Journal of Arts & Humanities

Volume 07, Issue 02, 2018, 59-65
Article Received: 14-02-2018
Accepted: 24-02-2018
Available Online: 03-03-2018
ISSN: 2167-9045 (Print), 2167-9053 (Online)
DOI: http://dx.doi.org/10.18533/journal.v7i2.1342

Building Social Interactions as a Creation of Networks in an RDF Repository

Eun G Park[1], Matthew Milner[2]

ABSTRACT

Humanities scholars are not likely to be thinking about their research findings as data, and the predominant models of organizing documents remain generally archival or bibliographic in nature for text-based documents. Although the linked data movement has greatly influenced information organization and search queries on the Web, in comparison to other fields, the adoption of the linked data approach to humanities collections is unequally paced. This study intends to explain how people or actors make social interactions, and how social interactions are formed in a type of network through the example of the Making Publics (MaPs) project. The objective of the MaPs project is to build collaborative common environments for tracing social interactions between people, things, places and times. To build social interactions, the Networked Event Model was designed in a collaborative environment. Events were defined as six types of nodes (e.g., people, organizations, places, things, events, and literals) in the RDF (Resource Description Framework) triple statements. The interaction vocabulary list is made of 173 verbs and predicates, offering 510 traceable events. The RDF repository runs on a Sesame server and MySQL architecture. Users can use digital tools to select and document events and visually present the selected events in interactive social web forms. The MaPs project sought to extract the network extant in the works of prose in large collaborative humanities documents. In this way, the dissemination of and access to humanities data can be made more connectable, available and accessible to both academic and non-academic communities.

Keywords: Event, Making Publics, Networks, RDF Repository, Social Interactions.
JEL classification: C88, C90, D83, D85.

1. Introduction: From the point of view of social interactions

As the ubiquity of digital media and technology alters the presentation of traditional humanities documents and artifacts in profound ways, existing data models are not yet well-suited to displaying

[1] Associate Professor, School of Information Studies, McGill University, Montreal, Canada. Email: eun.park@mcgill.ca
[2] Digital Historian for the Nano History project and Memorial University of Newfound, St. John's, NL. Canada. Email: m@matthewmilner.name

texts. With the introduction of a linked data approach in the field of Information Science, the linked data movement has greatly influenced information management and search queries and the Web has changed to a Web of data (W3C, 2015). *Linked Data* refers to "data published on the Web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked to from external data sets" (Bizer, Heath & Berners-Lee, 2009, p.2). Linked data approaches use technologies such as Uniform Resource Identifier (URI) (Berners-Lee et al., 2005), HyperText Transfer Protocol (HTTP) (Berners-Lee et al., 2006), RDF (Resource Description Framework) (Brickley & Guha, 2004) and others. Some cultural heritage institutions, including libraries, archives and museums, have implemented these technologies and transformed their legacy data to linked data in order to improve access to their collections. An example of a library community is the Canadian Linked Data Initiative, which is an association of Canadian university libraries, major Canadian libraries and archives, and aims to collaborate and coordinate the metadata production activities of "…technical services units, special collections and digital libraries in both English and French" (Admin & McOrmond, 2018, Canadian Linked Data Initiative Home). In addition, the Linked Open Data in Libraries, Archives and Museums (LODLAM) promotes the linked open data approach by making collaborative works with various professionals and user groups (Linked Open Data in Libraries, Archives and Museums, n.d.). The Europeana Linked Open Data project provides a single access point for books, paintings, films and archival records for cultural institutions in Europe and allows data to be semantically linked across European institutions (Isaac & Haslhofer, 2013; Europeana, 2013). The Bibliographic Framework (BIBFRAME) Initiative was launched in 2011 by the Library of Congress to replace libraries' cataloging records with a linked data model (Miller et al., 2012). These examples encourage the promotion of the linked data movement to their member institutions. Today when searching for information on the Web, users tend to search Google quickly rather than using libraries' catalogues. Thus, the institutions' collections should become more accessible, structured and semantically linked.

