



# VizCurator: a Visual Tool for Curating Open Data

Bahar Ghadiri Bashardoost\*, Christina Christodoulakis\*, Soheil Hassas Yeganeh\*, Okie Hassanzadeh^  
Renée J. Miller\*, Kelly Lyons\*

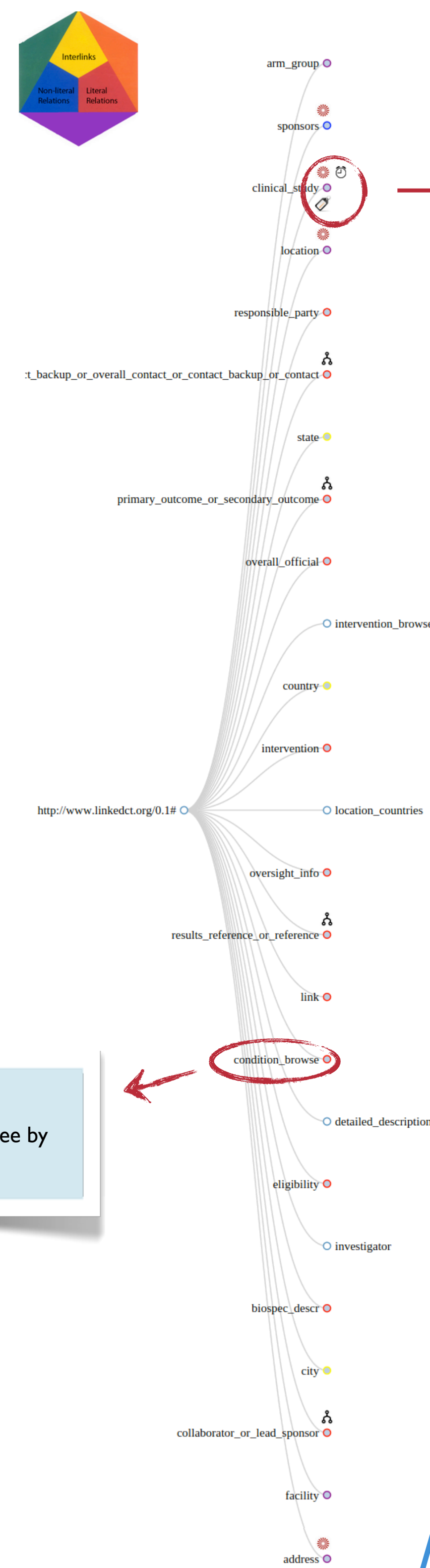
\*Department of Computer Science, University of Toronto, ^IBM Research  
ghadiri@cs.toronto.edu, christina@cs.toronto.edu, soheil@cs.toronto.edu, hassanzadeh@us.ibm.com, miller@cs.toronto.edu, klyons@cs.toronto.edu

## Introduction

VizCurator permits the exploration, understanding and curation of open RDF data, its schema, and how it has been linked to other sources. We provide visualizations that enable one to seamlessly navigate through RDFS and RDF layers and quickly understand the open data, how it has been mapped or linked, how it has been structured (and could be restructured), and how deeply it has been related to other open data sources. More importantly, VizCurator provides a rich set of tools for data curation. It suggests possible improvements to the structure of the data and enables curators to make informed decisions about enhancements to the exploration and exploitation of the data. Moreover, VizCurator facilitates the mining of temporal resources and the definition of temporal constraints through which the curator can identify conflicting facts. Finally, VizCurator can be used to create new binary temporal relations by reifying base facts and linking them to temporal resources. We will demonstrate VizCurator using LinkedCT.org, a five-star open data set mapped from the XML NIH clinical trials data (clinicaltrials.gov) that we have been maintaining and curating for several years.

