Relevancy Scoring for Knowledge-based Recommender Systems

Robert David[1] and Trineke Kamerling[2]
[1]*Semantic Web Company, Austria*
[2]*Rijksmuseum Amsterdam, The Nederlands*

Keywords: Cultural Heritage, Knowledge Representation, Semantic Web, Information Retrieval, Recommender, Relevancy.

Abstract: Knowledge-based recommender systems are well suited for users to explore complex knowledge domains like iconography without having domain knowledge. To help them understand and make decisions for navigation in the information space, we can show how important specific concept annotations are for the description of an item in a collection. We present an approach to automatically determine relevancy scores for concepts of a domain model. These scores represent the importance for item descriptions as part of knowledge-based recommender systems. In this paper we focus on the knowledge domain of iconography, which is quite complex, difficult to understand and not commonly known. The use case for a knowledge-based recommender system in this knowledge domain is the exploration of a museum collection of historical artworks. The relevancy scores for the concepts of an artwork should help the user to understand the iconographic interpretation and to navigate the collection based on personal interests.

1 INTRODUCTION

Complex knowledge domains like iconography are difficult to understand. Non-experts, who do not have domain knowledge, cannot interpret the symbols to determine the meaning of artwork content. Even for art historians, the analysis of the meaning of an artwork is a difficult task for complex scenes. To illustrate the problem, we present the motivating scenario of museum visitors who want to understand the content of artworks and want to navigate museum collections based on personal interests. However, visitors do not have domain knowledge. We call such a person a non-expert. Having no understanding about the artworks, visitors have to use the limited information that is provided by the museum. Usually, this is a description in natural language and provided as text or as an audio guide. The natural language description makes it difficult for non-experts to understand the details of the scene interpretation, because generic concepts, like religion or mythology, are mixed with specific concepts, like symbolic objects. However, this detail information helps to understand an artwork as well as related artworks. We present an approach to describe visual artwork content using concepts of multiple interlinked vocabularies. The concepts are automatically scored based on relevancy regarding the artworks description. By looking at the concept annotations and the associated relevancy scores, visitors get a better understanding because the most important concepts of the content are pointed out to them. As part of a knowledge-based recommender, these concepts can be selectively used by visitors to navigate artwork collections.

1.1 Knowledge-based Recommender Systems

Knowledge-based recommender systems are well suited for complex knowledge domains (Felfernig and Burke, 2008). The basis of the recommendations is a knowledge model representing domain knowledge that is used for similarity-based retrieval. Therefore knowledge-based recommender systems are not dependent on data from user decisions like other recommendation approaches. Users with only little domain knowledge can still use the system for recommendations based on their interests. They only need to have general knowledge about the set of collection items and an informal knowledge of their needs (Burke, 2000). Because the options for navigation are based on the model describing the items, this approach can be used to help users to better understand the knowledge domain. When navigating the infor-

David, R. and Kamerling, T.
Relevancy Scoring for Knowledge-based Recommender Systems.
DOI: 10.5220/0008068602330239
In *Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2019)*, pages 233-239
ISBN: 978-989-758-382-7

mation space, two navigation decisions can be taken. First, the user selects from the results of similarity-based retrieval (Burke et al., 1997), which will recommend items that are similar to the current item based on the knowledge model. Second, the user can tweak (Burke et al., 1997) the selection based on information needs to alter the resulting recommended items. Knowledge-based recommender systems help users explore and understand an information space (Burke, 2000). Users have to make decisions for retrieval and tweaking based on the model to navigate the information space. Although knowledge-based recommender systems only require general knowledge of the domain, users still might not have enough understanding to decide how to tweak the recommendations to satisfy their information needs. To help users understand and make these tweaking decisions, we can show how significant specific parts of the description are for an item. Knowing the significance is also helpful for understanding the meaning of items in the information space. The significance is expressed as a relevancy score. We focus on scoring concept annotations which are actually instances of concepts (Gicquel and Lenne, 2013) assigned to the items. They represent the semantic context of an item. Showing the significance of concepts therein helps the users make decisions for retrieval to satisfy information needs.

