Visualizing Intellectual Connections among Philosophers Using the Hyperlink & Semantic Data from Wikipedia

Sofia J. Athenikos
College of Information Science & Technology
Drexel University
Philadelphia, PA, USA
1-215-895-2474

sofia.j.athenikos@acm.org

Xia Lin
College of Information Science & Technology
Drexel University
Philadelphia, PA, USA
1-215-894-2482

xlin@cis.drexel.edu

ABSTRACT
Wikipedia, with its unique structural features and rich user-generated content, is being increasingly recognized as a valuable knowledge source that can be exploited for various applications. The objective of the ongoing project reported in this paper is to create a Web-based knowledge portal for digital humanities based on the data extracted from Wikipedia (and other data sources). In this paper we present the interesting results we have obtained by extracting and visualizing various connections among 300 major philosophers using the structured data available in Wikipedia.

Categories and Subject Descriptors
E 1.1 [**Data Structures**]: *graphs and networks*; H 3.5 [**Information Storage and Retrieval**]: Online Information Services – *web-based services*; H 5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia – *navigation, user issues*.

General Terms
Algorithms, Design, Human Factors.

Keywords
Wikipedia, Digital Humanities, Visualization, Social Network.

1. INTRODUCTION
Since its inception in 2001, Wikipedia (http://www.wikipedia.org) has become one of the most sought-after resources on the Web. The manner of information search and retrieval on Wikipedia, however, has remained that of conventional keyword-based search and retrieval. The objective of the ongoing project reported in this paper is to extract and visualize facts, relations, and networks involving influential intellectual/cultural figures using Wikipedia (and other data sources). By doing so, we aim at creating a useful Web-based portal for digital humanities. In this paper we present some of the early results we have obtained [1-3] concerning 300 major philosophers, with a focus on visualizing their connections.

2. WIKIPEDIA DATA EXTRACTION
We obtained a list of 300 major philosophers from Wikipedia's "Timeline of Western Philosophers" page. We extracted hyperlink and semantic data on the philosophers from their individual article pages, and stored the data in a MySQL database. We retrieved info needed for visualization by querying the database, and stored the results as XML files marked up with GraphML and TreeML.

3. PHIL NETWORK VISUALIZATION
In this section we illustrate discovering and exploring connections among philosophers via interactive information visualization using the hyperlink and semantic data extracted from Wikipedia. The visualization is done by using Prefuse (http://prefuse.org/).

3.1 Strongest Link Paths (SLP) Method
The method that we use to effectively highlight the most dominant nodes and their interconnections within a network is a graph simplification method we have developed, called the *strongest link paths* (SLP). Here we use two variations of SLP. In the case of strongest link/influence networks, we select, for each node, only the link with the highest hyperlink/influence count. In the case of extended link/influence networks, we add all 1st-degree links and, for each subsequent degree, we add links to nodes that are not already covered, thereby eliminating overlapping edges.

