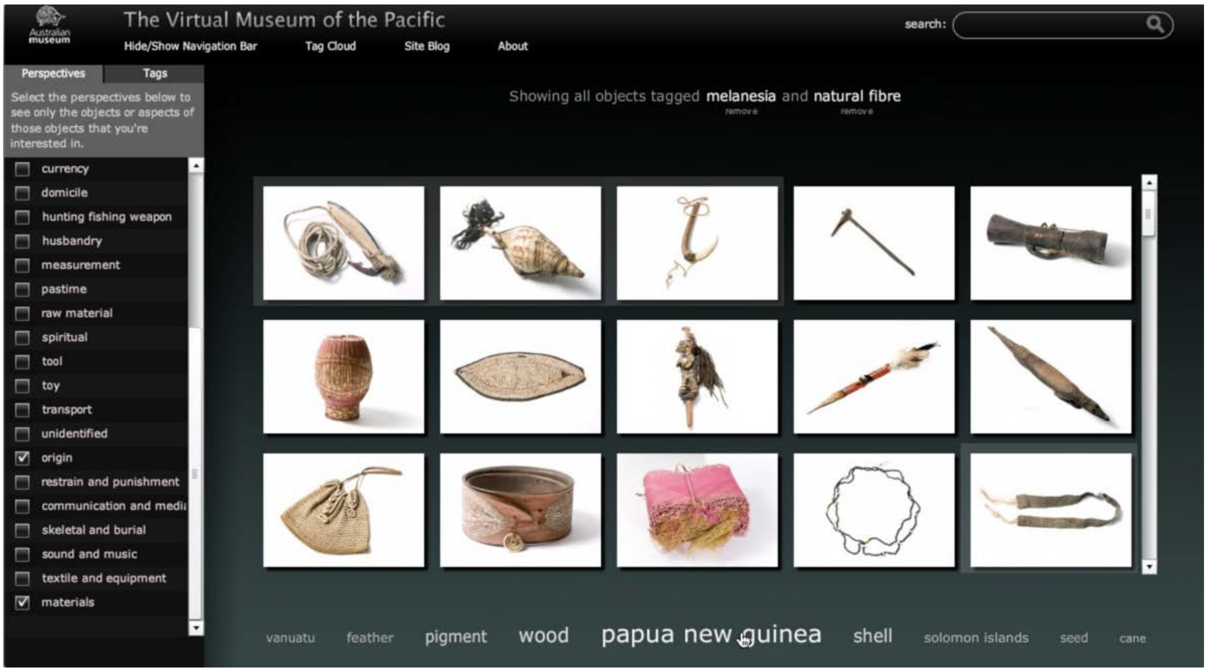

Figure 5. The Virtual Museum of the Pacific allowed users to explore the ethnographic Pacific collections of the Australian Museum.