1.2 Information Content

The notion of information content is described in (Resnik, 1995). It uses a hierarchic taxonomy and a corpus where the taxonomy concepts occur to determine the informativeness of concepts. The taxonomic hierarchy represents how general or specific a concept is and therefore can be used to determine how informative it is regarding a description. As probability to encounter a concept in the taxonomy increases, informativeness decreases, so the more abstract a concept, the lower its information content. If there is a unique top concept, its information content is 0 (Resnik, 1995). The information content score is not only based on the taxonomy, but also uses corpus-based statistical information. The frequency of concepts in a corpus is taken into account and is part of the calculation of the probability to encounter a concept. If a concept does not occur in the corpus, its information content is 0. Concepts that do not occur in the corpus do not contribute to the information content score. The information content of concepts can be used to measure the semantic similarity between two concepts based on the taxonomy and the corpus (Resnik, 1995).

2 RELEVANCY SCORING BASED ON INTERLINKING

The calculation of relevancy scores for concepts we propose is based on the notion of information content. We adapt the approach to use interlinked taxonomies instead of an annotated textual corpus to calculate informativeness and use it as a relevancy score. The basis is the use of two taxonomies to describe the collection items. These taxonomies must have different semantic expressivity regarding the knowledge domain they describe to provide an advantage to non-experts. Higher expressivity means a more concrete description requiring a higher level of domain knowledge to understand. Lower expressivity means a more common description that also non-experts understand. We use the taxonomy with higher expressivity as a basis to calculate scores for the concepts of the taxonomy with lower expressivity. That means we interpret the higher taxonomy as the annotated corpus according to information content and calculate the scores based on the concept occurrences. These occurrences are represented by the links between concepts of the two taxonomies. The concepts of the lower taxonomy are used to describe the concepts of the higher taxonomy. They therefore provide a description of these concepts that is understood by non-experts. The first step of the method is to provide these taxonomy interlinks. The second is the calculation of the information content scores which represent the relevancy.

2.1 Vocabulary Interlinking

Interlinking multiple vocabularies that are used to annotate items has several advantages. It automatically extends descriptions by adding additional linked concepts and it improves search by improving recall. Using multiple vocabularies, we can describe items on different vertical semantic levels (Hyvönen et al., 2007). Users with different views or different knowledge can use different levels for describing the same item. The levels are represented by vocabularies with different semantic expressivity or in general a different point of view on the domain. Interlinking vocabularies also provides different approaches to data integration. (Hollink et al., 2003) shows how to improve semantic search by interlinking concepts of different ontologies that are used to annotate artworks. This increases the recall of the search. (Hyvönen et al., 2007) uses three vertical levels of detail with increasing semantic expressivity for different kinds of content annotation. It can be used to integrate content annotated at different levels of granularity. Each new level of annotation granularity adds new information

with respect to the previous level. This means that semantically richer representations can be easily interpreted at the lower level. These examples show the advantages provided by vocabulary interlinking.

2.2 Relevancy Scoring

The interlinks between concepts of multiple vocabularies represent descriptions in vertical semantic directions. To explain concepts of a higher semantic level, and therefore items of a collection annotated with these concepts, we provide relevancy scores for the lower level concepts based on the notion of information content. This can only be done for vocabularies with a taxonomic structure because the hierarchy is needed to calculate informativeness. The scores represent the importance a concept has for the annotated item in the scope of the interlinked vocabularies. Relevancy scoring can be done automatically and independent of any collection that the vocabularies are applied to. The following example explains the method. Let taxonomy A and taxonomy B be taxonomies used to annotate items. Taxonomy A is of higher semantic expressivity than taxonomy B. By using the interlinks we calculate the relevancy of concepts of taxonomy B based on their occurrences in taxonomy A. If a concept from taxonomy B is not interlinked to concepts of taxonomy A, its information content is 0. Concepts that are not interlinked do not contribute to the information content score. The relevancy for a concept of taxonomy B is the information content score based on the hierarchy of taxonomy B and the number of interlinks with concepts of taxonomy A. This means we interpret taxonomy A as a corpus for taxonomy B and the frequency of interlinks is taken into account to calculate the probability to encounter a concept. In figure 1, we see several interlinks going from concepts of taxonomy B to concepts of taxonomy A. Concept b2 has two links (a3, a4), b3 has one link (a3) and b4 also has one link (a2). The total number of interlinks is 4. We calculate the information content scores (ic) for concepts of taxonomy B. The information content score for a concept of taxonomy B is the negative log of the likelihood to encounter this concept, including the child concepts.