However, some humanities scholars are not likely thinking about their research findings as data and tend to be hesitant to adopt rapid technologies related to linked data. In the humanities fields, the predominant model of organizing documents is still in collections of artifacts, since text-based documents are generally archival or bibliographic in nature. To some extent, the Web seems to amplify the problems of organizing and presenting humanities documents (Yakel *et al*., 2007). To find humanities information, humanities resources have been treated separately from Web resources. Therefore, accessing these resources requires physical visits to the holding institutions or entering an institution site; users have less and less incentive to search the institution site separately (Coyle, 2010). Because the linked data method requires technology capabilities and trained personnel, in comparison to other fields, the adoption of the linked data approaches in humanities and arts collections is unequally paced; this study intends to encourage it more actively.

In the form of networks on the Web, the term social interaction or social networking has emerged as a point of interest. To conceptually understand data as social interactions, in sociology social interactions indicate a sequence of social actions between individuals or groups forming a social relationship (Decoster, 2002). In cultural-historical activity theory, human activity is situated within historical and cultural contexts (Foot, 2014) and activity is what subjects (e.g. people) do together, as human activities are constantly modified by cultural and historical influences. In an activity system, a subject (or actor) and an object interact with each other through the use of tools that are employed by the subject to act upon an object or pursue a desired outcome. So, social interactions are considered a recreation of agents in texts. In archival science, where texts are dynamic and fluid, social actions are considered "persistent representations of occurrents" (Yeo, 2008, p. 136). Thus, an activity, action or event is central to activities and texts in building the context of the socio-cultural aspects of agents in the form of a network.

In line with this view, in the linked data approach, making a network is a means to understanding how groups of people build links around shared interests and dissipate objects around those activities of interest. In other words, people who appear in texts become agents who form networks through the use of and engagement with other agents on specific works of drama, portraits, political controversy, or certain discoveries, etc. They do not arise from the frequency of certain kinds of social and artistic phenomena; rather, interest is focused on interactions between human agents, who appear in particular places at a particular time from those texts or artifacts. Tracing these interactions is especially important

and requires a radically different approach because it inherently focuses on the nature of human interaction over time, not just on mere documents or artifacts themselves. Therefore, a reconceptualization of texts through networks is building a context in which the migration of the text travels through a collection and into an archive, which transforms it from prose-based text into a live network, and finally disseminates it to humanities scholars and users.

This study aims to explain how people or actors make social interactions and how the social interactions are formed in networks through the example of the Making Publics (MaPs) project. This approach is a means to understanding people, things, places and times in the text-based documents of the humanities and arts. It places a strong emphasis on the interactions of people and things, along with the need to account for their changes over time. The MaPs project uses a linked data method by designing an event model, defining events and event types and identifying the interaction vocabularies through the use of digital tools (e.g. event selection tools and a visual display tool). The outcome of this project will contribute to promoting linked data studies to the humanities fields, improving the dissemination of and access to humanities data, and making it more accessible to wide user groups. This study encourages the adoption of linked data technologies for making collections of humanities and arts more linkable, connectable, and approachable to the Web. The uniqueness of this study is in its attempts to adopt the linked data approach to humanities collections and develop two digital tools as a protocol for event selection and visualization when extracting events in an RDF repository.

This study is organized into five sections. The second section explains the method of the MaPs project by building networks as a collaborative environment. The third section explains the procedures for building social interactions. The fourth section details the results of the project. The last section discusses relevant issues regarding the building procedures of social interactions and concludes with further study plans.