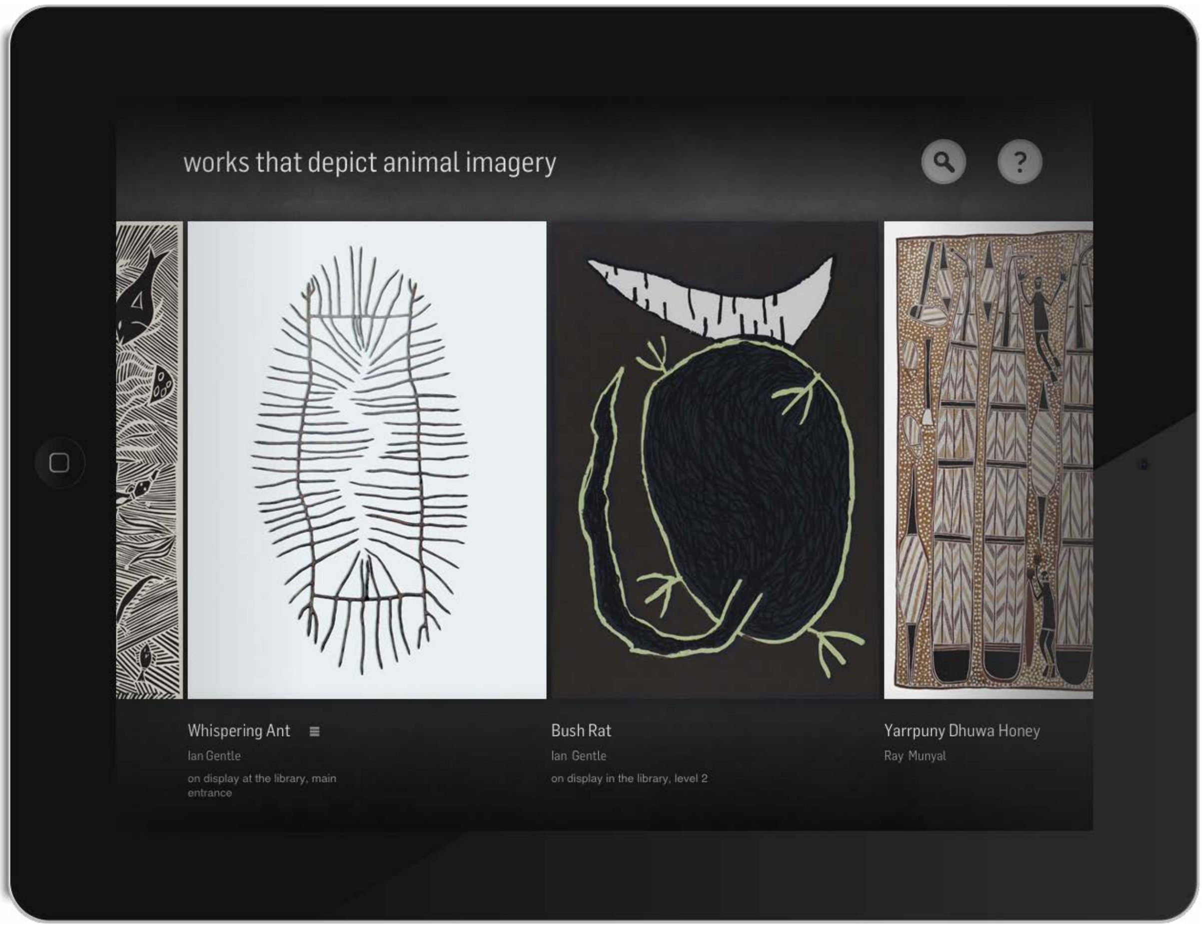

Figure 6. The A Place for Art iPad app allows users to explore the collection as pathways of related concepts.

Figure 7. The A Place for Art iPad app allows users to explore the collection as pathways of related concepts.

Figure 8. Salvatore Zofrea, 'Illawarra Flame Tree and Bowerbird': one of the works in the 'A Place for Art' collection.

FORMAL CONCEPT, EXPRESSED IN NATURAL LANGUAGE	SIMILARITY	# OBJECTS
intricate prints that depict animal imagery and nature	0.50	2
works that depict animal imagery and have red tones	0.38	2
intricate works that depict nature	0.21	3
works that have blue tones	0.13	8
works that depict animal imagery	0.11	11
vibrant works	0.08	23

Table 2. Concepts similar to 'Illawarra Flame Tree and Bowerbird', showing the multiplicity of contexts that the object appears in.

depicted as a horizontal lateral gallery of images that represent thematic concepts derived from the metadata of the artworks. Users can move back and forth along the pathway by swiping left and right with their finger, or they can tap on an individual work to view a larger image and read its label.

Exploration of the collection is done primarily via similarity-based interaction, so that a visitor can explore related works based on the themes and associations of works that they are interested in. Users can also 'branch out' and explore related pathways by double-tapping on a single work – for example, double tapping on the work titled 'Illawarra Flame Tree and Bowerbird' (Figure 8) would branch the user to another pathway of related works to that artwork, such as 'intricate prints that depict animal imagery and nature', 'works that depict animal imagery and have red tones', etc. The app employs the use of visual overlays to convey the sense that they are manipulating the pathway structure visually (Figure 7), with the intention that the user can directly manipulate the pathway structure and therefore drive their own narrative and exploration throughout the collection.