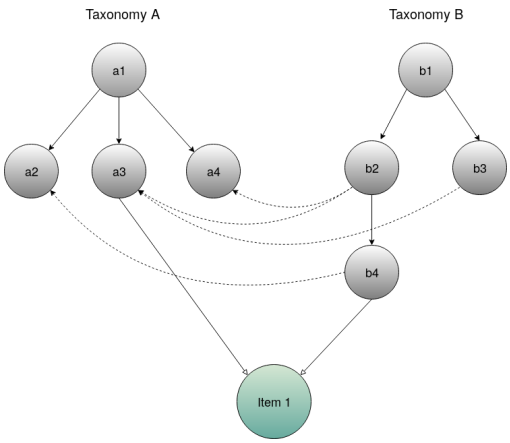

Figure 1: Taxonomy interlinking.

$$ic(b1) = -\log(\frac{2}{4} + \frac{1}{4} + \frac{1}{4}) = 0$$

$$ic(b2) = -\log(\frac{2}{4} + \frac{1}{4}) = 0.125$$

$$ic(b3) = -\log\frac{1}{4} = 0.6$$

$$ic(b4) = -\log\frac{1}{4} = 0.6$$

The information content scores represent the relevancy that each concept of taxonomy B has for describing concepts of taxonomy A. The scoring is done automatically and is not dependent on the item collection the two taxonomies describe.

3 ART HISTORIC IMPLEMENTATION

3.1 Vocabularies for Art Description

In this paper we focus on semantic artwork descriptions to be used by a knowledge-based recommender system for understanding and exploring an artwork collection. We select two vocabularies for doing the descriptions. Both are organized as hierarchic taxonomies. They are available as Semantic Web (Berners-Lee et al., 2001) data models using the Simple Knowledge Organization System (SKOS) (Alistair Miles, 2009). Both vocabularies were designed to describe artworks, so it makes sense to apply them both to artwork items for a semantic description. They represent information at different vertical levels of semantic expressivity, so they qualify for interlinking and relevancy score calculation.

Iconclass. Iconclass (Van de Waal, 1974) is a classification system for iconography. It is structured as a hierarchy of concepts representing themes of the iconographic tradition of art history. Therefore it is mainly intended for historical artworks of European origin. The concepts within Iconclass include common artwork themes throughout history and mythology. It is not intended to describe single parts of an artwork, but contains concepts that can be used to describe the scene as a whole. Historic and

mythological persons and stories are included. Even single scenes from stories are represented by concepts in Iconclass. For example, the concept `http://iconclass.org/71U4273`, titled "Holofernes beheaded by Judith with his own sword; the maidservant may be keeping watch", represents the scene of the Old Testament where Judith beheads Holofernes. This example shows the high level of detail the concepts of Iconclass have. With just one concept the whole content of an artwork can be described. Iconclass is intended to be used by domain experts to identify and annotate iconographic scenes for historical artworks.

AAT. The Getty Art and Architecture Thesaurus (Petersen, 1990) includes a broad range of concepts that can be used to describe art and architectural works. It has many concepts for simple items and physical objects, concepts for materials that can be used to describe the physical characteristics of works, and also concepts for immaterial things like emotions and seasons. AAT concepts can be used to describe the elements shown by an artwork. Therefore AAT can be used for the same purpose as Iconclass, although on a different semantic level because it does not contain iconographic concepts. Because both vocabularies have the same purpose, but different views on artworks, they are suitable for the use case evaluation.

3.2 Relevancy Scoring for Art History

Iconclass, as an expert taxonomy of the art historical domain, is difficult to understand for non-experts. In contrast, the AAT represents a more general taxonomy and is therefore easier to understand for non-experts. To provide a better understanding for the Iconlass concepts, we will relate them to associated AAT concepts. This is done by first interlinking the AAT concepts to Iconclass concepts and then calculating the information content scores for the AAT concepts based on the interlinks. These scores show the relevancy of the AAT concepts for iconographic descriptions in the context of the two vocabularies.