2. Method: Making publics in an environment of collaborative interaction

To illustrate social interactions, the Making Publics (MaPs) project presents the procedures and results of shaping social interactions. To understand the lives of people in Europe during 1500-1700, a large collection of texts and documents are collected about normal people, places and artifacts in Europe during the given time period. The database contains nearly 2000 works cited in the form of articles, papers, books, and essays. The collections include plentiful references to people, places, things, and their interactions on various topics in early modern Europe. The key aims of the MaPs project are: 1) to build collaborative common environments (also known as the context) for tracing social interactions between people; and 2) to create links between different types of texts in the existing architecture. Technically, this builds a topically construed middleware environment drawn from user interactions in the direction that users want it to go in. Existing event models, like those from Schema.org (n.d.) or Music Ontology Tools (2018), view the event or interaction as a noun. However, our approach sees an interaction as a socio-cultural process that needs to preserve agents with start and end points. In our approach, any form of interaction between agents or states of agents is extracted as an event and is supposed to be included for modeling. To model the extracted interactions, events are defined as unnamed entities that encapsulate an agent, a verb, and an object.

3. Procedures for building social interactions

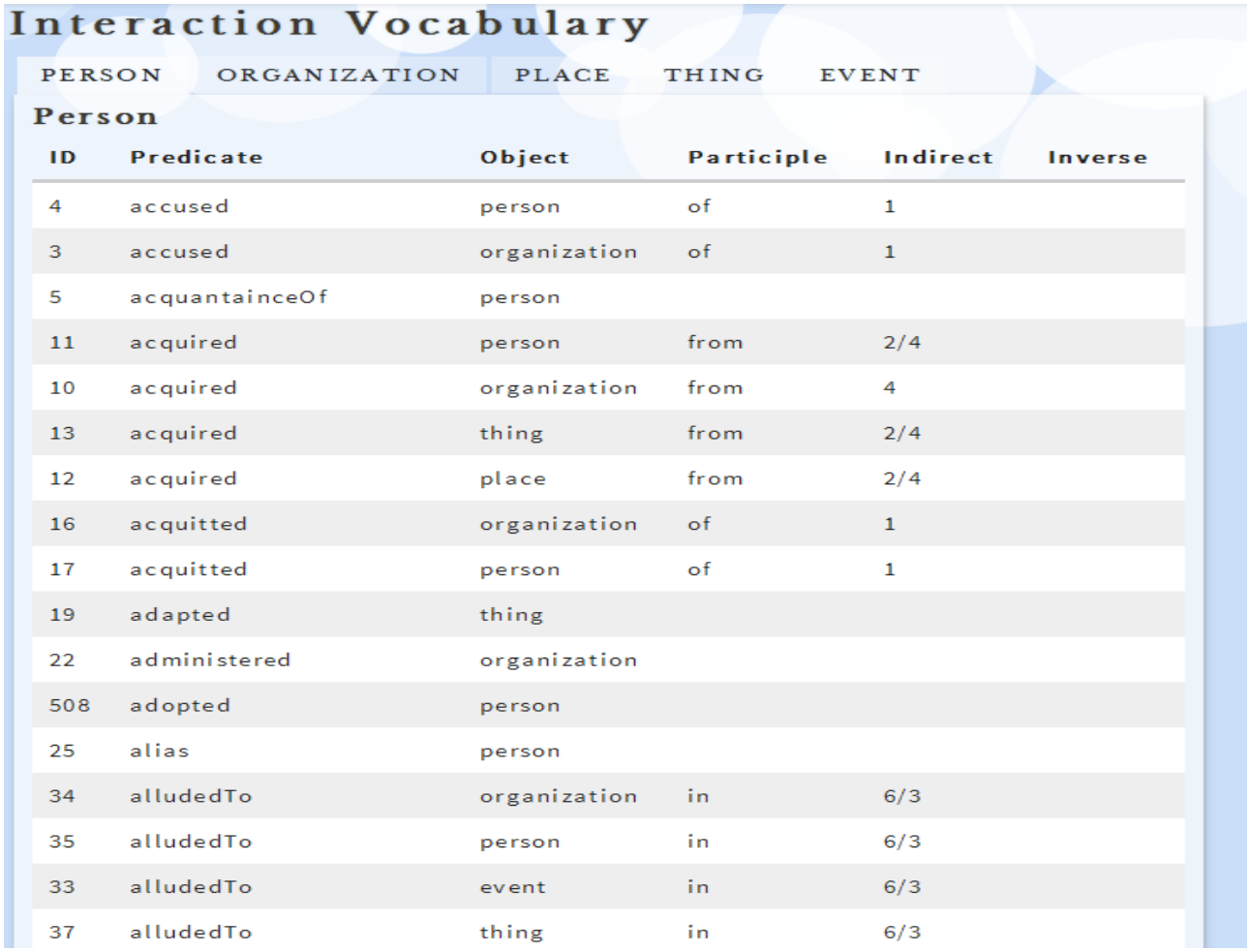
First, at the core of our design lies the Networked Event Model (NEM), which is a prototype for describing forms of social interactions between agents (i.e. people) and objects (i.e. a thing) that can account for change over time. In the social network approach, something that agents do becomes an event or interactions within a framework of social contexts. Thus, the primary tasks are identifying agents first and then identifying six kinds of events: people, organizations, places, things, events, and literals. NEM uses the RDF (Resource Description Framework) triple statements and adds a unique identifier, essentially modeling it as an RDF Quad or a named graph:

