

Customizing "General SPARQL" for visualisation of in-house data in Cytoscape

Martijn Van Iersel, Joanna Parmley and Liz Reynolds

General Bioinformatics

Abstract. The General SPARQL app lets you navigate the semantic web using federated SPARQL queries. The app is pre-configured to visualise public datasets from the EBI RDF platform as well as Bio2RDF and others, such as Reactome, Uniprot, HGNC, NCBI Taxonomy, ChEMBL and more. SPARQL results are mapped to node and edge attributes in Cytoscape, and assigned visual properties such as colour, node shape and arrow style. The network can be navigated and expanded dynamically using context menu actions.

The app is fully extensible. Any aspect can be customized, such as the location of SPARQL endpoints, and the use of visual properties. It can be made to work with any public or private RDF resource.

When General Bioinformatics first publicly presented the General SPARQL app, it was received with enthusiasm by both network visualisation and semantic data communities. The app is intuitive to use and works out-of-the-box with public data. Naturally the potential to customise and make it work with in-house datasets is appealing to companies with non-public data.

To further demonstrate customisation possibilities, we created a use case with data that is not readily available on the semantic web. We applied this to a disease-centric use case covering genetic variants, drug interactions and pathway data. This data can be used to find supporting evidence in drug target discovery. In a real-world research setting, customisation of the app would most likely be applied in two stages. The first stage to set up RDF and SPARQL queries would require knowledge of scripting. The second stage consists of a dynamic visualisation of the data inside Cytoscape, and can be performed by scientists without programming knowledge. Thus, this method could be used by a bioinformatician to create and deploy dynamic visualisations for a group of scientists.