Interlinking Iconclass and AAT. For interlinking Iconclass and AAT concepts we have to find a way to describe the Iconclass concepts using the AAT concepts. The hierarchic structure cannot be used for this because the two taxonomies, although intended for a similar purpose, represent different knowledge domains. An alignment based on the structure is therefore not possible. However, both taxonomies provide labels as textual descriptions of the concepts that can be used as a basis for matching. For Iconclass concepts, besides the labels, there is a context description for each concept that is done using a controlled vocabulary. Although this vocabulary does not have a structure, there are no ambiguous terms in it. It also represents a more general description of a concept which is something that is similar to what AAT provides. For the AAT, we also use the provided labels as a basis for interlinking. We match the labels of the AAT concepts to the subject terms of the Iconclass concepts to create an initial mapping. This is done by using a regex-based approach that removes terms in brackets and then directly matches the results. Although both labels and subjects are not ambiguous within each vocabulary, the mapping result contains many ambiguities. To disambiguate, we take the taxonomic hierarchy of AAT and try to resolve the match based on the labels of the path to the taxonomic root. If the subjects of an Iconclass concept are found among the labels of the AAT concepts on the path to the root, we identify this AAT concept as a correct match. After resolving the ambiguities using this method, we have an automatically generated interlink set for Iconclass and AAT.

Relevancy Scoring. The relevancy scoring is done for AAT concepts based on the result of the interlinking. We compute the information content values as described above. The result are the relevancy scores for all interlinked AAT concepts that represent how significant a concept is for the description of art historical items.

4 USE CASE EVALUATION

For evaluating the approach we apply it to an art historical collection of items. Evaluation is done regarding the correctness of the interlinking and scoring, as well as the art historical correctness of the results. The higher scored concepts are more significant than the lower scored concepts regarding an item description. That means that the order of the concept annotations of an item, created by the relevancy scores, has to be verified for correctness from an art historical point of view. The evaluation is presented in three parts. First, we describe the collections we used as a basis for the evaluation. Second, we show three example items from the collections to illustrate the approach. We discuss the correctness of the interlinking and the concept scores of these items. We then show how the concepts relate the different items with each other and how the relevancy scores represent the significance of these relations. Third, we describe how we evaluated

the approach on a larger scale by creating sample sets and performing an art historical analysis of the scored concept annotations.

4.1 Artwork Collection

For the collection of artworks used in the evaluation we decided to use the Linked Data representation (Dijkshoorn et al., 2014) of the Rijksmuseum collections, which is available as Linked Open Data. This dataset contains over half a million objects, including detailed descriptions and high-quality images released under a public domain license. The museum uses the Iconclass vocabulary to describe subject matter. We chose two collections from the Rijksmuseum dataset which contain visual artworks. The first collection contains paintings, while the second collection contains pottery. We evaluate the approach by first concluding AAT concepts based on the asserted Iconclass concepts for an item and then verifying the interlinks and the relevancy scores regarding the iconographic interpretation.

4.2 3 Examples

We present 3 examples of artworks that show different scenes. The Iconclass annotations were added by the Rijksmuseum and are asserted in the published dataset. Most of the annotations of the three artworks are located in different parts of the Iconclass taxonomy. There is not much relation between them regarding the iconographic description. However, by expanding to AAT concepts based on interlinking, we can conclude shared concepts among them. We show how the artworks relate and how significant the relation between them is based on the relevancy scores.

Example 1: Time (`http://hdl.handle.net/10934/RM0001.COLLECT.7192`) The first artwork shows time as an old man, naked and with wings on his back. He is kneeling on his right knee and holds up a cloth with his right hand. On the floor lies a golden scale between leaves. Time, being an abstract concept, is represented by the physical form of a person. Such symbolic representations are called allegories. The item is annotated with one Iconclass concept which describes the scene, as shown in table 1.

Table 1: Time.