3.2 Strongest Hyperlink/Influence Networks
Figures 1-7 show close-ups of the largest clusters in the strongest bidirectional-link network graph. Overall, each cluster represents a coherent intellectual theme, reflecting significant intellectual connections within/between (a) philosophical school/tradition(s). For example, the cluster in Fig. 1, with Hegel, Marx, & Kant as central nodes, contains many figures related to German idealism and Marxism. The cluster in Fig. 2, with Descartes & Leibniz (w/ Spinoza), represents the rationalist tradition, whereas the cluster in Fig. 3, with Locke & Hume (w/ Berkeley), represents the empiricist tradition. (The link between Hobbes and Descartes in the cluster in Fig. 2 is an example of connection in terms of contrast.) The cluster in Fig. 4, with Russell, Wittgenstein, & Carnap, is representative of the analytical philosophical tradition, while the cluster in Fig. 5, with Heidegger, Husserl, & Derrida, is representative of the continental philosophical tradition. The cluster in Fig. 6, with Plato, Aristotle, & Socrates, mostly consists of classical Greek philosophers and those of the Hellenistic and Roman eras. The cluster in Fig. 7, with Aquinas (w/ Augustine & Anselm), represents the theological tradition.
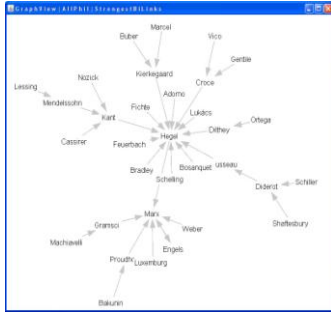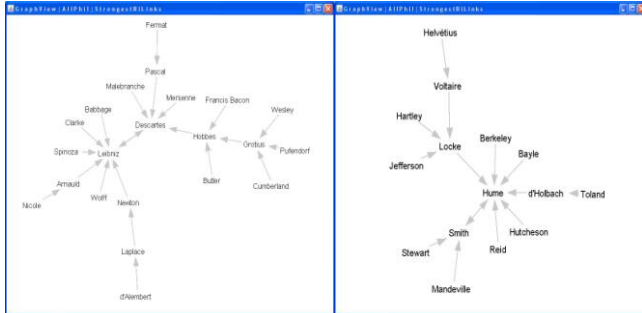
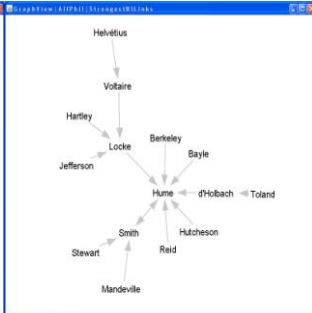**Figure 1. Hegel cluster.**

Figure 2. Descartes cluster.
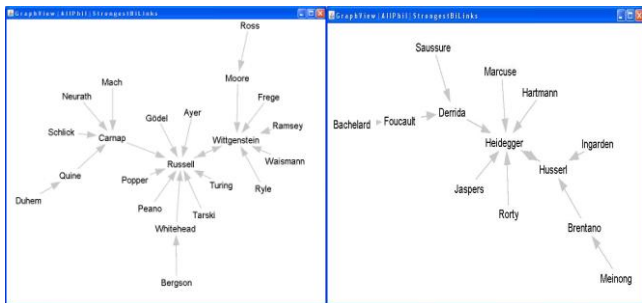

Figure 3. Hume cluster.
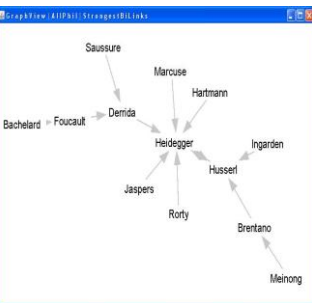

Figure 4. Russell cluster.

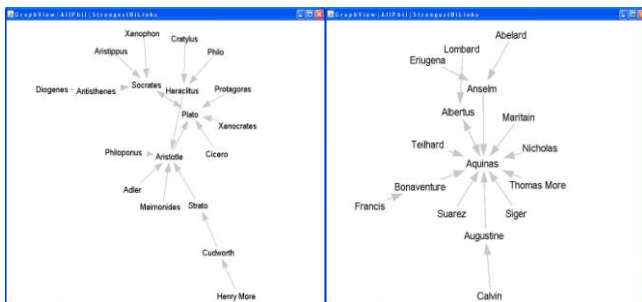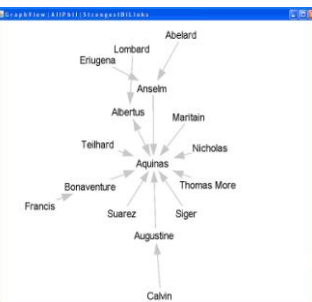**Figure 5. Heidegger cluster.**

Figure 6. Plato cluster.

Figure 7. Aquinas cluster.

