A SURVEY OF EXPLORATORY SEARCH SYSTEMS BASED ON LOD RESOURCES

Karwan Jacksi[1], Nazife Dimililer[2] and Subhi R. M. Zeebaree[3]

[1]*University of Zakho, (UOZ), Iraq,* karwan.jacksi@uoz.ac
[2]*Eastern Mediterranean University, (EMU), N. Cyprus* nazife.dimililer@emu.edu.tr
[3]*Duhok Polytechnic University, (DPU), Iraq,* subhi.rafeeq@dpu.edu.krd

ABSTRACT. The fact that the existing Web allows people to effortlessly share data over the Internet has resulted in the accumulation of vast amounts of information available on the Web. Therefore, a powerful search technology that will allow retrieval of relevant information is one of the main requirements for the success of the Web which is complicated further due to use of many different formats for storing information. Semantic Web technology plays a major role in resolving this problem by permitting the search engines to retrieve meaningful information. Exploratory search system, a special information seeking and exploration approach, supports users who are unfamiliar with a topic or whose search goals are vague and unfocused to learn and investigate a topic through a set of activities. In order to achieve exploratory search goals Linked Open Data (LOD) can be used to help search systems in retrieving related data, so the investigation task runs smoothly. This paper provides an overview of the Semantic Web Technology, Linked Data and search strategies, followed by a survey of the state of the art Exploratory Search Systems based on LOD. Finally the systems are compared in various aspects such as algorithms, result rankings and explanations.

Keywords: semantic web, linked data, exploratory search system

INTRODUCTION

As the World Wide Web (WWW) progressively made the process of uploading data easy for all users, the amount of information grew in an unprecedented manner and has transformed the Web to a huge semi-structured database. The new challenge to overcome is to retrieve relevant information from different resources efficiently. Even though search engines are generally employed to locate information on the web, to obtain high quality results, efficient searching skills are required.

Marchionini in (Marchionini, 2006) has distinguished between two categories of search strategies: lookup and exploratory search. Lookup search strategy, which is also known as keyword-based search, has as background database systems that are used to find information containing a specific keyword. This search strategy is considerably the most prevalent one in the existing Web, also known as Syntactic Web, where the sources of data are largely textual documents (Mirizzi et al., 2010). On the other hand, in exploratory search strategy, the final goal of end user may change. In this strategy, the interest of the end user focuses on learning activities such as understanding new ideas and concepts, knowledge achievement and information assessment, and investigation such as analysis, composition and evaluation more willingly than query answering and information retrievals (Jiang, 2014).

The Syntactic retrieval of information is restricted to keywords. Thus, to answer user queries, search engines search for keywords which do not retrieve the high quality results. To improve this, annotations can be added to the existing Web contents. Hence, the Semantic Web is formulated (Madhu et al., 2011; Jenice Aroma & Kurian, 2013).

The Semantic Web is the next generation and an augmentation of the WWW, where information has well-defined meaning and, is machine processable and understandable, and hence allows people and machines to work in cooperation (Berners-Lee, Hendler, & Lassila, 2001). The Semantic Web combines Resource Description Framework (RDF), Web Ontology Language (OWL), and Extensible Markup Language (XML) technologies to facilitate replacing the content of the current Web. These technologies provide machines a way to read and understand the meaning of data. Consequently, automated information processing capability is given to search systems (Jenice Aroma & Kurian, 2013).Berners-Lee in (Berners-Lee, 2006) remarks that the Semantic Web is not only about annotating data on the Web, but it is also about linking data with each other so that people and machines can discover the Web of data. Linked data provides a way to find other related data when only a subset is given. As a result, Berners-Lee has coined both terms, Semantic Web and Linked Data, and defines the Linked Data as “the Semantic Web done right”.