Iconclass Concept	Title
http://iconclass.org/23A1	Father Time, man with wings and scythe

Example 2: Autumn (Conversation) (`http://hdl.handle.net/10934/RM0001.COLLECT.6396`) The second artwork shows a landscape in the season of autumn. We can see a road on the banks of a water with some fishing boats. On the road there is a man who speaks with a woman. The artwork is part of a series about the four seasons. The item is annotated with three Iconclass concepts which describe the scene, as shown in table 2.

Table 2: Autumn (Conversation).

Iconclass Concept	Title
http://iconclass.org/46C22	harbouring
http://iconclass.org/23F44	autumn landscape; landscape symbolizing autumn (the four seasons of the year)
http://iconclass.org/33A35	conversation, dialogue; conversation piece

Example 3: Still Life with Fruit and Oysters (`http://hdl.handle.net/10934/RM0001.COLLECT.9119`) The third artwork shows a still life with fruits, oysters and a watch. The item is annotated with four Iconclass concepts which describe the scene, as shown in table 3.

Table 3: Still Life with Fruit and Oysters.

Iconclass Concept	Title
http://iconclass.org/23U25	watch
http://iconclass.org/25F72(OYSTER)	molluscs: oyster
http://iconclass.org/41C653	fruit
http://iconclass.org/41C38	laid table as still life

Interlinks and Relevancy Scores. Using the interlinks we conclude the following AAT concepts with relevancy scores for the items shown in table 4, table 5 and table 6. These concepts do not only represent what scenes are shown on the artworks, but also what the meaning of these scenes is. They represent iconographic interpretations using general concepts that non-experts understand more easily compared to the iconographic concepts of Iconclass. When we look at the concluded AAT concepts, we can see some of them are incorrect for the description of the artwork. These false positives were produced because the disambiguation did not have sufficient context information in these cases. We show them in the tables marked in red. They are incorrect and therefore will not be considered any further. Then we look at the relevancy scores and the order of concepts that

Table 4: Time.

AAT Concept	Title	Score
http://vocab.getty.edu/aat/300024371	scythes	10.3707
http://vocab.getty.edu/aat/300375053	wings (animal components)	8.9653
http://vocab.getty.edu/aat/300133089	time	6.8398
http://vocab.getty.edu/aat/300202507	allegories (document genre)	5.3514
http://vocab.getty.edu/aat/300055866	allegory (artistic device)	5.3514
http://vocab.getty.edu/aat/300179372	Nature	5.1355

Table 5: Autumn (Conversation).

AAT Concept	Title	Score
http://vocab.getty.edu/aat/300178227	conversation pieces (portraits)	12.2425
http://vocab.getty.edu/aat/300026185	dialogues	10.8562
http://vocab.getty.edu/aat/300133093	autumn	10.7384
http://vocab.getty.edu/aat/300008678	harbors	10.1631
http://vocab.getty.edu/aat/300018959	Company (style)	9.7576
http://vocab.getty.edu/aat/300160084	companies	9.2221
http://vocab.getty.edu/aat/300133091	seasons	7.2487
http://vocab.getty.edu/aat/300055250	traffic	6.8398
http://vocab.getty.edu/aat/300133089	time	6.8398
http://vocab.getty.edu/aat/300082981	ships	6.6367
http://vocab.getty.edu/aat/300053892	transporting	6.2485
http://vocab.getty.edu/aat/300179372	Nature	5.1355
http://vocab.getty.edu/aat/300008626	landscapes (environments)	4.7171
http://vocab.getty.edu/aat/300265711	Homo sapiens (species)	4.3210
http://vocab.getty.edu/aat/300055806	civilization	4.1848
http://vocab.getty.edu/aat/300026009	societies	4.1838
http://vocab.getty.edu/aat/300055768	culture	4.1259

Table 6: Still Life with Fruit and Oysters.

