

Demonstration of multi-perspectives exploratory search with the Discovery Hub web application

Nicolas Marie^{1,2} and Fabien Gandon¹

¹ WIMMICS, INRIA Sophia-Antipolis, Sophia Antipolis, France

² Alcatel-Lucent Bell Labs, Nozay, France
{firstname.lastname}@inria.fr

Abstract. This paper describes a demonstration of the Discovery Hub exploratory search system. The demonstration focuses on the exploration of topics through several perspectives.

1 Introduction

Exploratory search refers to cognitive-consuming search tasks like learning or investigation. There is a need to develop systems optimized for supporting exploratory search as the today widely-used search engines fail to efficiently support exploratory search [1]. Linked data offers exciting perspectives in this context and several systems were already published. We want to reach a new level in linked data based exploration by allowing the users to unveil *knowledge nuances* corresponding to specific facets of interest of the topic explored. In this demonstration paper we first give a brief overview of linked data based exploratory search systems. Then we present the Discovery Hub web application and focus more particularly on its multi-perspectives exploration capacity. Finally we present the demonstration scenario we propose.

2 Linked data based exploratory search

The amount of contributions at the crossroad of semantic search and exploratory search is increasing today. The linked data knowledge, and especially DBpedia, allows to design new information retrieval approaches and interaction models that efficiently support exploratory search tasks. **Yovisto**³ (2009) is an academic videos platform that retrieves topics' suggestions that are semantically related to the users' query. The objective is to ease the exploratio of the videos collection. **Lookup Explore Discover**⁴ (2010) helps the users to compose queries about topics of interest by suggesting related query-terms. Once the query is composed the system retrieves the results from others several services such as search engines and social networks. **Aemoo**⁵ (2012) offers a graph-view on topics of interest. The graph shows their neighborhood filtered by a semantic pattern.

³ <http://www.yovisto.com/>

⁴ <http://sisinflab.poliba.it/led/>

⁵ <http://wit.istc.cnr.it/aemoo>

The users can reverse the filtering to show more surprising knowledge. They can also ask for explanations (cross-references in Wikipedia) about the relations between the shown resources. The **Seevl**⁶ (2013) demonstrator is a music discovery platform implementing a linked data based recommendation algorithm. The DBpedia semantics are also used in Seevl to support browsing (e.g. by music genres, band members) and to provide explanations about the recommendations (showing the shared properties between the artists). **Linked Jazz**⁷ (2013) aims to capture the relations within the American jazz community in RDF. The authors rely on a large amount of jazz people interviews transcripts. These transcripts are automatically processed and then finely analyzed through a crowd-sourced approach.

The approaches recently published in the literature produced good results when evaluated. Nevertheless a common limit of the existing linked data based exploratory search systems is the fact they constrain the exploration through single results selection and ranking schemes. The users cannot influence the retrieved results to reveal specific aspects of knowledge that interest them in particular.

3 Multi-perspectives exploratory search

The framework and models implemented by the Discovery Hub application was presented in [2]. Contrary to other systems it does not pre-compute and store the results for later retrieving. Instead it computes the results on demand thanks to a semantics-sensitive graph traversal algorithm. The algorithm is applied on a small amount of data stored in a local and transient triple store. The data is incrementally imported at query time using SPARQL queries sent to the targeted SPARQL endpoint (DBpedia in the case of Discovery Hub). The objective of this step is to identify a set of relevant results related to the initial topics of interest that will be explored by the user [2]. The web application demonstrating the framework, called Discovery Hub, is available online⁸ and was showcased in several screencasts⁹.

The fact the results are computed at query-time allows to let the users control several computation parameters through the interface and to offer multi-perspective exploratory search. Indeed, the objects described in linked data datasets can be rich, complex and approached in many manners. For example, a user can be interested in a painter (e.g. *Claude Monet* or *Mary Cassat*) in many ways: works, epoch, movement, entourage, social or political contexts and more. The user may also be interested by basic information or unexpected ones depending on his actual knowledge about the topic. He may also want to explore the topic through a specific culture or area e.g. impressionism in American or French culture.

⁶ <http://play.seevl.fm>

⁷ <http://linkedjazz.org/>

⁸ <http://discoveryhub.co>, current CPU-intensive experiments might slow down the search temporarily

⁹ <https://www.youtube.com/user/wearediscoveryhub/videos>

The framework allows three operations for building such exploration perspectives, detailed in [3]. (1) The users can **specify criteria of interest** and disinterest that are used by the framework during the sample importation and its computation. The DBpedia categories are used for this purpose, see Figure 1. The objective is to guide the algorithm in order to retrieve results that are more specifically related to the aspects that interest the user, see example of queries and results in Table 1. (2) It is possible to **inject randomness** in the algorithm values in order to modify the ranking scheme and expose more unexpected results¹⁰. (3) With the proposed framework it is easy to **change the data source** used to process the query¹¹. In the context of DBpedia it enables the use of the DBpedia international chapters like the French, German, Italian ones¹² to leverage cultural bias.