## Discovery Via Visualization

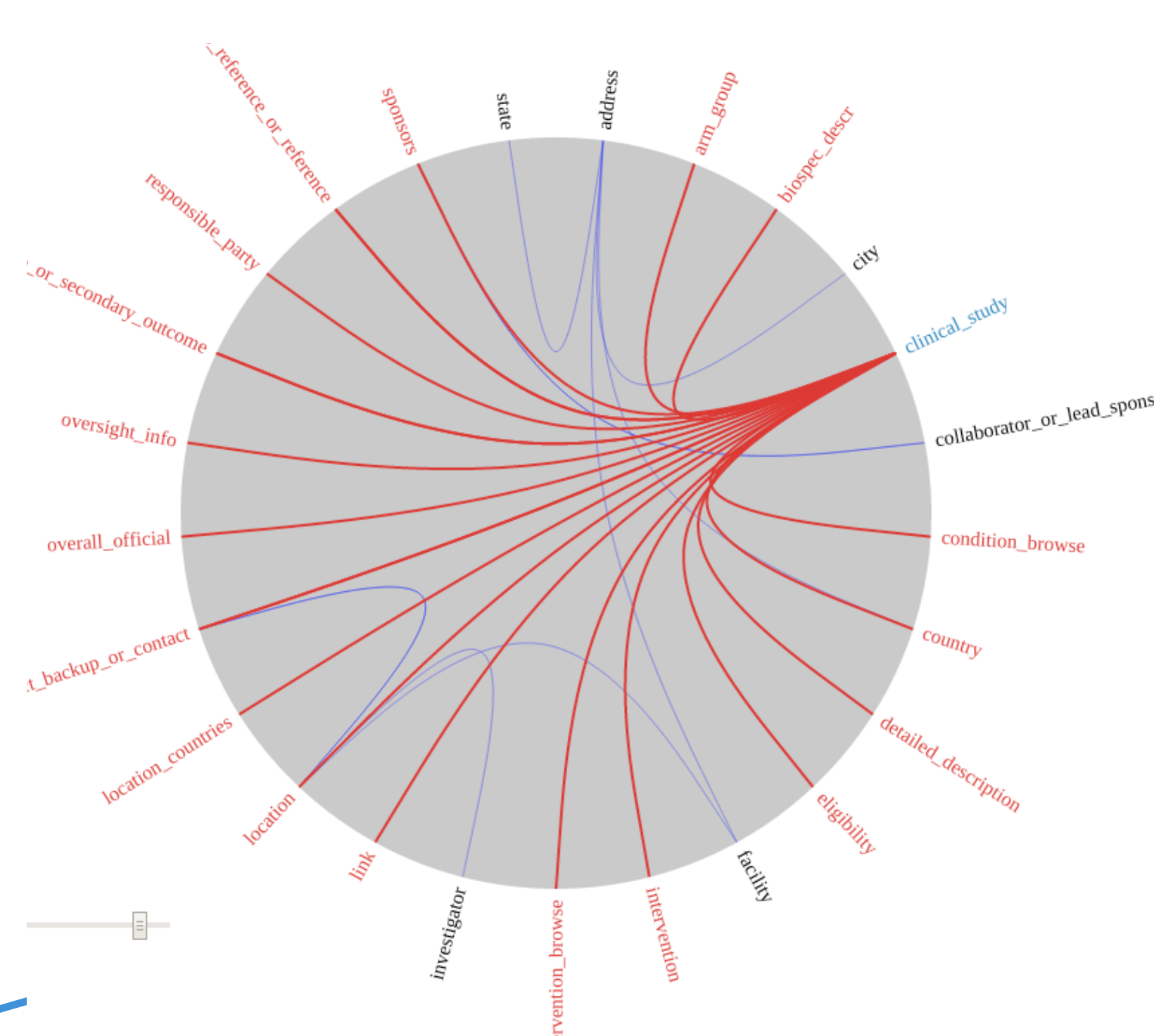
### Schema Tree View



**Seamless Connection to Data Browsing:**  
1. RDF layer can be explored while navigating the schema tree by clicking on the desired resource to open a LodLive[!] view.