The Semantic Web provides data well defined meanings allowing automated data processing and thus raises the efficiency and success of the searching methods (Berners-Lee et al., 2001). The data on the Semantic Web is described using W3C standard named RDF which is purely an XML language. It is a standard model for interchanging of data on the Web and it simply describes resources on the Web as triplets in the form of <subject, predicate, object> expressions. The linking structure of the Web is extended by RDF to use URIs to denote the relationship between resources. Hence, using this simple technique, a combination of structured and semi-structured data on the Web is formed creating a directed, labelled graph, where the edges of the graph represent the labelled link between two resources. One of the vital components of the Semantic Web infrastructure is ontology (Fensel, 2005). OWL and RDF Schema (RDFS) are the data representation models recommended by W3C to provide elements for the description of ontologies (Hitzler, Krotzsch, & Rudolph, 2011). The Linked Data is a term that refers to a set of steps for publishing and linking structured data on the Web. Berners-Lee has introduced these steps in his design issues’ notes on Web architecture. These steps have turned out to be the Linked Data Principles (Heath & Bizer, 2011). These steps state: 1)Identify things with universal URIs; 2)Use HTTP URIs to allow people and agents to look these things up and dereference them, 3)Information on the Web should be served alongside the URIs on top of using standards such as RDF and SPARQL[1], 4) Make links to other URIs so more things can be discovered (Le Hors, Nally, & Speicher, 2012).

Hypertext Web connects HTML documents with each other using untyped hyperlinks, while Linked Data depends on documents having RDF data formats (Klyne, Carroll, & McBride, 2004). Thus, instead of simply linking these documents with each other, Linked Data uses RDF to create typed links that connect arbitrary things globally. The main requirements for Linked Data are explicit semantic definition and links to/from other external data sets (Heath & Bizer, 2011). Linked Open Data (LOD) refers to Linked Data is published un-

[1] www.w3.org/TR/rdf-sparql-query

der an open license. However, in general, the Linked Data does not have to be open, since it can be used internally, for personal and groups (Berners-Lee, 2006).

EXPLORATORY SEARCH SYSTEMS

Exploratory search is a new approach to seek information on the Web, where the aim is not to find merely what have been sought for, but to find information related to it as well. This approach represents the searcher activities when the final goal is not predetermined, the domains of the goals are unknown and the goal itself is not known. So, a larger class of activities, such as learning, investigation, comparison and evaluation, are covered by exploratory search compared to classical information retrieval. The Exploratory Search Interface workshop in 2005 was the first significant event in the history for this approach (White, Kules, & Bederson, 2005). The Semantic Web technology with its nature of robotic information processing offers a great potential for search systems; especially for exploratory search subdiscipline (Dimitrova, Lau, Thakker, Yang-Turner & Despotakis, 2013). In this section, the state of the art exploratory search systems which use the Semantic Web technologies are reviewed.

Yovisto

Yovisto is a search engine dedicated to academic lecture and conference videos. This system utilizes Semantic Web and LOD for implementing explorative search for video data (Waitelonis & Sack, 2012).

Searching within the content of videos is possible with Yovisto through its time based video index, which makes it distinct among other video search engines. Metadata generation uses robotic analysis systems such as intelligent character recognition methods and scene detection techniques. Furthermore, at any point within a video, users can annotate their comments and tags through time dependent collaborative annotation. Hence, the index of the system is developed from fine granular time dependent metadata.

Yovisto uses LOD resources for the search process provision; it exposes further information or suggestions, which is semantically related to users query, from its indexed resources. The approach utilizes DBpedia1 as a resource, but since DBpedia provides a large amount of information about each of its entities, retrieving and presenting all relevant information is computationally intensive. Hence, a set of heuristics, based on statistical and structural features of the DBpedia's RDF graph, are developed to determine the most relevant information to the users query. These heuristics are used to rank the relations between DBpedia entities to conclude their relevance. Furthermore, since processing online queries with DBpedia leads to performance issues, the system uses an offline handling to process every term in advance.

Semantic Wonder Cloud (SWOC)

SWOC is a linked data based exploratory search system for DBpedia (Mirizzi et al., 2010). The system allows users to explore DBpedia nodes using semantic connections in the RDF dataset, in addition to exploiting DBpedia nodes via associations which are computed using knowledge from external sources such as Web search engines and social tagging systems. Therefore, a hybrid approach is used to rank pairs of DBpedia resources which makes the SWOC a unique system among other RDF explore systems. The backend of the system com-

[1] www.dbpedia.org

putes the associations between pairs of DBpedia nodes, and a flash based Web application that presents the information achieved from the backend.