Each pathway within the app represents a single formal concept with its objects represented by the artworks and its attributes represented by its title. The app employs the use of natural language processing and expression so that formal concepts are represented as statements rather than attribute sets – for instance, a user could be viewing a set of objects with attributes 'painting' and 'the Illawarra' expressed as "paintings that depict the Illawarra", and then explore the related concept "other works that depict social critique and the Illawarra", containing the related attribute set 'social critique' and 'the Illawarra'.

The app uses concept similarity (Saquer and Deogun 2001) to highlight the relationships between objects within the collection. Referring to the example of the print work 'Illawarra Flame Tree and Bowerbird' (Figure 8) which is, according to the natural language description of its object concept, "an intricate and vibrant print that depicts animal imagery, the Illawarra and nature and has red and blue tones", we can observe the multiplicity of contexts that this object can be interpreted in and determine which concepts are 'most similar' to the artwork (shown in Table 2). A user navigating from this work can explore another pathway of related works in any number of contexts, depending on the set of works that they have already viewed within the collection.

The app underwent a user experience evaluation that was conducted with 24 participants under the age of 35 who had an interest in visiting museums and art galleries. While participants praised the aesthetics, visual design and 'sense of flow' conveyed by the app's interaction, along with 'categoryless' nature of presentation, participants noted that the app also needed to convey a sense of overview of the collection and provide more conventional navigation options, such as the ability to view a list of artists or browse by medium – a finding supported by related research in the field of interactive visualizations for cultural heritage collections (Dörk, Carpendale, and Williamson 2012; Dörk, Carpendale, and Williamson 2011; Whitelaw 2015). In designing A Place for Art, we determine how users can interact with narrative path-based exploration of a digital cultural heritage collection, and how the approach applies to a collection of print works, paintings and sculptures that are distributed on university-wide campus environment.

Scalability of Approach with the Brooklyn Museum Collection

The ImageSleuth, Virtual Museum of the Pacific and A Place for Art case studies experimented with relatively small collections containing no more than 500 objects. However, online museum collections often showcase tens-of-thousands or even hundreds-of-thousands of objects. Therefore, it is important assess the scalability of the Formal Concept Analysis algorithms with respect to much larger data-sets. To explore this issue, we assessed the browsing framework's performance on 15,000 objects from the Brooklyn Museum's online collections.

Traditional visualization approaches for Formal Concept Analysis rely on computing and displaying the complete concept lattice, as shown in Figure 1. For relatively small data-sets covering dozens of objects and attributes, this approach is an effective and sufficient means of conveying the connections among objects within an information space. However for larger data-sets approaching hundreds of objects, this approach becomes problematic as the number of formal concepts generated (and hence the complexity of the displayed concept lattice) is quadratic with respect to the number of objects (Carpineto and Romano 2004). This poses obvious scalability and complexity issues for larger formal contexts such as those generated from museum collections.

ImageSleuth, The Virtual Museum of the Pacific and A Place for Art uses the NearestNeighbours algorithm to implement the conceptual neighborhood approach (Eklund, Ducrou, and Brawn 2004) that computes a partial view of the concept lattice in the form of a single formal concept and its immediate neighbors. This approach overcomes the scalability limitations of computing a complete concept lattice. In these applications the conceptual neighborhoods are computed dynamically at query-time, so that any changes to the underlying formal context are immediately reflected in its underlying concept lattice. This allows the collection and the relationships among the objects to dynamically respond to user tagging and other changes. However, we find that this approach does not scale for the 15,000 objects from the Brooklyn Museum's data-set. To overcome this limitation, we have implemented a system that pre-computes and caches the conceptual neighborhoods of all possible formal concepts from the collection of 15,000 objects, and then employ a fast incremental algorithm that updates a set of formal concepts without the need to compute a complete lattice (Outrata 2013). When new objects are added to the collection, we update the cache of conceptual neighborhoods as a background process, allowing the system to scale to a much larger collection of objects.