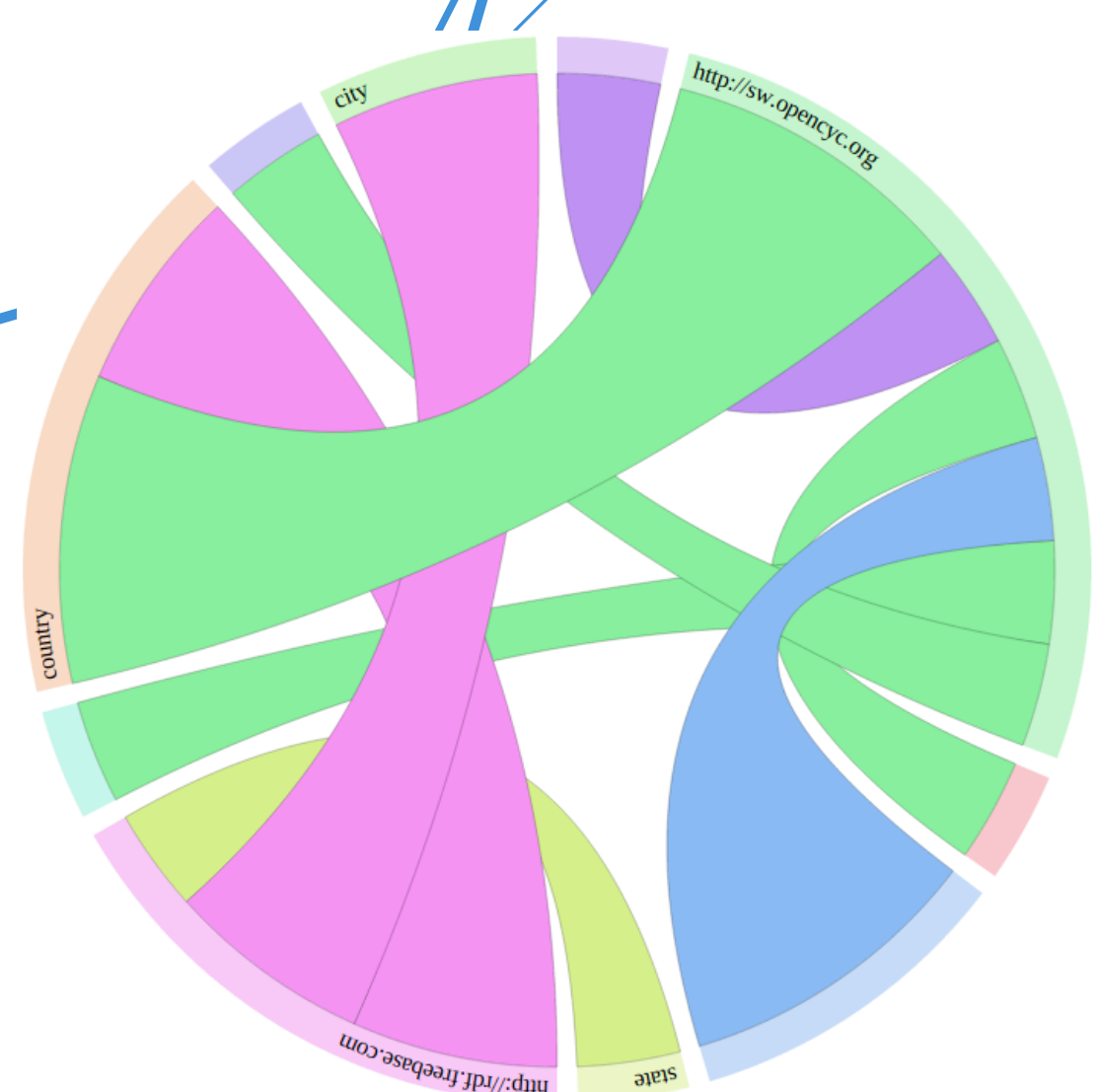
### Relation-mapping View



### Node-link Network

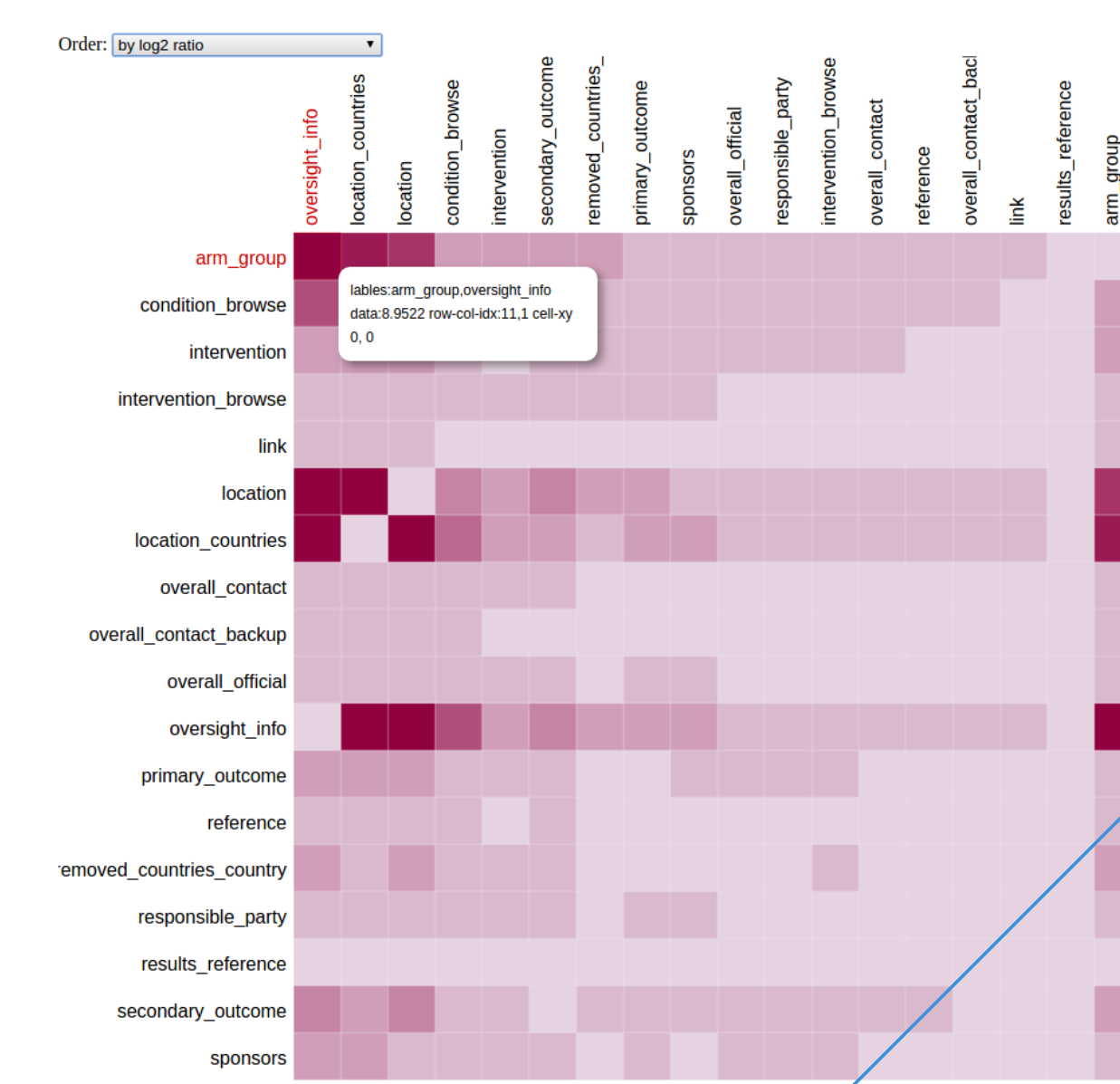


