Curran Kelleher
Previous Research Experience

I have been involved with a number of large software development and research projects. Here I will only discuss those from which I plan to draw ideas and designs for my proposed research. When I was first learning to program, I had created an interactive 3D graphing calculator. While working at the Institute for Visualization and Perception Research (IVPR) I was exposed to many systems and platforms for interactive visualization and analysis. During an internship at the New England Complex Systems Institute, I worked with a team to develop a collaborative simulation of global scale socioeconomic simulations. During a summer stay in Germany, I worked on a data publication infrastructure for use with visual analytics systems. From all these experiences, I have gained valuable perspective and insight which prepare me well for the pursuit of larger projects such as my proposed research.

My first significant software project was a personal one dedicated toward visually representing multivariate mathematical functions as interactive three dimensional surfaces. The purpose of this work was to see how interactive visualizations of 3D mathematical functions can serve as an aid for mathematics students in building intuition about the structure of surfaces. I worked independently on this project for a number of years in transition from high school to college. From this project I learned how to create modular interactive software, and how various visualization and interaction techniques can generally aid significantly in development of intuition and understanding.

During my first year at IVPR I worked with a team of graduate students on implementing the Ph. D. thesis concept of Howard Goodell - rich session histories in the Universal Visualization Platform (UVP). The UVP was the primary testbed of the IVPR group at the time for research and visualization tool development. It was comprised of a multiple coordinated view framework supporting standard visualizations (such as scatter plot, pie chart, parallel coordinates, heatmap, etc.) and interactions (such as brushed selection and probing). Our task was to implement a navigable session history graph within the system which included additional features such as text and voice annotations. The purpose of this work was to explore how these rich session histories can enhance the collaborative visual analysis process. From this work came my first publication, called “Collecting and Harnessing Rich Session Histories” [1].

The UVP had grown very large and unmanageable, and creators of visualization tools still needed to deal with a set of low level APIs. This made it difficult to create new visualization tools. I had an idea for how to redesign the UVP to make creation of visualization tools very easy while still supporting the primary features of the system: high performance, rich interaction, and session history. I also wanted to make it easy for third parties to create plug-in visualization tools which could be loaded from the Web at runtime, making the system much more extensible. I was given an entire summer to work independently (and with the help of an undergraduate student) on this idea, called JyVis. By the end of that summer, we had implemented the system and published a technical report [1].

The IVPR group was funded for several years to create a fully web based version of the UVP concept, called Weave (WEb based Analysis and Visualization Environment) by a nationwide consortium of non-profit planning agencies, governments and universities called the Open Indicators Consortium. When released, Weave will be freely available to the public (including source code) for not-for-profit use, while commercial users will be charged a licensing fee. After I returned from my junior year abroad in Germany, I became involved with the Weave project. I worked with a large team of software developers and social science researchers to implement Weave. Example tasks included the creation of visualization tools, development of algorithms for multi-scale polygon generalization [cite], design of the Weave server architecture, and the redesign of the Weave administration console (for importing and publishing tabular and geospatial data sets).

While working on Weave, we were confronted with a large data set whose structure didn't seem to fit into the tabular model supported by Weave. The data set was the Bureau of Labor Statistics (BLS) Employment Data Set, which consists of multiple intersecting hierarchies: time, geographic region, industry and ownership domain. This problem led me to investigate alternative data models, and I discovered the notion of a hierarchical data cube, which fit the BLS data perfectly. The concept of a data cube also fits with the goal of the Open Indicators Consortium for so called "nested indicators" - indicators across many geographic and temporal scales.

The Weave team also realized that in order to achieve runtime integration of data from multiple providers (one of the original goals), a metadata infrastructure was required. Weave needed standard names for measures and indicators along with their descriptions, standard record identifiers of various types, and specification of unit of measure for each indicator. In the UVP, JyVis, and Weave to date, these characteristics of the data are not represented.

I investigated the space of possible solutions for solving the two problems faced by the Weave team: support for hierarchical data cubes, and support for data integration across providers. In my reading I had learned that the Semantic Web paradigm had already addressed most of the issues of vocabulary definition and data integration. The only unsolved requirement was support for hierarchical data cubes.

My work led me to develop a blueprint for an extension of the Semantic Web to support semantically annotated interlinked data cubes, called the Universal Data Cube (UDC). Data published using the UDC framework would include all metadata required for integration hierarchical data cubes from multiple sources, and also would include human usable metadata such as unit of measure used, indicator description and provenance information.

During the summer of 2010 I was given the opportunity to work exclusively on the Universal Data Cube (UDC) during a stay as a guest researcher in Daniel Keim's visualization group at the University of Konstanz in Germany. I argued that the development of this system would transform and greatly enhance the visualization community if it were successfully implemented and achieved widespread adoption as part of a data dissemination standard. I worked on the Universal Data Cube for two months, giving presentations to the group as I progressed. By the end of the summer I had refined the specification of the UDC, fully implemented a prototype of the system, and written a research paper draft [1]. This work lays the foundation of my proposed future research.

[1] Additional material for consideration is available at http://www.curransoft.com/grfp/.