Conclusion and Summary

During the last 10 years our team has worked on case studies that examine how Formal Concept Analysis could be used to analyze, visualize and browse digital art collections. Our research contributes to the design and implementation of a story-telling based view of the collection within visualization systems – views that combine the traditional presentation of cultural heritage objects in the form of 'exhibitions' combined with the flexibility of allowing visitors to choose their own narrated path through the collection. The motivation of this approach was based on the idea that, by being able to visualize and browse a collection based on a method that is rooted heavily in the philosophical logic of human thought, one could meaningfully gain insights about the connections that objects have with one another. Our work has examined how Formal Concept Analysis could be used to support the exploration of image-based collections; how it could be used to provide a means of linking objects together within a museum; and how it could be used to show the subjective relationships between works of art. We assess the scalability of our approach to a large museum data-set. Our results demonstrate that Formal Concept Analysis can be used as a data science technique in digital art collections - one that aims to draw the implicit relationships between objects, artworks and their features.

NOTES

[1] https://www.rijksmuseum.nl/en/rijksstudio
[2] https://www.artsy.net/
[3] https://www.flickr.com/commons
[4] https://lha.uow.edu.au/uowac/index.html

BIBLIOGRAPHY

Anderson, Gail. 2004. *Reinventing the Museum: Historical and Contemporary Perspectives on the Paradigm Shift*. Rowman Altamira.

Andrews, Keith, Michael Pichler, and Peter Wolf. 1996. "Towards Rich Information Landscapes for Visualising Structured Web Spaces." In Proceedings of the 2nd IEEE Symposium on Information Visualization, InfoVis '96, 62–63. San Francisco.

Boyd, D, M Davis, C Marlow, and M Naaman. 2006. "HT06, Tagging Paper, Taxonomy, Flickr, Academic Article, To Read." In Proceedings of the ACM Conference on Hypertext and Multimedia. Odense, Denmark.

Cairns, Susan. 2013. "Mutualizing Museum Knowledge: Folksonomies and the Changing Shape of Expertise." Curator: The Museum Journal 56 (1): 107–19. doi:10.1111/cura.12011.

Cameron, Fiona. 2001. "Wired Collections - the next Generation." Museum Management and Curatorship 19 (3): 309-- 312.

Cameron, Fiona, and Sarah Kenderdine. 2007. Theorizing Digital Cultural Heritage: A Critical Discourse. MIT Press.

Carpineto, C., and G. Romano. 2004. Concept Data Analysis: Theory and Applications. J. Wiley.

Clough, Paul, Mark Stevenson, and Nigel Ford. 2009. "Personalizing Access to Cultural Heritage Collections Using Pathways."

Cole, Richard, Peter Eklund, and Gerd Stumme. 2003. “Document Retrieval for E-Mail Search and Discovery Using Formal Concept Analysis.” Applied Artificial Intelligence 17 (3): 257–80.

Colquhoun, Brooklyn. 2013. “Making Sense of Historic Photographic Collections on Flickr The Commons: Institutional and User Perspectives.” In Museums and the Web 2013: Proceedings, edited by Nancy Proctor and Rich Cherry. Portland, OR: Archives & Museum Informatics. http://mw2013.museumsandtheweb.com/paper/making-sense-of-historic-photographic-collections-on-flickr-the-commons-institutional-and-user-perspectives/.

Davey, Brian, and Hilary Priestley. 2002. Introduction to Lattices and Order. Cambridge University Press.

Davis, B. 1994. “Digital Museums.” Aperture, no. 136: 68–70.

Dork, M., S. Carpendale, and C. Williamson. 2011. “Visualizing Explicit and Implicit Relations of Complex Information Spaces.” Information Visualization 11 (1): 5–21. doi:10.1177/1473871611425872.

Dörk, Marian, Sheelagh Carpendale, and Carey Williamson. 2011. “The Information Flaneur: A Fresh Look at Information Seeking.” In Proceedings of the 29th SIGCHI Conference on Human Factors in Computing Systems, 1215–24. doi:10.1145/1978942.1979124.

———. 2012. “Fluid Views: A Zoomable Search Environment,” 233–40. http://dl.acm.org/ft_gateway.cfm?id=2254599&-type=pdf.

