



Figure 2: Heat map: darker cells are more populated

For each type in a KB, VizCurator selects types that use dates and potential temporal types. A potential temporal type is shown with a clock icon besides its name in the schema tree view. The curator can set a constraint by right clicking on a property of an entity type in the schema tree view. For example, for the `Completion_date` property of the `clinical_study` one can define a constraint: `UNIQUE AND AFTER http://www.linkedct.org/0.1#start_date`. The conflicting facts will be shown as warnings.

Sometimes a temporal relation is needed to better describe facts. For example, in our dataset each `clinical_study` has an `overall_status` that can be recruiting, active, etc. It is very useful to know approximately how long a `clinical_study` is in a specific state because a push for an update can be triggered if a fact stays in a state longer than a typical average time. VizCurator allows the curator to create new temporal relations and assign them to the facts by right clicking on a property of an entity type in the schema tree. A dialogue will open and ask for the name of the temporal property which can be one of `since`, `onDate`, `until`. The curator can set the object of that property to be either Null or a default value. VizCurator will then automatically create the necessary resources both in the RDFS and RDF layers. In order to do so, we adopted T-YAGO’s approach [11] where we first reify the base facts and give them a new URI, and then use that URI in the new relation.

4. DEMONSTRATION PLAN

We will invite the audience to explore LinkedCT while highlighting different aspects of VizCurator.

Exploring the LinkedCT KB: The audience will be able to explore and discover the LinkedCT KB both in RDFS and RDF levels. As they explore, they can view how complex the structure of a trial is and what parts are used sparsely and what parts are coherent and used in almost all trials. They can drill down to components, such as sponsoring agencies and see the heterogeneity in the way information about these agencies is structured. Or, they can have a bird’s-eye view of LinkedCT and inspect the linkage points between LinkedCT and other external KBs.

Identifying possible problems in the KB: To make things even more interesting, we will ask the audience to use VizCurator’s cues to identify curation actions that can improve the structure of LinkedCT. These actions include finding N-ary relations and picking out the best binary candidates to be extracted from those relations, finding temporal resources, and inspecting the external links.

Re-curating the KB: The audience can create new binary relations from the N-ary relations and inspect the changes using VizCurator. Also, in this step, the audience will define temporal rules and find the conflicting triples.

5. ACKNOWLEDGMENTS

This research was partially funded by the GRAND NCE as part of the MEOW project, and by NSERC BIN.

6. REFERENCES

- [1] H. Bast, F. Baurle, B. Buchhold, and E. Haußmann. Easy Access to the Freebase Dataset. In *WWW, WWW Companion '14*, pages 95–98. International World Wide Web Conferences Steering Committee, 2014.
- [2] C. Bizer, P. Boncz, M. L. Brodie, and O. Erling. The Meaningful Use of Big Data: Four Perspectives – Four Challenges. *SIGMOD Rec.*, 40(4):56–60, Jan. 2012.
- [3] M. Bostock, V. Ogievetsky, and J. Heer. D³ Data -Driven Documents. *IEEE Trans. Vis. Comput. Graph.*, 17(12):2301–2309, 2011.
- [4] P. Buneman, J. Cheney, W. C. Tan, and S. Vansummeren. Curated Databases. In *PODS, PODS '08*, pages 1–12. ACM, 2008.
- [5] D. V. Camarda, S. Mazzini, and A. Antonuccio. LodLive, Exploring the Web of Data. In *I-SEMANTICS*, pages 197–200, 2012.
- [6] A.-S. Dadzie and M. Rowe. Approaches to Visualising Linked Data: A Survey. *Semantic Web*, 2(2):89–124, 2011.
- [7] V. Geroimenko and C. Chen. *Visualizing the Semantic Web: XML-based Internet and Information Visualization*. Springer, 2006.
- [8] O. Hassanzadeh, A. Kementsietsidis, L. Lim, R. J. Miller, and M. Wang. LinkedCT: A Linked Data Space for Clinical Trials. *CoRR*, abs/0908.0567, 2009.
- [9] A. Jentzsch, R. Isele, and C. Bizer. Silk - Generating RDF Links while Publishing or Consuming Linked Data. In *ISWC'10*. Citeseer, 2010.
- [10] S. S. Sahoo, W. Halb, S. Hellmann, K. Idehen, T. Thibodeau Jr, S. Auer, J. Sequeda, and A. Ezzat. A Survey of Current Approaches for Mapping of Relational Databases to RDF. *W3C RDB2RDF Incubator Group Report*, 2009.
- [11] G. Weikum and M. Theobald. From Information to Knowledge: Harvesting Entities and Relationships From Web Sources. In *PODS, PODS '10*, pages 65–76. ACM, 2010.
- [12] S. H. Yeganeh, O. Hassanzadeh, and R. J. Miller. Linking Semistructured Data on the Web. In *ACM SIGMOD (WebDB Workshop 2011)*, 2011.