### Inter-linkage View

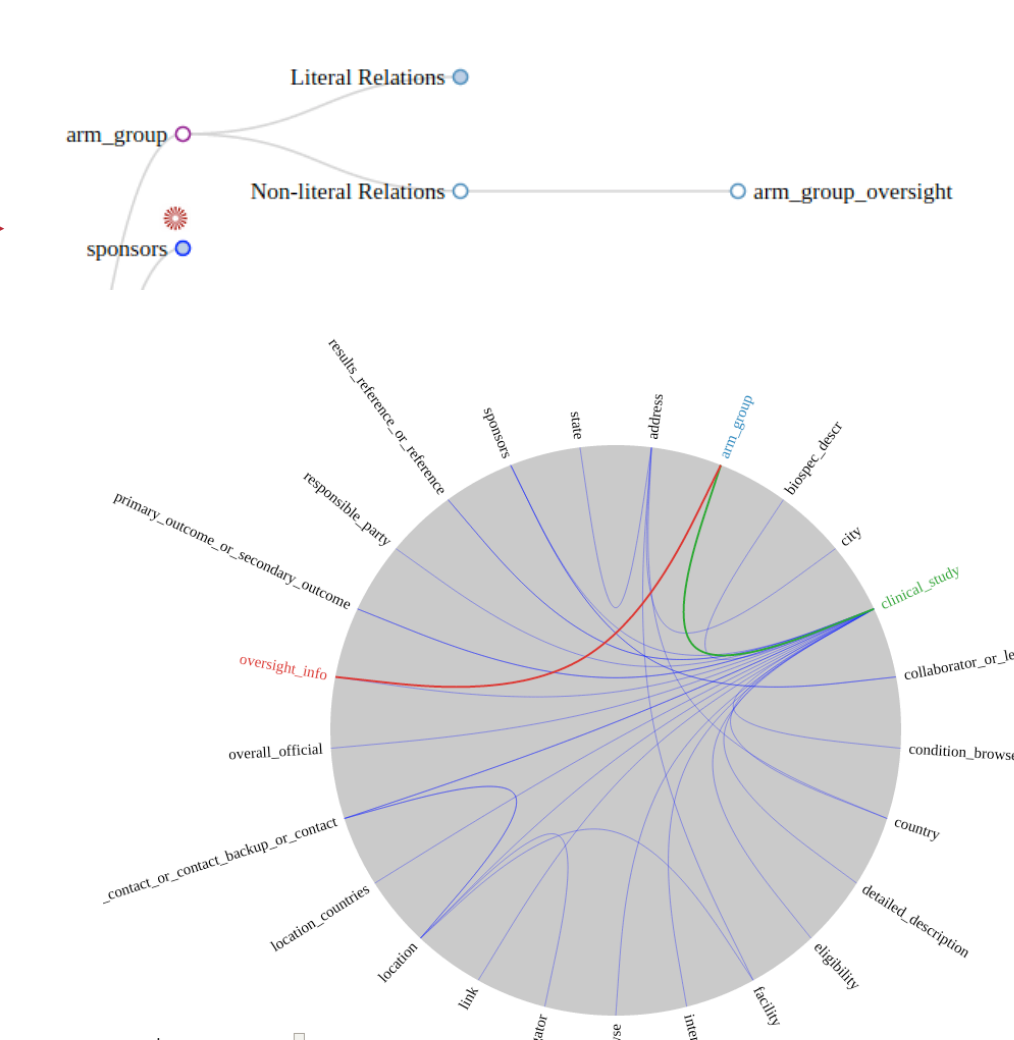


## Curation Via Visualization