ID: agent → verb → object.

The use of quads clearly identifies the interactions. Using identifiers also enables the nesting of events, turning events themselves into entities that can be referenced as a part of larger and more complex networks of interactions. The resulting datatype or entity that aims to represent an unnamed

event as an interaction is carried out by an agent on an object within a given timeframe. Practically speaking, NEM limits the kinds of agents and objects possible in a network to nouns, such as people, organizations, places, and things or artifacts.

Next, to provide a means of documenting concrete forms of interactions, the NEM interaction vocabulary is produced. The list focuses on common verbs in assertive actions, such as 'wrote', 'killed', 'sued' rather than subjective, affective, or epistemological ones (e.g. knows, might have eaten, loves, etc.). Subject specialists in English and History have worked through scenarios based on categories of interactions for each datatype (e.g. 'creation', 'destruction', 'possession', 'transactions', or 'movement', etc.). Figure 1 demonstrates the interaction vocabulary list that includes 173 verbs for predicates and 510 traceable interactions, participles for transactions and content, and inverse for two-way relationships and indirect objects.



Interaction Vocabulary

PERSON ORGANIZATION PLACE THING EVENT

Person

ID	Predicate	Object	Participle	Indirect	Inverse
4	accused	person	of	1	
3	accused	organization	of	1	
5	acquantainceOf	person			
11	acquired	person	from	2/4	
10	acquired	organization	from	4	
13	acquired	thing	from	2/4	
12	acquired	place	from	2/4	
16	acquitted	organization	of	1	
17	acquitted	person	of	1	
19	adapted	thing			
22	administered	organization			
508	adopted	person			
25	alias	person			
34	alludedTo	organization	in	6/3	
35	alludedTo	person	in	6/3	
33	alludedTo	event	in	6/3	
37	alludedTo	thing	in	6/3	

Figure 1. Interaction vocabulary

Dates are stored alongside all of the data, allowing for multiple users to insert multiple dates for any datatype. By breaking down the entire statement into components, the RDF Schema is designed, applied and published as social interactions in the Web environment. The RDF is used as an independent repository running on a Sesame server and MySQL architecture. NEM provides a means for users to document both options as simply adding new data to the system.

4. Results

4.1 User-based network design and tools

The MaPs allows users to build their own historical or social networks and express and manage the interests of a community of humanities scholars and users in a network model. Thus, the environment can be realized as a tool that users can use to identify and mediate between users, the data of users, and existing data resources. The MaPs designed an easy way for often digitally-hesitant users to comprehend

the event and make links in the model. The MaPs provides an interface that would allow users to easily and quickly select, document events and efficiently visualize the models for their networked work. For example, new methods of inputting and visualizing the presentation of texts are used. Figure 2 shows that Workbench is a 'drag and drop' event builder for users to select nodes to be used for an event, and then it filters verbs accordingly.