The backend of the system, DBpediaRanker, is used to rank resources in the DBpedia, which is based on the hybrid ranking algorithm. This algorithm is used to calculate the similarity of DBpedia resources with respect to an initial node. So, while DBpediaRanker discovers the DBpedia graph, it queries external data sources in order to calculate a similarity value for each of the resource pairs touched throughout the discovery.

In the interface, the user selects a starting topic to be discovered; then an autocomplete drop down list is presented from the system containing a list of labels, returned by DBpedia lookup service[1], referring to DBpedia resources. . The selected resource is returned in a graph form, and it is surrounded by 10 most similar concepts. The size of each node in the generated graph increases proportionally with the computed similarity value for each resource. The bigger the size of the node is, the more relevant to the selected concept.

Lookup Explore Discover (LED)

LED is an exploratory search system that aims at improving the search process by allowing users to accurately explore information related to a query (Mirizzia, Di Noiaa, & Di Sciascioa). The system depends on DBpedia dataset to discover the semantics of the query keywords. When a user enters a keyword of interest, DBpedia lookup service is used to return a drop down list containing a set of DBpedia resources. Once the user selects a resource from the list the LED suggests a set of semantically related resources that are exposed in the form of a tags cloud.

The users of the system can refine their query by adding new tags available from the tag cloud. When the user adds a new tag to the query, a new tab is created containing relevant tags to all concepts in the refined query. The generated tags in the cloud are semantically related with each other and to the query, in addition each of them is associated to an RDF resource in the DBpedia. The system uses several external sources such as the major search engines: Google, Bing and Yahoo to retrieve information. In addition, results from a microblogging service, Twitter, and a news feed are also integrated into the result space. DBpediaRanker, a hybrid ranking algorithm, is used in this approach to rank resources in DBpedia regarding a given keyword. The system offers a RESTful JSON API that can be read and accessed from any method having HTTP requests. LED can also be used as a web service which is called Not Only Tag (NOT[2]); it can become a semantic tagging system with trivial interface modifications, without modifying the backend of the system. NOT is the essential of a semantic social tagging system for an active annotation and any Web resources type retrievals.

Aemoo

Aemoo is a Web application that uses Encyclopedic Knowledge Pattern (EKP)[3] to support its exploratory search system (Musetti et al., 2012). When users query Aemoo from its simple keyword based search interface, the system exploits DBpedia first to resolve users' query, then external sources such as Wikipedia, Google News and Twitter are used to collect and aggregate the knowledge which is achieved through the use of knowledge patterns, the structure of hypertext links, and exploitation of the Semantic Web techniques.

[1] wiki.dbpedia.org/Lookup
[2] sisinflab.poliba.it/not-only-tag
[3] www.ontologydesignpatterns.org/ekp

In order for the system to decide what information should be presented, the retrieved knowledge from the resources is filtered using the EKP so that Aemoo presents a set of relevant and reasonable information. In addition, the motivation of why a certain piece of knowledge is presented is included.

A supplementary function called curiosity is offered also by Aemoo. This function displays additional knowledge that is not included in the previous presentation due to the EKP filtration.

Seevl

Seevl is a Linked Data based application for musical discovery (Passant, 2012). The Web is used by Seevl to mine music connections so that context, search and discovery are brought for the music a user likes. Multiple Web resources are collected and unified as Linked Data. As a result, a huge graph of musical entities, such as Artists, Labels, Bands and etc., are built. On top of this collected data, further services, for instance search and recommendations, are provided.