AAT Concept	Title	Score
http://vocab.getty.edu/aat/300041615	watches	12.9356
http://vocab.getty.edu/aat/300310174	oysters	11.5493
http://vocab.getty.edu/aat/300380162	vegetables	9.9399
http://vocab.getty.edu/aat/300265736	tables (architectural elements)	9.8446
http://vocab.getty.edu/aat/300027364	tables (documents)	9.8446
http://vocab.getty.edu/aat/300216559	bellies	9.6034
http://vocab.getty.edu/aat/300039548	tables (support furniture)	9.1290
http://vocab.getty.edu/aat/300053578	measuring	8.4813
http://vocab.getty.edu/aat/300266510	Mollusca (phylum)	8.1650
http://vocab.getty.edu/aat/300024838	instruments	8.0681
http://vocab.getty.edu/aat/300015638	still lifes	7.3983
http://vocab.getty.edu/aat/300254496	food	7.0114
http://vocab.getty.edu/aat/300011868	fruit	6.8789
http://vocab.getty.edu/aat/300133089	time	6.8398
http://vocab.getty.edu/aat/300011734	earth (soil)	5.2898
http://vocab.getty.edu/aat/300311363	earth (color)	5.2898
http://vocab.getty.edu/aat/300179372	Nature	5.1355
http://vocab.getty.edu/aat/300055806	civilization	4.1848
http://vocab.getty.edu/aat/300026009	societies	4.1838
http://vocab.getty.edu/aat/300055768	culture	4.1259
http://vocab.getty.edu/aat/300249395	Animalia (kingdom)	3.0796

they create for the artwork description. We can see that the order is valid regarding the depicted scenes and the iconographic interpretations. It goes from the more specific concepts to the more generic ones and provides additional detail information to non-experts compared to the Iconclass concepts. We also see that there are concepts added that have thematically to do with the artwork, but are not actually shown. An example is the concept for scythes for the artwork Time (table 1), which is the top scoring concept (table 4). This is an often used symbol of the allegorical representation of time as an old man, although it is not shown on this specific artwork example. Still we would want to use this concept when retrieving related artworks.

Relations between the Artworks. We showed three examples of artworks with their interlinked and scored AAT concepts. We can see the significant concepts as a description of what the artwork scene represents. Now we can compare the three artworks to see how they are related. This comparison can be used by a knowledge-based recommender system to retrieve navigation options. The three artworks show different scenes and have different iconographic interpretations. Still they have some concepts in common. These concepts, time and nature, are shown in table 7. The first artwork, Time, is an anthropomorphic representation of time. This is an art historical interpretation that requires domain knowledge. The second artwork, Autumn (Conversation), shows a scene in autumn. Although not clearly visible from the content, the title of the artwork indicates this to the viewer. Autumn, as a season, is one of the four time periods in a year. The third artwork, Still Life with Fruit and Oysters, is a still life, which is a genre of artworks showing mostly inanimate subject matter. Still lives often have symbolic elements. This artwork depicts a watch as part of the scene, which is a symbolic representation of the brevity of life. We can see that the three artworks have the concept of time in common. Also, they have the concept of nature in common, although this is less significant. Table 7 shows the scores for both concepts. We can see

Table 7: Shared concepts.

AAT Concept ID	Title	Score
http://vocab.getty.edu/aat/300133089	time	6.8398
http://vocab.getty.edu/aat/300179372	Nature	5.1355

that the concept of time is in the upper score range of the scored concepts, while nature has a lower score. The relevancy scores of these concepts represent how strongly the three artworks are related. There is a relation, although not very obvious. This is also reflected in the scores. It is difficult for a non-expert to understand that the concept of time is part of the meaning of the three scenes. The first artwork has time in the title, while the second artwork mentions autumn, and the third artwork depicts a watch. By providing the interlinked and scored AAT concepts, we can show these shared concepts and the significance regarding the three artwork scenes. The three examples illustrate how the scored concepts represent a detailed description of artwork content. The concept order based on the scoring shows how significant each concept is for the description of the artwork content. By comparing the concepts and the relevancy scores, we can show relations between artworks as well as how significant these relations are.

4.3 Art Historical Evaluation

The evaluation was done on samples from the selected two collections of the Rijksmuseum dataset. We generated 3 sets of item samples from these collections containing 100 items each. As part of this research a prototypical web application was developed to calculate and retrieve scored AAT concepts for the items.