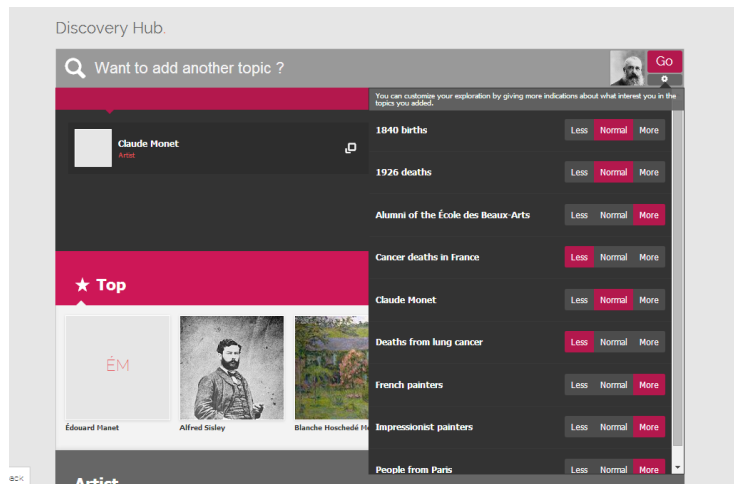


Fig. 1. Discovery Hub interface, criteria of interest specification query

4 Demonstration scenario

The demonstration will be constituted of a sequence of interactions like the ones presented in the previously published screencasts. First the user launches a simple query (query 1 in Table 1), he examines the results. At this point we show the audience the Discovery Hub functionalities supporting the exploration and understanding. We will focus on the faceted browsing aspect, the explanations and the redirections toward third-party platforms. During this step we will engage the conversation about how we compute the results and what is the software architecture.

¹⁰ this advanced query mode is supported by the framework but currently not available through the interface and will be integrated soon

¹¹ idem 10

¹² wiki.dbpedia.org/Internationalization/Chapters?v=190k

Table 1. Results of three queries about Claude Monet using the criteria specification

Query	Claude Monet (1)	Claude Monet (2)	Claude Monet (3)
Criteria	None	Impressionist painters + Artists from Paris - People from Le Havre - Alumni of the École des Beaux-Arts - French painters -	Impressionist painters - Artists from Paris + People from Le Havre + Alumni of the École des Beaux-Arts + French painters +
Results			
1	Pierre-Auguste Renoir	Theodore Robinson	Pierre-Auguste Renoir
2	Alfred Sisley	Edouard Manet	Gustave Courbet
3	Édouard Manet	Alfred Sisley	Edgar Degas
4	Mary Cassatt	Władysław Podkowiński	Jacques-Louis David
5	Camille Pissarro	Leslie Hunter	Jean-Baptiste-Camille Corot
6	Edgar Degas	Theodore Earl Butler	Jean-François Millet
7	Charles Angrand	Lilla Cabot Perry	Paul Cézanne
8	Gustave Courbet	Frank Weston Benson	Marc Chagall

During the results examination we will voluntarily focus on the French impressionist painters that were close to Monet. At this point the user might be interested in the relations of Monet with the non-French impressionists (query 2 in Table 1). We will explain the querying system for criteria of interest specification and then emphasize the differences between the results obtained with query 1 and 2.

To continue in the same logic we will submit the query 3 as well as a query with a high level of randomness and one using the French chapter of DBpedia in several tabs. We let the audience compare the results. We seek an interactive demonstration by encouraging the audience to try the application while commenting the system more than a strict and pre-defined sequence of interactions (which serves only to start the interactions).

5 Conclusion and perspectives

Discovery Hub is a linked data based exploratory search system built on the top of DBpedia. With this demonstration we want to show the value of linked data for exploratory search. Mature datasets like DBpedia allow the creation of new information retrieval approaches as well as new interaction models. More specifically we want to demonstrate the multi-perspectives exploratory search capacities of Discovery Hub. Thanks to the demonstration track we hope to have discussions with other researchers about the perspectives we envision for Discovery Hub. It notably includes an approach where the user can specify or change the specified criteria of interest interactively in order to re-rank the results without relaunching the whole query-process.

References

1. G. Marchionini. Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4):41–46, 2006.
2. N. Marie, F. Gandon, M. Ribière, and F. Rodio. Discovery hub: on-the-fly linked data exploratory search. In *Proceedings of the 9th International Conference on Semantic Systems*. ACM, 2013.
3. N. Marie, M. Gandon, Alain Giboin, and E. Palagi. Exploratory search on topics through different perspectives with dbpedia. In *Proceedings of the 10th International Conference on Semantic Systems*. ACM, 2014.