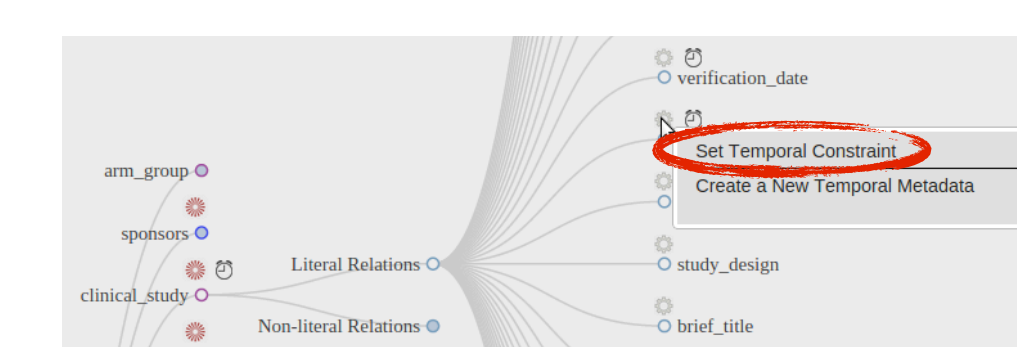
### Extracting Binary from N-ary Relations



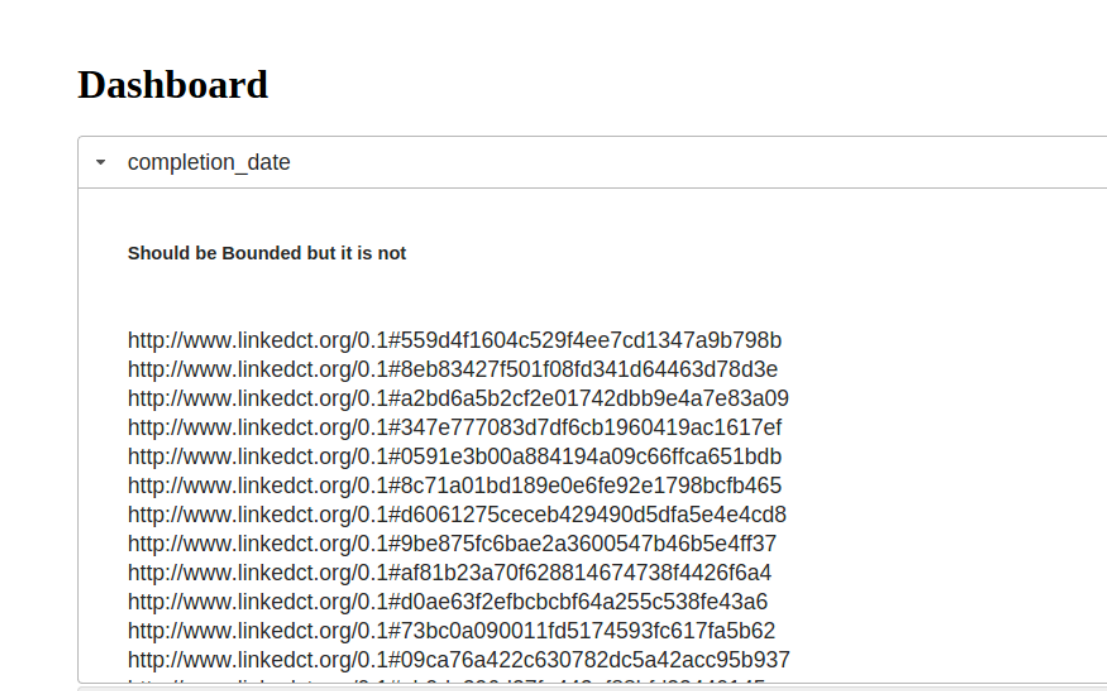
Set Name  
Please enter the name of the new relation  
arm\_group\_oversight  
arm\_group\_oversight\_oversight\_info  
Generate Mapping Cancel