The resulting item information for the sample sets was published as a report. This report was then evaluated by experts of the art historical domain. The evaluation had to verify the correctness of the interlinks as well as the correctness of the order of the concept annotations. As already seen in the presented examples, the interlinks contain incorrectly disambiguated concepts. This not only created incorrect concept annotations, but also influenced the relevancy scores. We ignored the incorrect concepts from the artwork descriptions and evaluated the remaining correct concept annotations regarding the correctness of the order. It showed that the correct concept annotations provide a valid description of the details of the content regarding the relevance. The concluded concepts automatically extend the asserted descriptions and help non-experts to understand the details using the relevancy scoring. They can be used as a basis for retrieval in knowledge-based recommender systems.

5 CONCLUSION

In this paper we presented an approach for relevancy scoring based on the interlinking of vocabularies. It is not dependent on instance data and does not have drawbacks regarding the size of the collection. It will work for small collections well as for large ones. However, it is dependent on the vocabularies selected for interlinking and the quality of the interlinks. To help users to better understand the meaning of artworks, the two vocabularies must have different semantic expressivity, so the concluded information is represented using a view that is understandable to non-experts. The data model can be generated automatically by vocabulary interlinking and calculating information content scores for the interlinked concepts. We performed a interdisciplinary evaluation to verify the approach. It showed that the interlinking resulted in false positive concept annotations that were incorrectly disambiguated. We ignored these false positives for the further evaluation and only considered the correct ones for relevancy scoring. It showed that the order of the correct concept annotations for the sample artworks was valid. However, the incorrect interlinks influence the relevancy scores and might result in a changed order for some artwork descriptions regarding correct concepts. This, as well as improving the disambiguation, has to be evaluated in future work. The next step is the extension of the implemented prototype for recommendations where the options are given by the scored concept annotations. We can then evaluate how these artwork descriptions help visitors to navigate the museum collections based on their interests by using the scored concepts for retrieval and as tweaking options.

REFERENCES

Alistair Miles, S. B. (2009). Skos simple knowledge organization system reference. http://www.w3.org/TR/skos-reference/. [Online; accessed 31-March-2015].

Berners-Lee, T., Hendler, J., Lassila, O., et al. (2001). The semantic web. *Scientific american*, 284(5):28–37.

Burke, R. (2000). Knowledge-based recommender systems. In *Encyclopedia of Library and Information Systems*, page 2000. Marcel Dekker.

Burke, R. D., Hammond, K. J., and Young, B. C. (1997). The findme approach to assisted browsing. *IEEE Expert: Intelligent Systems and Their Applications*, 12:32–40.

Dijkshoorn, C., Jongma, L., Aroyo, L., van Ossenbruggen, J., Schreiber, G., ter Weele, W., and Wielemaker, J. (2014). The rijksmuseum collection as linked data. *Semantic Web*, (Preprint):1–10.

Felfernig, A. and Burke, R. (2008). Constraint-based recommender systems: Technologies and research issues. In *Proceedings of the 10th International Conference on Electronic Commerce*, ICEC '08, pages 3:1–3:10, New York, NY, USA. ACM.

Gicquel, P.-Y. and Lenne, D. (2013). Semantic and Contextual Proximities for Informal Learning: The Case Study of Museum Visits. *International Journal of Computer Applications*, 69(26):12134–8410.

Hollink, L., Schreiber, G., Wielemaker, J., and Wielinga, B. (2003). Semantic annotation of image collections. In *Second International Conference for Knowledge Capture (KCAP2003) Workshop on Knowledge Markup and Semantic Annotation*.

Hyvönen, E., Ruotsalo, T., Haggström, T., Salminen, M., Junnila, M., Virkkilä, M., Haaramo, M., Mäkelä, E., Kauppinen, T., and Viljanen, K. (2007). *CultureSampo–Finnish Culture on the Semantic Web: The Vision and First Results*, pages 33–58. LIT Verlag, Berlin.

Petersen, T. (1990). *Art & architecture thesaurus*. Oxford University Press. http://www.getty.edu/research/tools/vocabulary/aat/.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 1*, IJCAI'95, pages 448–453, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Van de Waal, H. (1974). *Iconclass: an iconographic classification system*, volume 2. North-Holland.