Ducrou, J. 2007. “DVDSleuth: A Case Study in Applied Formal Concept Analysis for Navigating Web Catalogs.” In Conceptual Structures: Knowledge Architectures for Smart Applications, 496–500. LNCS 4604. Springer-Verlag.

Ducrou, J, B Vormbrock, and P Eklund. 2006. “FCA-Based Browsing and Searching of a Collection of Images.” In Proceedings of 14th International Conference on Conceptual Structures, 203–14. LNAI 4068. Springer.

Ducrou, Jon. 2007. “Design for Conceptual Knowledge Processing: Case Studies in Applied Formal Concept Analysis.” University of Wollongong Thesis Collection. University of Wollongong. http://ro.uow.edu.au/theses/760.

Eklund, Peter, Jon Ducrou, and Peter Brawn. 2004. “Concept Lattices for Information Visualization: Can Novices Read Line-Diagrams?” In International Conference on Formal Concept Analysis, 57–73. Springer.

Eklund, Peter, Tim Wray, Peter Goodall, and Amanda Lawson. 2012. “Design, Information Organisation and the Evaluation of the Virtual Museum of the Pacific Digital Ecosystem.” Journal of Ambient Intelligence and Humanized Computing 3 (4). Springer-Verlag: 265–80.

Engberg, Maria. 2017. “Digital Archives , the Museum and the Culture Snacker.” In Openness: Politics / Practices / Poetics, 14–23. Living Archives. https://livingarchives.mah.se/files/2015/01/openness_final.pdf.

Ennis-Butler, Ben, Sam Hinton, and Mitchell Whitelaw. 2011. “Playing with Complexity: An Approach to Exploratory Data Visualisation.” In Proceedings of the 2011 Conference of the Australian Council of University Art and Design Schools, edited by Gordon Bull.

Gaver, W. 2002. “Designing for Homo Ludens.” I3 Magazine. http://luna.gold.ac.uk/media/27gaver.ludens.02.pdf.

Godin, Robert, C Pichet, and J Gecsei. 1989. “Design of a Browsing Interface for Information Retrieval.” In ACM SIGIR Forum, 32–39. ACM.

Goldfarb, Doron, Max Arends, Josef Froschauer, and Dieter Merkl. 2011. “Revisiting 3D Information Landscapes for the Display of Art Historical Web Content.” Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology - ACE ’11. New York, New York, USA: ACM Press, 1. doi:10.1145/2071423.2071480.

Goodale, Paula, Paul Clough, Nigel Ford, Mark Hall, Mark Stevenson, Samuel Fernando, Nikolaos Aletras, Kate Fernie, Phil Archer, and Andrea De Polo. 2012. “User-Centred Design to Support Exploration and Path Creation in Cultural Heritage Collections.” In Proceedings of the 2nd European Symposium on Human-Computer Interaction and Information Retrieval, i:75-- 78.

Gorgels, Peter. 2013. “Rijksstudio: Make Your Own Masterpiece!” In Museums and the Web 2013: Proceedings, edited by Nancy Proctor and Rich Cherry. Portland, OR. http://mw2013.museumsandtheweb.com/paper/rijksstudio-make-your-own-masterpiece/.

Hall, Mark, Eneko Agirre, Nikolaos Aletras, Runar Bergheim, Konstantinos Chandrinos, Paul Clough, Samuel Fernando, et al. 2012. “PATHS – Exploring Digital Cultural Heritage Spaces.” In Theory and Practice of Digital Libraries, edited by Panayiotis Zaphiris, George Buchanan, Edie Rasmussen, and Fernando Loizides, 7489:500–503. Lecture Notes in Computer Science. Springer. http://www.springerlink.com/index/0UG4X6381VH7Q783.pdf.

Hooper-Greenhill, E. 1992. Museums and the Shaping of Knowledge. Routledge.

———. 1994. “Museum Education: Past, Present and Future.” In Towards the Museum of the Future: New European Perspectives, edited by R Miles and L Zavala. Routledge.

Hoptman, G. 1992. “The Virtual Museum and Related Epistemological Concerns.” In Sociomedia. Multimedia, Hypermedia and the Social Construction of Knowledge, edited by E Barrett. MIT Press Cambridge, MA, USA.