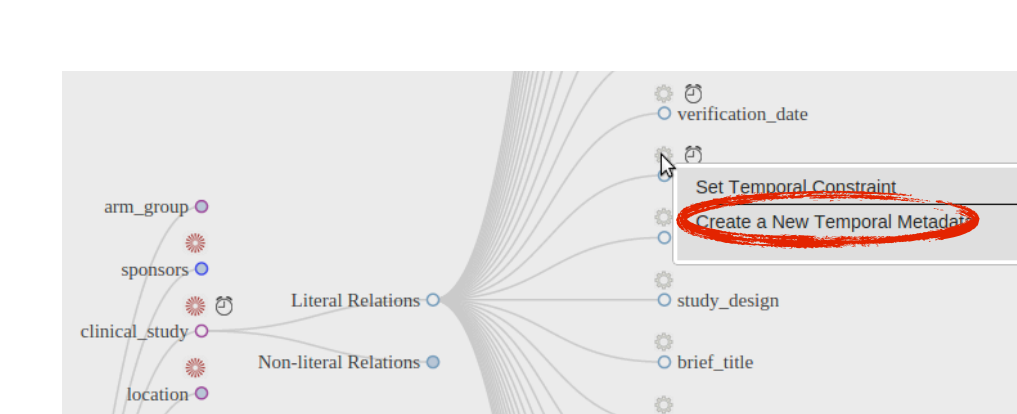
### Adding Temporal Constraints



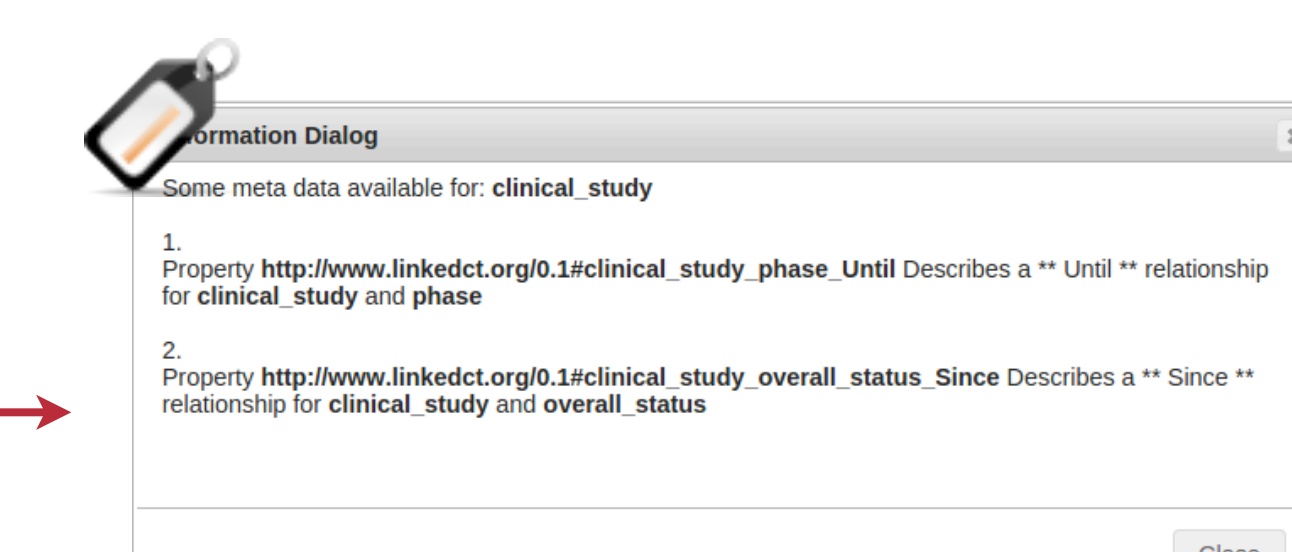
Set Temporal Constraint  
AND  
AND  
AND  
Unique within (uri: start\_date, uri: system\_completion\_date)  
Apply Cancel



### Adding Temporal Meta data



Adding meta data  
Default Value: June 2015  
Metadata Type: Since  
OnDate  
Unit  
Apply Cancel



**References:**  
I. D.V. Camarda, S. Mazzini, and A. Antonuccio. LodLive, Exploring the Web of Data. In I-SEMANTICS, pages 197–200, 2012.  
**Acknowledgments:**  
This research was partially funded by the GRAND NCE as part of the MEOV project, and by NSERC BIN.