A dedicated RDF store is used to store the collected data; this RDF store is powered by OpenLink Virtuoso. The store is hosted on an Elastic Compute Cloud (EC2) with the aim of obtaining the benefits of EC2 architecture such as load balancing and elastic cache. Later on, a dedicated data browser is developed on top of the data so that the available data can be explored by users and recommendations for each artist in the store are given. Explanations of these recommendations are shown to users so as to help them understand why a certain artist is recommended.Seevl is a fully Semantic Web based architecture, where all of its parts built using Semantic Web technologies. Data from various sources are modelled as RDF, stored in RDF store and rendered through SPARQL queries.

Discovery Hub

Discovery Hub is an exploratory search system that is based on the technology of Linked data (Marie, Gandon, Ribière & Rodio, 2013). To query the system, an input box is available so users can enter their topics of interest. When a topic is entered, the system displays a drop down list of concepts from the Web of Data that are related semantically to the input topic. Once the topic is selected, it pushes to a stack located to the right of the search box. This stack is used to let users enter their new topics of interests. Afterwards, the engine searches for the selected topics and categorizes them according to their relevant labels. Faceted browsing is offered by the search engine and several results explanations characteristics.

The platform uses DBpedia datasets for resource selection, and addresses DBpedia SPARQL endpoints remotely with the aim of performing the results selection and ranking on the fly without any pre-processing requirements. The system depends on an extension of semantic spreading activation to typed graphs of the Web of Data formality and combined it with a graph sampling method to calculate result without the need of in advance processing.

Linked Jazz

Linked Jazz is an exploratory system that reveals the relationships of the jazz community (Pattuelli, Miller, Lange, Fitzell, & Li-Madeo, n.d). The system aims at improving the discovery of digital cultural heritage materials and enriching the semantics describing them examining the power of LOD technology. Another objective of the system is to discover the associations between musicians and uncover their community's networks. Thus, the goal of the system is a support of finding meaningful relations concerning the professional lives and personal details of jazz musicians from the materials based on interviews' transcripts from jazz archives. As a consequence, an RDF dataset that describes these meaningful relations is built as LOD. A directory of individual names of jazz musicians joined up with their URIs, a fundamental mechanism for the LOD, is built which is called Linked Jazz Name Directory. Data from bibliographic name authority and DBpedia have been used for the creation of the Name Directory by using Mapping tool. The outcome of this process is a dataset which is in turn gets refined by using the Curation tool.

Linked Jazz Network Visualization tool is used to discover and explore the final accomplished dataset. This tool provides a number of modes, such as Fixed, Free, Similar and Dynamic modes, to view and explore an overview of the jazz community network. The Fixed mode is used to pin the individuals with the most connections to the outside which makes the presentation of nodes clearer; grouping individuals together based on their number of connections is done by the Free mode; the Similar mode groups individuals together by their number of shared connections; and finally, the Dynamic mode which lets users to manually add individuals and view their shared connections.

inWalk

inWalk is an interactive Web application for exploration of linked data based on the concept of inCloud (Castano, Ferrara, & Montanelli, 2014). An inCloud is a high level graph of thematically connected vertices, where vertices of the graph represent clusters of related linked data and edges denote relations of proximity between vertices. inCloud is constructed using hierarchical clustering algorithm HCf+ (Ferrara, Genta, & Montanelli, 2013). The aim of the system is to: 1) provide a thematic and a high level view of data created through the methods of similarity based combination, 2) provide toolkits for intuitive keyword based and SQL related query languages in order to help inexperienced users to query the linked data with RDF query languages, such as SPARQL and MQL.

The features of the inWalk system can be described as 1) abstraction and combination: the system offers a conceptual view of data through the concept of inCloud, which is the outcome of an abstraction process that is built on linked data combination; 2) exploration by walks: the thematic walk and the inside walk are the two potential movements that the system relies on to discover the linked data; 3) filtering by patterns: filtering actions through the inCloud structure is promised by the system which allows users to explore the parts of inCloud with a specific choice principle.

The system consists of two main components, the interface and the engine. The engine converts a collection of linked data retrieved from a certain repository, such as DBpedia or freebase, into an equivalent inCloud cluster. Subsequently, the HTML5 interface performs discovery walk on the clusters.

