Capstone Research Plan Document
Human Computer Interaction
School of Informatics
2003

Name:
Sean Ellis

Advisor:
Dennis Groth

Working Title:
Annotating Dynamic Data Visualizations

Target Interest Group:
People who collaborate on data that is represented visually. This project’s focus is to develop a framework that could be adapted to many different fields. Our solution tackles the very general problem of the inefficiencies of current written collaboration methods. A good example of a professionals which might be interested in this framework are cardiologists, as demonstrated in the picture below of cardiologists gathered around a display, looking at images of a heart.

www.sbcardiology.com/

Introduction to the Problem:
With ever-increasing amounts of data being generated by the corporate and scientific communities, it is imperative that novel techniques for garnering knowledge from large collections of data be developed. This project proposes research of visualization applications, coupled with interactive techniques, deployed on mobile devices. The mobile devices will communicate with more powerful stationary devices
equipped with very large displays. Our hope is to create a new paradigm for the way people interact with visualization spaces.

We propose to explore the impact of using mobile devices to remotely connect to very large visualizations. Our solution will give users the ability to not only view data, but add annotations to the data in the visualization. This process will create an interactive environment, where users can collaborate via visualizations. Currently, there is no standard for connecting to a visualization, and allowing annotation in a shared way.

Information visualization is important because it helps us gain insight on information that is clustered into large displays. Understanding and extracting knowledge from large sets of data is important to any field, or profession, where large amounts of data are produced. It is important for users to be able to make observations and annotations to data, as observations and annotations add to our understanding of the underlying data.

An analogy to a discussion forum helps clearly explain the benefits of adding annotations to a set of data. Every discussion forum must have a topic, to motivate responses. A visualization, is itself a topical representation of data. In the visualization below, data is plotted on a 2D graph with Salary on the y-axis and Age on the x-axis. The topic of this visualization could be, “Is salary amount related to age?”

Topics is discussion forums are analogous to the visualization itself; threads in a discussion forum are analogous to annotations in a visualization.

Predispositions:

1. People must collaborate
2. People collaborate over data
3. Groups of data form data visualizations
4. When people collaborate, they share ideas
5. Annotating pictures is a useful tool when that picture is shared among people
6. Annotations lead to insights
7. Annotations lead to more annotations
8. Insights can be motivated from other insights
9. Different people view the same data with different perspectives
10. Annotating electronic ‘documents’ is non-trivial
11. Annotating electronic documents with handwriting is non-trivial
12. Tablet PC’s allow for electronic handwriting in it’s truest form
13. Drawing/marking are naturals way for people to annotate a picture
14. Drawing/marking is intuitive
15. Computers are efficient
16. Computers make processes efficient
17. Computers are evolving into writing machines (Tablet PC)
18. Large collections of data are informative
19. When something is interesting, people want to see it
20. Collaboration is effective when users can share insights
21. Collectively, different perspectives are interesting
22. Capturing perspectives via digital annotations is interesting
23. Annotations can be as of, or more interesting that the data itself
24. True collaboration can occur when each user has the ability to view all annotations made to the data
25. Wireless computing is booming
26. Different input and output methods are inevitable
27. Handwriting applications are convenient
   o If handwriting recognition is good

Research Plan – Literature Review:

I have created an annotated bibliography for this project as a means to organize the literature that I have considered. Over the course of this project, this annotated bibliography will continue to grow. It can be found at:

http://www.informatics.indiana.edu/dgroth/Research/Projects/Annotation/bib.htm
Research Plan – User Studies:

Observations of cardiologists, or chemists have been discussed. Nevertheless, I will observe a group of professionals who must collaborate as an integral part of their occupation.

Insights and Vision:

Concepts:

There are a number of possible concepts which could stem from our work. This application could be applied to many different fields; any field where people collaborate on visualizations might have an interest in such a product.

I am developing a prototype which belongs to a much larger framework, which addresses the general problem of collaborating via data in an electronic environment. In order to add value to this project, I will develop one or more specific uses for this prototype as the development process continues.

Prototype:

Visualization transforms the symbolic into the geometric, enabling researchers to observe their simulations and computations. Visualization offers a method for seeing the unseen. It enriches the process of scientific discovery and fosters profound and unexpected insights. In many fields it is already revolutionizing the way scientists do science. (McCormick et al., 1987)

Tablet PC

For this project we will be using Tablet PC’s as our mobile device of choice. We feel that Tablet PC’s are well suited for this task because they are mobile, have relatively large, high resolution, screens, and can support multiple forms of input:

- Handwriting recognition via a stylus
- Voice recording
- Typing/text input

While our goal is not to restrict users from using an input device other than a Tablet PC (e.g. PDA), development for devices other than a Tablet PC is not within the scope of this project.

.NET

We will be developing this project using the Microsoft .NET framework. .Net is an ideal solution because, “It (.NET) enables a high level of software integration through the use of XML Web services – small, discrete, building-block applications that connect to each other as well as the other, larger applications over the Internet. .NET is infused into the products that make up the Microsoft platform, providing the ability to quickly
and reliably build, host, deploy, and utilize connected solutions using XML Web services, all with the protection of industry-standard security.” (Microsoft.com)

**XML**

One of the goals of this project is to standardize the process by which visualizations make themselves available to outside users. In doing this, it will be possible for any person or group to share their visualization with multiple users. In order to create an open-source, non-proprietary standard, we will use XML.

“XML allows the flexible development of user-defined document types. It provides a robust, non-proprietary, persistent, and verifiable file format for the storage and transmission of text and data both on and off the Web.” (The XML FAQ <http://www.ucc.ie:8080/cocoon/xmlfaq>)

**Process**

Once a user finds interesting data points on a visualization, they can ‘highlight’ those points by touching them on the display. Users can then ‘tear off’ those points to a Tablet PC. ‘Tearing off’ is a term we use to represent copying data from one display to another. Once the data is ‘local’, the user can make annotations to the data. Users will then be able to put the data back to the original visualization. Other users will then be able to see the annotations as they interact with the visualization.

One of the strengths of this project is its flexibility, in that it is not specific to any set of data. Once there is a general framework built, it should be possible to apply this application to many different visualization spaces.

The prototype for this project is going to be an extensive program written with C#, and XML. In developing in C# we are taking full advantage of the Microsoft .NET framework, which gives us the ability to create a fully Internet enabled program which can be accessed from virtually any platform. Our prototype will allow user’s to remotely
connect to a data visualization, view a data set, make annotations to the data, and view other user’s annotations. Through this process, we hope that user’s will generate interesting insights based on the annotations of other user’s. Also, we hope to solve the problem of collocation, with respect to annotating data. With our system, users will remotely connect to a data visualization, housed on a server; as a result of the server being accessible via the Web, users will be able to connect wherever, whenever.