3.3 Extended Hyperlink/Influence Networks

Figure 8 shows a graph representing extended out-links from Foucault, who is connected to only 11 philosophers directly but to a total of 271 philosophers via extended out-link connections. Figure 9 shows a radial graph representing extended influences from Thales, which shows that Thales, the first philosopher on the chronological list of 300 philosophers, can reach Foucault, the last one, within just 3 degrees of separation.
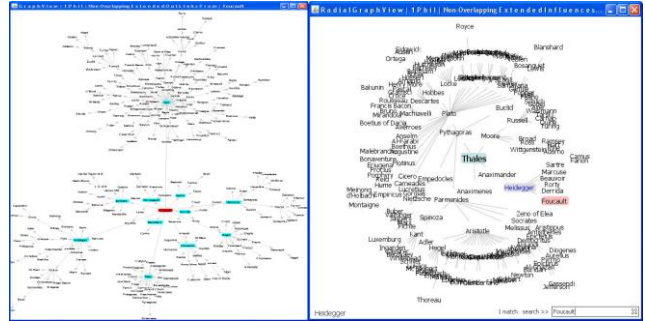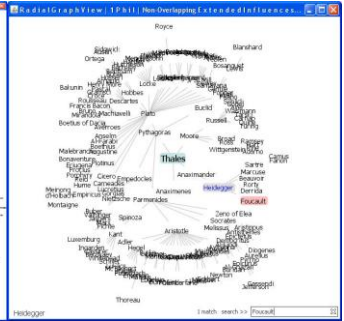

Figure 8. Foucault network.

Figure 9. Thales network.

4. SEMANTIC SEARCH INTERFACE

We created a prototype Web portal interface to enable semantics-based and visually-enhanced search and exploration of the facts, relations, and networks concerning philosophers. The Web portal is accessible at: http://research.cis.drexel.edu:8080/sofia/WPS/. Table 1 shows query options result display modalities.

Table 1. Options for semantics-based search and exploration

Foci	Facets	Visualization Modalities
One Phil	Academic/Biographical Facts	Radial Graph, Graph, Tree
	Direct Hyperlinks/Influences	Radial Graph, (Colored) Graph
	(All) Extended 3-Degree Hyperlinks/ 6-Degree Influences	Tree
Two Phil	Direct Relations	Radial Graph
	Commonalities	Radial Graph, Graph
	Direct (Common) Hyperlinks/Influences	Radial Graph, Graph
All Phil	Strongest Hyperlink/Influence Networks	Graph
	(Non-Overlapping) Extended Hyperlink/Influence Networks	Radial Graph, Graph
	(Purely Statistical) Rankings	Tag Cloud

5. CONCLUSION

In this paper we have demonstrated visualizing the connections among philosophers using the hyperlink/semantic data extracted from Wikipedia. The results have shown that we can obtain fairly meaningful representations of the intellectual connections among philosophers even from the simple hyperlink data and that we can further explore interesting connections by using the semantic data.

6. REFERENCES

[1] Athenikos, S.J., and Lin, X. The WikiPhil Portal: visualizing meaningful philosophical connections. In *Proceedings of the Chicago Colloquium on Digital Humanities and Computer Science Vol. 1, No. 1 (DHCS 2008)* (Univ. of Chicago, Chicago, IL, USA, 1-3 November 2008). 2009. Available at: http://jdhcs.uchicago.edu.

[2] Athenikos, S.J. Interactive visualization and exploration of information on philosophers (and artists, scholars, & scientists) in an e-learning portal for digital humanities. In *Proceedings of the 2009 Symposium on Interactive Visual Information Collections and Activity (IVICA 2009)* (Austin, TX, USA, 19 June 2009). 2009, 1-6. Available at: http://www.csdl.tamu.edu/~shipman/IVICA/2009proceedings.pdf.

[3] Athenikos, S.J., and Lin, X. WikiPhiloSofia: extraction and visualization of facts, relations, and networks concerning philosophers using Wikipedia. In *Conference Abstracts of the 2009 Digital Humanities Conference (DH 2009)* (Univ. of Maryland, College Park, MD, USA, 22-25 June 2009). 2009, 56-62. Available at: http://www.mith2.umd.edu/dh09/wp-content/uploads/dh09_conferencepreceedings_final.pdf.