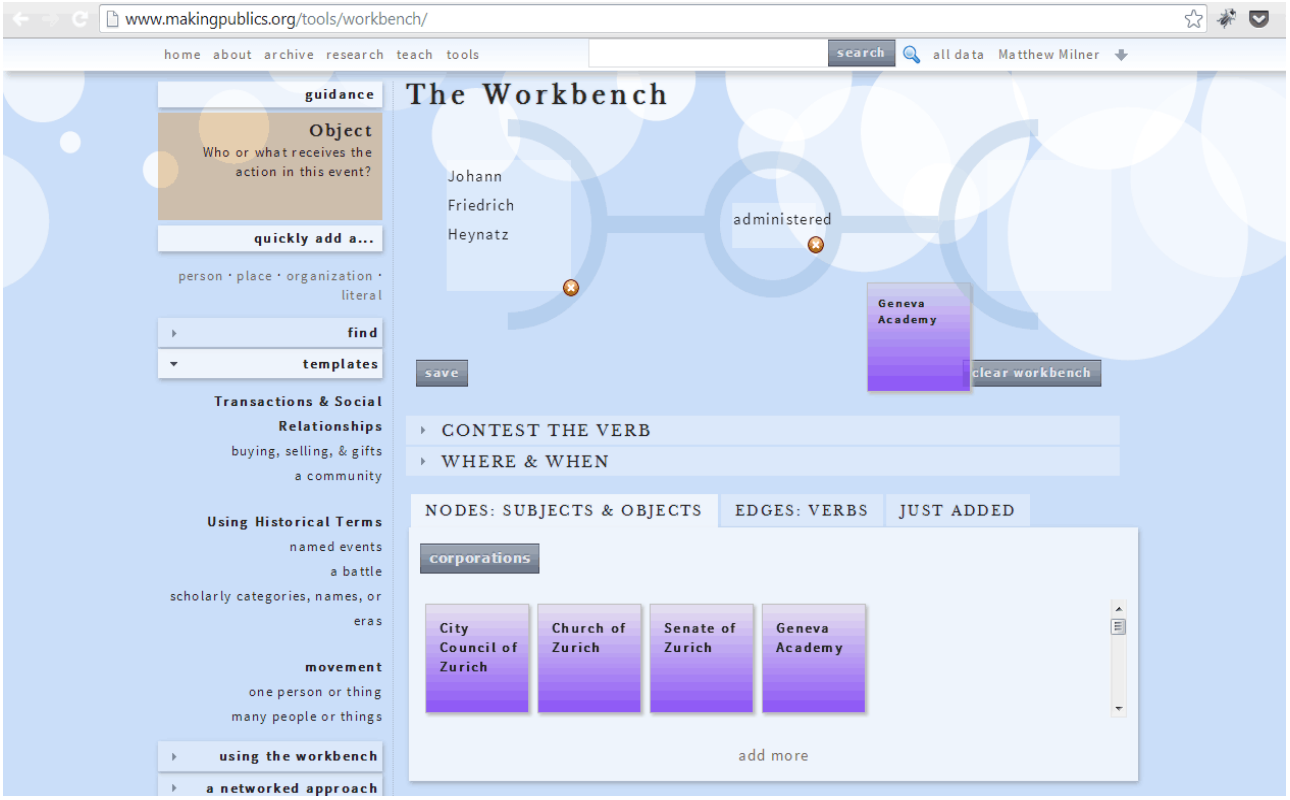


Figure 2: Workbench

This function can control user input and contest verbs, define dates and locations, and create complex events.