The interface of the system allows user to search and discover a certain inCloud cluster through either a thematic walk or inside walk. In the thematic walk, users build their thematic paths concerning their interest from the default available inCloud clusters without querying the system. While the inside walk is achieved by typing any interested topic into the search box and selecting the cluster of interest from an autocomplete list.

DISCUSSION

Table 1 provides an overview of the systems that have been reviewed in the previous section with additional features summarized from (Marie & Gandon, 2014). As shown in the table, all systems except for inWalk, use DBpedia as main datasets to extract data from. SWOC, LED and Aemoo additionally utilize external services such as search engines and tagging systems to compute associations between DBpedia nodes. DBpedia lookup service has been notably used by most of the systems as query model. Aemoo and Discovery Hub systems make use of linked data for multi-purpose search systems where all types of information can be retrieved without directing to any specific domain. All of the systems offer breadcrumb feature (history) using sessions which helps users easily retract to their previous state allowing comparison and analysis of the results. This feature reduces the load on the servers by avoiding sending the already sent request twice. Yovisto, Seevl and Discovery Hub offer registration feature as well, this keeps search history for registered users at the server side.

Explanation of results helps users to grasp and understand the space of retrieved information. This feature is proposed by Aemoo, Seevl, Discovery Hub and Linked Jazz systems using textual and graph modes. Ranking of the results are not provided by Aemoo, Linked Jazz and inWalk since the results are in the graph form. The rest of the systems offer results ranking in various forms. SWOC uses the node size of the graph for the retrieved results where node size increases proportionally with the relevancy. Yovisto and Seevl offer a list of thumbnail picture style rankings; LED ranks results as list of texts, whereas Discovery Hub ranks results into list of categories with thumbnail pictures.

Table 1. Overview of advanced exploratory search systems

System Name	Yovisto	SWOC	LED	Aemoo
Release Date	2009	2010	2010	2012
Web Address	www.yovisto.com	sisinflab.poliba.it/semantic-wonder-cloud	sisinflab.poliba.it/led	wit.istc.cnr.it/aemoo
Main Data	DBpedia EN+DE	Dbpedia EN	Dbpedia EN	DBpedia EN
Auxiliary Data	No	Search Engines & Tagging Systems	Search Engines & Tagging Systems	External services
Query Model	Keyword search	Dbpedia lookup	Keyword search	Lookup
Matching	String-match	Direct match (lookup)	String matching	Direct match (lookup)
Domain in (Purpose)	Academic Videos	IT Domain	ICT	General
Database Method	Freebase Parallax	SWOC Storage	LED Storage	Knowledge Pattern (KP) Repository Manager
Databse Purpose	map queries to entities	Stores popularity & similarity values for pairs of resources	Stores results computed by Ranker	responsible for the storage, indexing & fetching of KPs
Principal Layout	Query suggestions	Graph	Tags cloud	Graph
Results Explanations	No	No	No	Wikipedia-based
Breadcrumb	Sessions & registeration	Sessions	Sessions	Sessions
Algorithms	Set of heuristics	DBpedia Ranker	DBpedia Ranker	EKP filtered view
Ranking	Yes	Yes, graph size	Yes	No
Offline Processing	Yes	Yes, similarity of pairs	Yes	Yes , EKP part
API	RDF triple-store	No	Yes, RESTful	Yes, RESTful
Faceted Navigation	Yes	No	No	No
User Interface	HTML	Flash based	HTML	HTML

To compute similarity values for resources SWOC and LED integrate DBpediaRanker into their system. Yovisto uses a set of heuristics to determine the most relevant resources related with a given DBpedia entity. Aemoo uses EKP which specifies the relevant types of resources and defines the typical classes used to explain entities of a certain type. LDSD and DBrec algorithms are used by Seevl to retrieve artist recommendations. Discovery Hub is based on a semantic spreading activation algorithm coupled to a graph sampling technique. Linked Jazz uses Mapping tool to build Linked Jazz Name Directory dataset which is refined by the Curator Tool. InWalk system uses HCf+ clustering algorithm to build a high level graph called inCloud. All of the systems, excluding Discovery Hub, needed an offline environment for different tasks processing; whereas results selection and ranking in Discovery Hub is executed on the fly without any pre-processing requirements.

