Sophia: An Altmetric Analysis Prototype

John Lonican[1], Mohan Timilsina[1], Waqas Khawaja[1], Ioana Hulpus[1], Michael Taylor[2], Conor Hayes[1] and Brian Davis[1]

1. National University of Ireland, Galway 2. Elsevier, Oxford, UK

Centre for Data Analytics

ELSEVIER

Insight

Introduction

Sophia is the first version of a custom analysis prototype built in a collaboration between Insight Galway and Elsevier. The prototype aims to assist users in understanding research dissemination in sources other than traditional scientific publications, such as in social media, government documents and news sources.

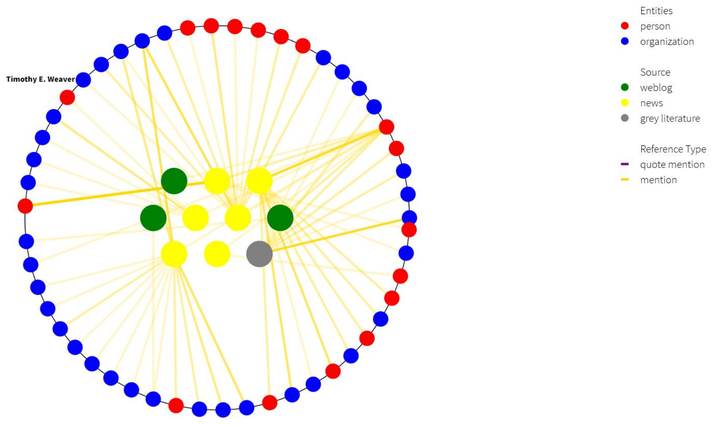

Fig. 2 – Graph representation of documents and mentioned entities

Fig. 3 – key-terms extracted from documents

- Clicking on an entity will fade the graph, showing only the entity, the documents which mention that entity, co-mentioned entities within those documents and the corresponding mention edges.

- Clicking on a term in the tag-cloud will remove all documents which do not contain that term, allowing co-mentions for a sub-graph of a given topic to be analyzed.

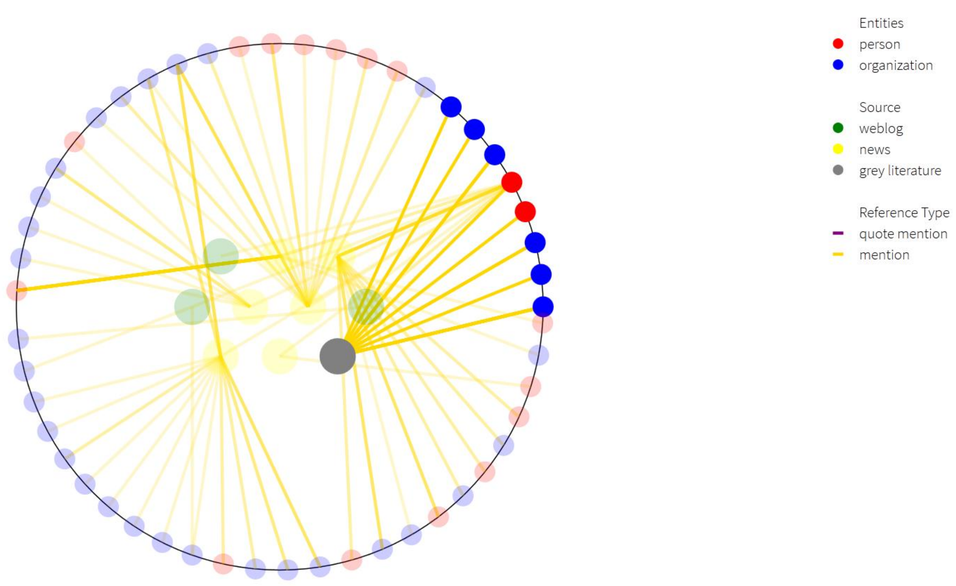

Fig 5 – entity co-mentions highlighted on faded graph

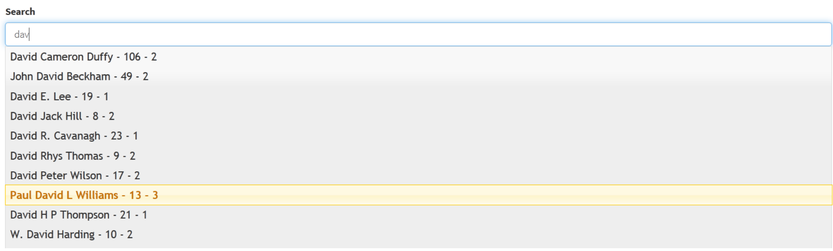

Fig. 1 – Author Suggestion

Features

- Ability to search scientific author by name based on Scopus id.

- The tag-cloud displays a number of key-terms extracted from the resulting document collection. These terms are scaled according to a modified tf-idf score calculated for each term, see Fig. 3.

- The graph visualization displays the mentioning documents, surrounded by a ring of mentioned entities. Edges between nodes on the graph represent a mention.

- Clicking on a document will display the content of that document in a pop-up box, with entities and mentioning statements highlighted for the convenience of the user, in the colour of the entity type.

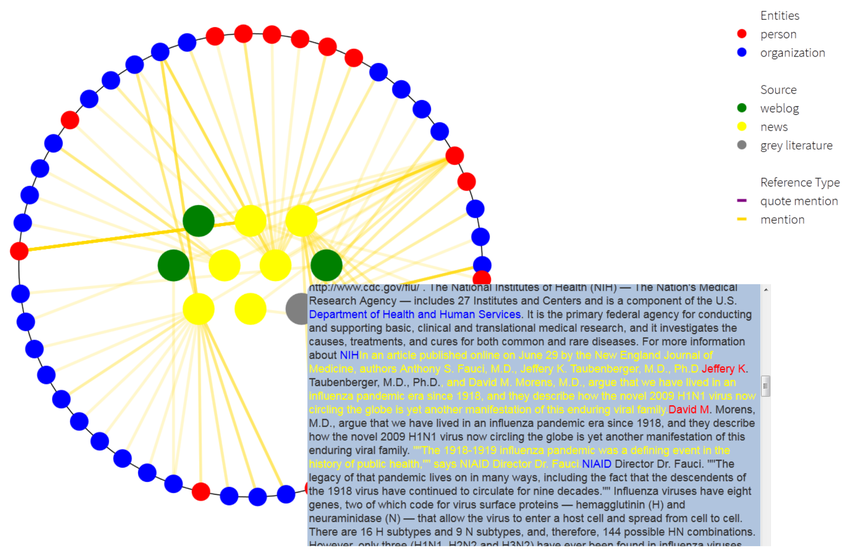

Fig. 4 – annotated document content, displaying highlighted entities

Future Work…

Future work on this application will include the following:

- improvements to visualizations
- improvements to text mining and annotation of the documents
- incorporating graph metrics
- displaying temporal information
- increasing input sources

This project has been funded by Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289 and by targeted project funding from Elsevier.

DCU NUI Galway OÉ Gaillimh UCC University College Cork, Ireland Coláiste na hOllscoile Corcaigh UCD DUBLIN SFI Science Foundation Ireland