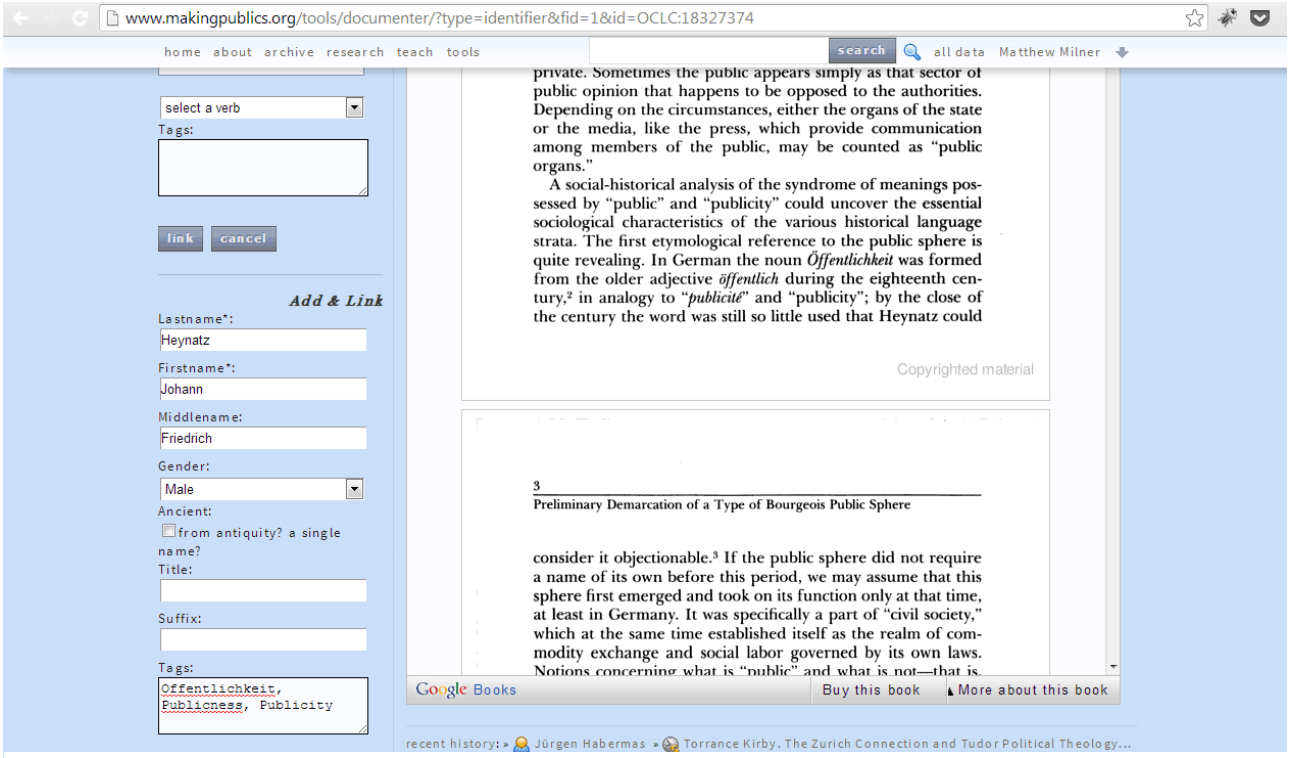


Figure 3. Documenter

Figure 3 illustrates that the Documenter allows a user to manually select and document entities in texts from other sources, from websites or PDF files.

The network representations of the input text can be visualized through Network Webs that use D3js in interactive forms in web browsers.

