Abstract: The Semantic Web brings us a wealth of data in a powerful and flexible ontology-driven, graph-based format that allows us to create a web of data as opposed to a web of documents. Semantic web technology helps to break down existing data silos and allows information to be freely linked together to provide richer and more meaningful analytics. Nexus is an experimental 3D semantic web visualization system that (SPARQL-driven), in the same spirit, also tries to attempt to break down the barriers between the people by allowing them to work collaboratively in real-time by leveraging virtual world technology (OpenSimulator) and the latest in-browser 3D graphics technology (WebGL) and new HTML5 communications technologies. Nexus uses Resource Description Framework (RDF), the foundation of the Semantic Web, for both its data model and visualization model. By tightly coupling data and visualization together with Semantic Web technology, it allows Nexus to have data-controlled visualizations such as molecular visualization.

Nexus Design Methodology: Nexus uses a back-end server written in Java using Jena, Jetty, Virtuoso, and a custom RDF ontology that controls visuals and layout aspects (nucleus, nucy2x, nexshiny, nexnodes etc.) The flow of data from the back-end server to the front-end user interfaces is nothing more than a stream of semantic triples. Front-end code processes the nes triples for instructions on how information should be presented and displayed for the user(s). Since 3D models are entirely represented in RDF, they can be persisted and shared by different front-end locations and/or triple stores. Nexus is still under heavy code revision.

WebGL/HTML5 Client - WebGL is a standards-based API for accessing low-level graphics functions via JavaScript in the browser (i.e. high-speed 3D graphics in the browser with no plug-in) with support in current versions of Firefox, Chrome, and support in the developmental versions of Safari and Opera. WebSockets is a new HTML5 bi-directional communications protocol supported in Firefox, Chrome, Opera (including iPhone/iPad) that is being used to synchronize client positions (represented as avatars in OpenSimulator) in the shared data visualization.

Nexus Architecture

Figure above – DNA RDF data initially loaded into Nexus. Data element positions set randomly before force-directed layout is applied as seen on right.

Figure right – A 2D layout of the linked DNA RDF data for a fragment of the over-all DNA structure based around a single C4 atom.

“The Semantic RDF Triple”

Subject
Predicate
Object

The figure on the left shows a RDF representation of a small strand of DNA structural data with a layout determined by a force-directed algorithm. The figure on the right is the same visualization but after the crystallographic coordinates have been mapped onto the positional coordinates of the atom types with all non-visual components hidden.

Problems and Challenges
• Create a semantically-driven visualization ontology → Nexus Ontology (Nex)
• Leverage virtual-world technology → Developed OpenSimulator client
• Leverage new graphics technology for browser-based visualizations → Developed WebGL/HTML5 Client
• Synchronize client activity with other clients for shared real-time experience → Implementing WebSockets ala embedded Jetty.

Future Directions
• Add and experiment with additional layout algorithms in Nexus – tree-form, clustered, layered, spectral
• Support of WebRTC (Real-time communications protocol) to allow video teleconferencing within the browser
• Improve scalability of RDF visualizations through use of multi-core coding - the goal is 2,000,000 triples of information displayed
• Additional visualization elements like mesh support.
• Align Nexus visualization ontology with Collada RDF representation.
• Add WebSockets to OpenSimulator client
• Embedded faceted browser, RDF data editing capability, WebID authentication
• SWRL-driven visualizations (Semantic Web Rules Language)

Acknowledgements
• Dr. Janos Hajagos – RDF UMLS and PubMed 2011 UMLS alignment
• Tammy DiPrima – VIVO REACH implementation
• Dr. Moises Eisenberg, Director of Medical Informatics
• DNA Data – Rutgers Protein Databank
• Michel Dumontier’s Lab pdb2rdf program

Erich Bremer, Associate Director Medical Informatics
erich.bremer@stonybrook.edu
http://www.ebremer.com/nexus