Yovisto, Seevl and Discovery Hub systems offer faceted navigation which let the users to explore a collection of retrieved information by applying multiple filters. Various database methods have been used with the systems for different purposes. Yovisto use Freebase Parallax to let users' query find a set of entities then find the videos for these entities. SWOC and LED use Storage which stores similarity values between the pairs of resources and the "popularity" of each resource in a DBMS for the purpose of an efficient retrieval at runtime. Aemoo uses Knowledge Pattern (KP) Repository Manager that is responsible for the storage, indexing and fetching of KPs. A Scalable RDF/NoSQL storage powered by OpenLink Virtuoso has been utilized by Seevl for scalability purposes. Discovery Hub uses Virtuoso for user accounts and follows. Linked Jazz created a directory of personal names of jazz artists paired with URIs called Name Directory. InWalk repository is created by inWalk system to provide a high-level view of relevant linked data.

Table 2. Overview of advanced exploratory search systems (continued)

System Name	Seevl	Discovery Hub	Linked Jazz	inWalk
Release Date	2012	2013	2013	2014
Web Address	play.seevl.fm	http://discoveryhub.co	https://linkedjazz.org	http://islab.di.unimi.it/inwalk
Main Data	Dbpedia, Freebase & MusicBrainz	DBpedia EN, FR, IT	Dbpedia -> Linked Jazz DB	Freebase & Twitter
Auxiliary Data	No	No	No	No
Query Model	Lookup	Dbpedia Lookup	Manual selection from a list	Lookup & selection
Matching	Direct match (lookup)	Direct match (lookup)	Selection	Direct match (lookup)
Domain in (Purpose)	Music	General	Jazz Musicians	Athletes or Twitter News
Database Method	Scalable RDF/ NoSQL storage by OpenLink Virtuoso	Virtuoso, MySQL	Linked Jazz Name Directory	inWalk repository
Databse Purpose	Scalability	User account & follows	Stores individuals represented by literal triples	Provides a high-level view of relevant linked data
Principal Layout	List	List	Graph	Graph
Results Explanations	Shared properties	Yes, Text & Graph	Yes, Interview Transcripts	No
Breadcrumb	Sessions & registeration	Sessions & registeration	Sessions	Sessions
Algorithms	LDSD, DBrec algorithm	Semantic spreading activation	Mapping & Curator Tool & the Transcript Analyzer	HCf⁺ clustering algorithm
Ranking	Yes	Yes	No	No
Offline Processing	Yes	No, on the fly	Yes, Linked Jazz DB	Yes
API	Yes, Content negotiation JSON-LD	No	Yes, JSON, RDF & GEXF files	No
Faceted Navigation	Yes	Yes	No	No
User Interface	HTML, ajax	HTML	HTML5 + jQuery	HTML5 + JS

CONCLUSION

The ease of uploading data online caused a proliferation of data on the Web providing new challenges to return relevant data. The evolution of the syntactic Web to the Semantic Web technology increases the efficiency of searching approaches by allowing automated semantic processing of data through exploratory search. The interest in this area will continue to expand constituting a critical improvement for the future of the search experience and its results.

REFERENCES

Berners-Lee, T. (2006, July 27). Linked Data - Design Issues. Retrieved January 9, 2015, from http://www.w3.org/DesignIssues/LinkedData.html

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*(5), p28.

Castano, S., Ferrara, A., & Montanelli, S. (2014). inWalk: Interactive and Thematic Walks inside the Web of Data. Presented at the *EDBT*, 628–631.

Dimitrova, V., Lau, L., Thakker, D., Yang-Turner, F., & Despotakis, D. (2013). Exploring Exploratory Search: A User Study with Linked Semantic Data. In *Proceedings of the 2Nd International Workshop on Intelligent Exploration of Semantic Data*, (pp. 2:1–2:8). NY, USA: ACM.

Fensel, D. (2005). *Spinning the Semantic Web: bringing the World Wide Web to its full potential*. MIT Press.