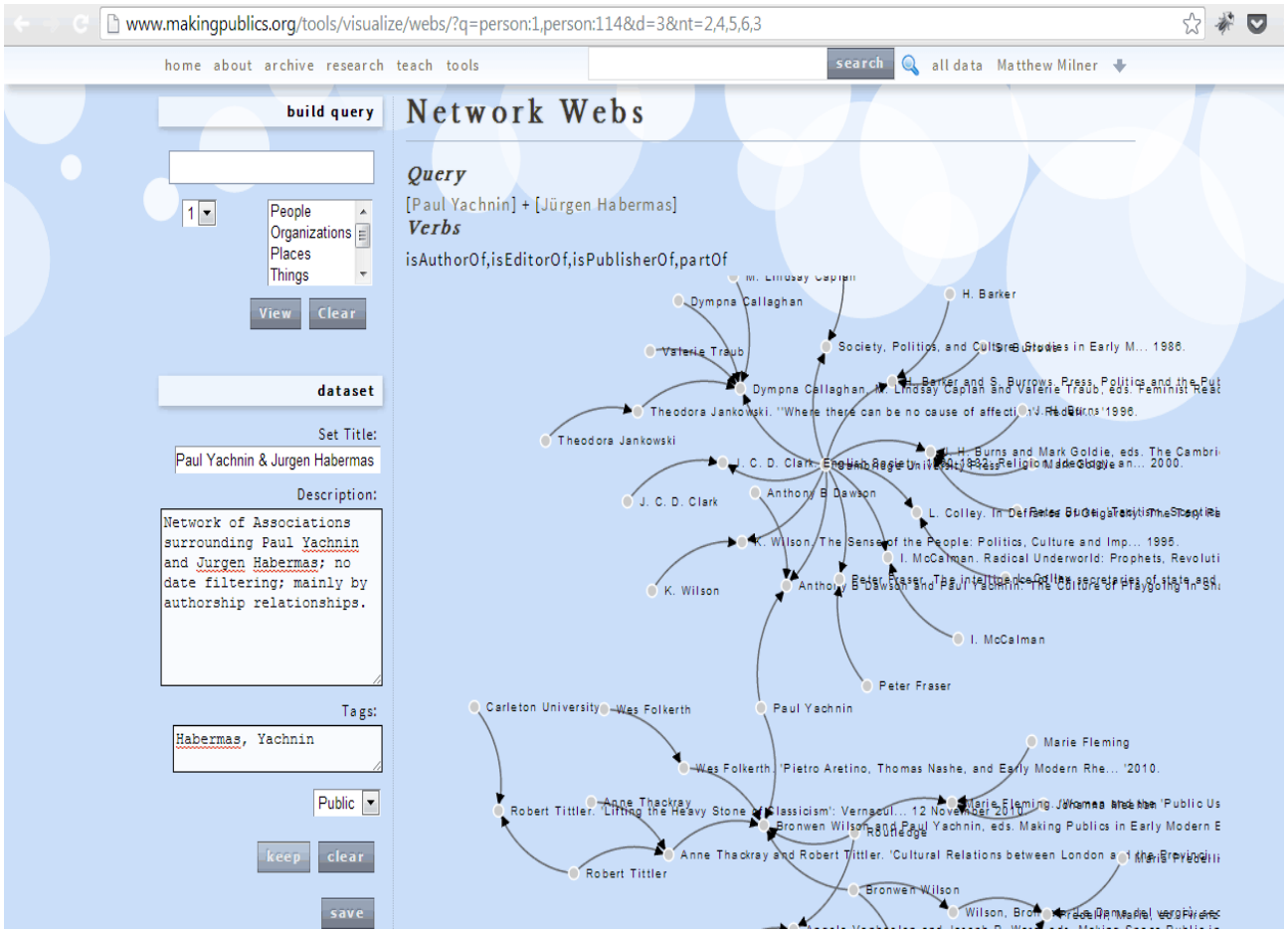


Figure 4: Network webs

5. Discussion

To present social interactions, the MaPs project builds network interactions between agents and things as events with different kinds of text or artifacts. As the Networked Event Model (NEM) is fully deployed in the MaPs environment, it comprises our six node types of events (e.g., people, organizations, places, things, events, and literals). The interaction vocabulary list includes 173 verbs and predicates, offering 510 traceable events. The RDF schema is adopted to create events and make a repository for handling networked humanities data. The MaPs project seeks to extract the network from the prose works of large, collaborative humanities documents.

In this project, the way in which RDF principles allow humanities researchers to consider humanistic knowledge as networked data is important. Turning the MaPs data into an RDF repository can reform the traditional way of dissemination and access to humanities research findings themselves, linking ongoing humanities research and making it more dynamic and integrated with other data. Therefore, humanities data will become more collaborative and interdisciplinary and humanities research can be made available and accessible to both academic and non-academic communities outside of traditional prose-based forms of dissemination.