Huizinga, Roy. 1971. Homo Ludens: A Study of the Play Element in Culture. Beacon Press.

Lawson, Amanda. 2012. A Place for Art. University of Wollongong Press. http://www.cornerhouse.org/wp-content/uploads/book-35695.pdf.

Manovich, L. 2012. “Media Visualization.” In The International Encyclopedia of Media Studies. Wiley.

Manovich, Lev. 2011. “Media Visualization: Visual Techniques for Exploring Large Media Collections.” Media Studies Futures, no. June: 1–21. http://www.manovich.net/DOCS/media_visualization.2011.pdf.

———. 2015. “Data Science and Digital Art History.” International Journal for Digital Art History, no. 1.

Marchionini, Gary. 2006. “Exploratory Search.”

Communications of the ACM 49 (4). ACM: 41. doi:10.1145/1121949.1121979.
McCay-Peet, Lori, and Elaine Toms. 2011. “Measuring the Dimensions of Serendipity in Digital Environments.” Information Research 16 (3): 1–7.
———. 2013. “Proposed Facets of a Serendipitous Digital Environment Conceptualization of the Construct.” In IConference 2013 Proceedings, 688–91. doi:10.9776/13318.
Murray, Janet. 2012. Inventing the Medium. Cambridge, MA: MIT Press.
O'Day, Vicki L., and Robin Jeffries. 1993. “Orienteering in an Information Landscape.” In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93, 438–45. New York, New York, USA: ACM Press. doi:10.1145/169059.169365.
Outrata, Jan. 2013. “A Lattice-Free Concept Lattice Update Algorithm Based on FCbO.” In Proceedings of the 10th International Conference on Concept Lattices and Their Applications, edited by Manuel Ojeda-Aciego and Jan Outrata, 261-- 274. La Rochelle, France.
Pearce, S. 1994. “Thinking about Things.” In Interpreting Objects and Collections, edited by S Pearce. London: Routledge.
Ramsay, Stephen. 2010. “The Hermeneutics of Screwing Around ; or What You Do with a Million Books.”
Ross, Max. 2004. “Interpreting the New Museology.” Museum and Society 2 (2): 84–103.
Ruffaldi, E., C. Evangelista, V. Neri, M. Carrozzino, and M. Bergamasco. 2008. “Design of Information Landscapes for Cultural Heritage Content.” In Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts, 113–19.
Ruffaldi, Emanuele, Chiara Evangelista, Veronica Neri, Marcello Carrozzino, and Massimo Bergamasco. 2008. “Design of Information Landscapes for Cultural Heritage Content.” In Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts, 113–19.
Samadani, Ramin, Suk Hwan Lim, and Dan Tretter. 2007. “Representative Image Thumbnails for Good Browsing.” In IEEE International Conference on Image Processing, 193–96.
Saquer, J, and J Deogun. 2001. “Concept Aproximations Based on Rough Sets and Similarity Measures.” International Journal of Applied Mathematics and Computer Science 11: 655–74.
Schweibenz, W. 1998. “The ‘Virtual Museum’: New Perspectives For Museums to Present Objects and Information Using the Internet as a Knowledge Base and Communication System.” In Knowledge Management Und Kommunikationssysteme, Workflow Management, Multimedia, Knowledge Transfer. Proceedings Des 6. Internationalen Symposiums F{ü}r Informationswissenschaft (ISI 1998), edited by H Zimmermann and B Schramm, 3–7.
Segel, Edward, and Jeffrey Heer. 2010. “Narrative Visualization: Telling Stories with Data.” IEEE Transactions on Visualization and Computer Graphicss 16 (6): 1139-- 1148.
Skov, Mette. 2009. “The Reinvented Museum : Exploring Information Seeking Behaviour in a Digital Museum Context.” Royal School of Library and Information Science.
Skov, Mette, and Peter Ingwersen. 2008. “Exploring Information Seeking Behaviour in a Digital Museum Context.” In Proceedings of the Second International Symposium on Information Interaction in Context - IIiX '08, 110. New York, New York, USA: ACM Press. doi:10.1145/1414694.1414719.
Smite, Rasa, Lev Manovich, and Smits Raitis. 2015. Data Drift. Archiving Media and Data Art in the 21st Century. RIXC.
Styliani, Sylaiou, Liarokapis Fotis, Kotsakis Kostas, and Patias Petros. 2009. “Virtual Museums, a Survey and Some Issues for Consideration.” Journal of Cultural Heritage 10 (4): 520–28. doi:10.1016/j.culher.2009.03.003.
Thudt, Alice, Uta Hinrichs, and Sheelagh Carpendale. 2012. “The Bohemian Bookshelf: Supporting Serendipitous Book Discoveries through Information Visualization.” In Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems, 1461–70. http://dl.acm.org/citation.cfm?id=2208607.
Toms, Elaine, and Lori McCay-Peet. 2009. “Chance Encounters in the Digital Library.” In International Conference on Theory and Practice of Digital Libraries, 192–202. Springer.
Trant, J. 2009. “Tagging, Folksonomy and Art Museums: Results of Steve.Museum's Research.”
Weinberger, D. 2005. “Taxonomies to Tags: From Trees to Piles of Leaves.” Release 1.0 23 (2).
Whitelaw, Mitchell. 2015. “Generous Interfaces for Digital Cultural Collections.” DHQ: Digital Humanities Quarterly 9 (1).
Whitelaw, Mitchell, and Sam Hinton. 2010. “CommonsExplorer.” http://mtchl.net/cex/.
Wille, R. 2005. “Formal Concept Analysis as Mathematical Theory of Concepts and Concept Hierarchies.” In Formal Concept Analysis, edited by Bernhard Ganter, Gerd Stumme, and Rudolf Wille, 3626:47–70. Lecture Notes in Computer Science. Springer Berlin / Heidelberg. http://dx.doi.org/10.1007/11528784_1.
Wille, Rudolph. 1999. “Conceptual Landscapes of Knowledge: A Pragmatic Paradigm of Knowledge Processing.” Edited by W. Gaul and H. Locarek-Junge. Classification in the Information Age, 344–56.
Wille, Rudolph, and Bernhard Ganter. 1999. Formal Concept Analysis: Mathematical Foundations. Berlin: Springer-Verlag.
Windhager, Florian, Paolo Federico, Günther Schreder, Katrin Glinka, and Marian Dörk. 2018. “Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges.” IEEE Transactions on Visualization and Computer Graphics.
Wray, Tim, Peter Eklund, and Karl Kautz. 2013. “Pathways through Information Landscapes: Alternative Design Criteria for Digital Art Collections.” In ICIS 2013 Proceedings. Milan, Italy.