Ferrara, A., Genta, L., & Montanelli, S. (2013). Linked Data Classification: A Feature-based Approach. In *Proceedings of the Joint EDBT/ICDT 2013 Workshops*, 75–82. NY, USA: ACM.

Heath, T., & Bizer, C. (2011). *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Morgan & Claypool. Retrieved from http://linkeddatabook.com/editions/1.0/#htoc8

Hitzler, P., Krotzsch, M., & Rudolph, S. (2011). *Foundations of semantic web technologies*. CRC Press.

Jiang, T. (2014). Exploratory Search: A Critical Analysis of the Theoretical Foundations, System Features, and Research Trends. In *Library and Information Sciences*, 79–103. Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-54812-3_7

Klyne, G., Carroll, J. J., & McBride, B. (2004, February 10). Resource Description Framework (RDF): Concepts and Abstract Syntax. Retrieved January 15, 2015, from http://goo.gl/8P7WyQ

Le Hors, A., Nally, M., & Speicher, S. (2012). Using read/write Linked Data for Application Integration-Towards a Linked Data Basic Profile. Presented at the *LDOW*.

Madhu, G., Govardhan, D. A., & Rajinikanth, D. T. (2011). Intelligent Semantic Web Search Engines: A Brief Survey. *arXiv Preprint arXiv:1102.0831*. Retrieved from http://arxiv.org/abs/1102.0831

Marchionini, G. (2006). Exploratory Search: From Finding to Understanding. *Commun. ACM*, *49*(4), 41–46. http://doi.org/10.1145/1121949.1121979

Marie, N., & Gandon, F. (2014). Survey of linked data based exploration systems. Presented at the *IESD 2014-Intelligent Exploitation of Semantic Data*.

Marie, N., Gandon, F., Ribière, M., & Rodio, F. (2013). Discovery Hub: On-the-fly Linked Data Exploratory Search. In *Proceedings of the 9th International Conference on Semantic Systems*, 17–24. New York, NY, USA: ACM. http://doi.org/10.1145/2506182.2506185

Mirizzia, R., Di Noiaa, A. R. T., & Di Sciascioa, E. (n.d). Lookup, Explore, Discover: how DBpedia can improve your Web search.

Mirizzi, R., Ragone, A., Noia, T. D., & Sciascio, E. D. (2010). Semantic Wonder Cloud: Exploratory Search in DBpedia. In F. Daniel & F. M. Facca (Eds.), *Current Trends in Web Engineering* (pp. 138–149). Springer Berlin Heidelberg.

Musetti, A., Nuzzolese, A. G., Draicchio, F., Presutti, V., Blomqvist, E., Gangemi, A., & Ciancarini, P. (2012). Aemoo: Exploratory search based on knowledge patterns over the semantic web. *Semantic Web Challenge*.

Passant, A. (2012). seevl: mining music connections to bring context, search and discovery to the music you like. From challenge.semanticweb.org/submissions2011/swc2011_submission_11.pdf

Pattuelli, M. C., Miller, M., Lange, L., Fitzell, S., & Li-Madeo, C. (n.d). The Code4Lib Journal – Crafting Linked Open Data for Cultural Heritage: Mapping and Curation Tools for the Linked Jazz Project [Journal]. Retrieved February 5, 2015, from http://journal.code4lib.org/articles/8670

Jenice Aroma R., & Kurian, M. (2013). A Survey on Tools essential for Semantic web Research. *International Journal of Computer Applications*, *62*(9), 26–29.

Waitelonis, J., & Sack, H. (2012). Towards Exploratory Video Search Using Linked Data. *Multimedia Tools Appl.*, *59*(2), 645–672. http://doi.org/10.1007/s11042-011-0733-1

White, R. W., Kules, B., & Bederson, B. (2005). Exploratory search interfaces: categorization, clustering and beyond: report on the XSI 2005 workshop at the Human-Computer Interaction Laboratory, University of Maryland (Vol. 39, pp. 52–56). Presented at the *ACM SIGIR Forum*, ACM.