Thus, the project is a good way to transform texts and artifacts in the humanities fields, including information science, literature, language, history, music, music technology, geography and more. It is recommended that humanities scholars should promote the use of linked data approaches to organize humanities and arts collections.

There was a challenge when recording events automatically in NEM. Although each option for making an automatic or manual selection has certain limitations, a combination of these two options is a

reasonable compromise in order to reach a better outcome. For further study, text extraction techniques can be used to automatically process extraction for event selection in a large amount of collections. In addition, it could be expanded upon to build social networks with more entities, such as entities with temporal and spatial relationships, or sub-entities with diverse kinds of topics.

References

Admin, J. & McOrmond, R. (January 23 2018). Canadian Linked Data Initiative Home. Retrieved 14 February 2018 from https://connect.library.utoronto.ca/display/U5LD/Canadian+Linked+Data+Initiative+Home.

Berners-Lee, T., Fielding, R. & Masinter, L. (2005). Universal Resource Identifiers (URI): Generic Syntax. Retrieved 14 February 2018 from http://www.ietf.org/rfc/rfc3986.txt.

Berners-Lee, T., Shadbolt, N. & Hall, W. (2006). The semantic web revisited. IEEE Intelligent Systems, 21 (3), 96-101.

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data: The story so far. International Journal on Semantic Web and Information Systems, 5, (3), 1-22.

Brickley, D. & Guha, R. (2004). RDF Vocabulary Description Language 1.0: RDF Schema. Retrieved 14 February 2018 from http://www.w3.org/TR/rdf-schema.

Coyle, K. (2010). Understanding the semantic web: bibliographic data and metadata. Library Technology Reports, 46(1), 1-31.

Decoster, V. A. (2002). Predicting emotions in everyday social interactions: a test and comparison of affect control and social interactional theories. Journal of Human Behavior in the Social Environment, 6(3), 53-73.

Europeana. (2013). Europeana Data Model primer. Retrieved 14 February 2018 from http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf.

Foot, K. A. (2014). Cultural-historical activity theory: exploring a theory to inform practice and research. Journal of Human Behavior in Social Environments, 24(3), 329-347.

Isaac, A. & Haslhofer, B. (2013). Europeana linked open data-data.europeana.eu. Semantic Web, 4(3), 291-297.

Linked Open Data in Libraries, Archives and Museums (LODLAM). (n.d.). About. Retrieved 14 February 2018 from http://lodlam.net/about.

Making Publics. (n.d.). Making Publics 1500-1700 Media, Markets & Association in Early Modern Europe. Retrieved 14 February 2018 from https://www.matthewmilner.name/projects/maps or http://makingpublics.mcgill.ca.

Miller, E., Ogbuji, U., Mueller, V. & MacDougall, K. (2012). Bibliographic framework as a web of data: linked data model and supporting services." Washington, DC: Library of Congress. 2012. Retrieved 14 February 2018 from https://www.loc.gov/bibframe/pdf/marcld-report-11-21-2012.pdf.

Music Ontology Tools. (2018). Project Activity. Retrieved 14 February 2018 from https://sourceforge.net/projects/motools.

Schema.org. (n.d.). Organization of Schemas. Retrieved from http://schema.org.

Semantic Services, Interoperability and Web Applications: Emerging Concepts, 205-277.

W3C (2015). Linked Data. Retrieved 14 February 2018 from https://www.w3.org/standards/semanticweb/data.

Yakel, E., Shaw, S., & Reynolds, P. (2007). Creating the next generation of archival finding aids. D-Lib Magazine, 13(5/6), Retrieved 14 February 2018 from http://www.dlib.org/dlib/may07/yakel/05yakel.html.

Yeo, G. (2008). Concept of record (2): prototypes and boundary objects. The American Archivists, 71(1), 118-143.