Data Sciences Summer Institute (DSSI)
Projects Summary
University of Illinois at Urbana-Champaign

Expert-Finder Project
2010 Data Sciences Summer Institute (DSSI)

We propose to develop an expert and expertise search system which tracks and retrieves researchers and professors based on their research expertise, visualizes their related information via an automatic generated webpages and displays their geographic locations via a Google map panel. Our system will have many useful implications for the academic, commercial and government fields. For example, academic funding agencies frequently search for researchers which are not only experts in their fields but also satisfy from geographical constraints. Commercial implications include, but are not limited to, an email agent that could prompt people of events when and *only* when the person is likely to be interested in that event. Government implications include military intelligence gathering, in that, a ranking intelligence officer may want to communicate with an expert on a topic who is nearby.

For a proof of concept we will limit our scope to experts in the realm of American computer science researchers.

A solution to this problem will likely contain the following pieces:

1. Data crawling and information extraction – crawl the Internet to find Computer Science researchers' web pages and extract their publication-text.
2. Data integration – match the web page information with structured information (e.g., DBLP)
3. Citation-Author-Topic Modeling (CAT) – determine the topics and cluster experts or professors based on the similarity of their expertise
4. Map Visualization – display the results on a map and provide functions to search, filter, cluster and constrain the results.

++

Crime Map Project
2010 Data Sciences Summer Institute (DSSI)

With crime accompanies difficult questions: Where is crime happening around me? Is the area less safe than before? Where will criminals strike next?

CrimeMap is a project that aims to study crime patterns from a variety of sources and provide real-time information about where crime is occurring and statistically where it is likely to occur in the future. CrimeMap uses unstructured police reports, news articles, and location-aware services (such as Twitter) to depict on a map where crime is occurring, who was nearby (and could be a potential witness), as well as look at trends of where crime is moving and where it is

likely to occur. CrimeMap provides “heatmap” visualizations over a map that allow users to see where hotter (or more frequent) areas of crime are occurring as well as when they are occurring.

Users, such as citizens and/or Police Departments, can study these trends online as well as take the information with them via a mobile iPhone/iPad application. For example, citizens could use the application to find the ‘safest-route’ home at night rather than just the shortest-route. Police officers could use the iPad application in the field to increase patrols in ‘hotter’ areas or become better prepared during a stop when there is a higher chance of the suspect using a weapon.

++

Computer Vision Project
2010 Data Sciences Summer Institute (DSSI)

The goal of the computer vision project is to compare numerous images and identify which ones are similar to each other.

How it works: We locate the interesting areas of a picture (interest points) and label them according to their type. If two pictures contain a similar amount of each type of point, we can say that they are similar. If they contain the exact same number of each type, they are probably identical.

We are looking at applying our image recognition software to compare the pictures in news reports and social networks (eg. Facebook).