RICHARD COLE is a software engineer. Richard's PhD explored the use of formal concept analysis for the organization of text documents. His subsequent postdoc explored the use of formal concept analysis for the identification of "concepts" from text and for the inherent structure implicit in software code bases.

Correspondence e-mail: richard.j.cole@gmail.com

FRITHJOF DAU has a PhD in Mathematics from TU Darmstadt and worked closely with the late Rudolf Wille who invented formal concept analysis. Frithjof now works as a software developer for SAP in Dresden.

Correspondence e-mail: frithjof.dau@sap.com

JON DUCROU has a PhD in Computer Science from the University of Wollongong. Jon's work explored navigation paradigms in information retrieval in formal concept analysis and he has worked at Amazon Inc. since graduation.

Correspondence e-mail: ducrou@amazon.com

PETER W. EKLUND is professor of Artificial Intelligence and Machine Learning at Deakin University. Peter's interests include applied formal concept analysis and its application to text, images and artwork.

Correspondence e-mail: peter.eklund@deakin.edu.au

TIM WRAY is a postdoc and front-end Web developer at the IT University of Copenhagen and holds a PhD in interaction design. Tim's PhD explored new ways of designing rich interactions with digital art collections. Tim is also part of the EU Horizon 2020 project "GIFT: Meaningful Personalization of Hybrid Virtual Museum Experiences Through Gifting and Appropriation": a project that combines the physical and digital aspects of museums to create personal encounters with cultural heritage.

Correspondence e-mail: tiwr@itu.dk