IPRO 303 : Chujio

Project Plan Fall 2010

Faculty Advisor: James Burstein

Table of Contents

Team Information....3
Team Roster....3
Team Identity....7
Team Purpose....8
Team Vision....8
Team Goals....8
Team Objectives....9
Background Information....10
Company History....10
Company Challenges....10
Technology and the Solutions....10
Previous Work....11
Ethical Dilemmas in Research....11
Sociological Effects and Costs....11
Proposed Implementations....12
Previous Research....12
Team Values....13
Team Ethics....13
Conflict Resolution....13
Work Breakdown Structure....14
Problem solving process....14
Team Structure....14
Meeting Times....15
Gantt Chart....16
Expected Results....17
Expected Activities....17
Research and Testing Results....17
Potential Project Tasks' Outputs....18
The Deliverables....18
Project Challenges and Constraints....19
Results as a Solution....19
Project Budgets....19
Designation of Roles....20
Works Cited....21
Appendix....22

Team Information

Team Roster

Aditi Kumar
Year: Third Year
Major: Computer Science
Minor: Business
Description: Aditi's technical knowledge and skills in the area of computer programming are not only helpful in terms of software development but also provide a significant base when it comes to systematic planning, logical thinking, problem solving, and creative thinking. She aims to use those skills to add value and achieve the team's project goals for this semester.

Christopher Curtis
Year: Fifth Year
Major: Computer Science
Description: Christopher's particular interests include engineering software and mobile applications which enables him to give practical insights and creative solutions to the various issues faced by IPRO303. Currently working as a mobile software developer and with experience in web design Christopher plans to help design a fresh, innovative social product for the NAVTEQ group.

Dmitriy Vysotskiy
Year: Fourth Year
Major: Electrical and Computer Engineering
Description: Dmitriy has years of experience working in team environments and has previously collaborated on a social project. His skills include digital design, product management and team administration. Dmitriy also has an interest in the growing mobile market and social interaction. He hopes to harness his skills and interests to guide his team in providing NAVTEQ with an innovative new product.

Team Information

Ernesto Ramirez
Year: Fifth Year
Major: Computer Engineering and Computer Science
Description: Ernesto wishes to use his knowledge of programming to help develop software that will be user friendly. He is a part of an Illinois Institute of Technology and Dominican University joint program. This will broaden the demographics of the project and bring a different concept to the group. He also has experience with some of the methods that will play a major role to the project. He is also a returning member of the previous IPRO and will bring that knowledge with to make this IPRO a success.

Hashem Abu-Amara
Year: Fourth Year
Major: Electrical Engineering
Minor: Psychology
Description: Hashem's technical background adds to the team's ability of resolving problems and engineering solutions and also enabling him to grasp new concepts. Hashem's minor in psychology adds to the team a new way of thinking with a different perspective and through experience from areas like motivation and success and industrial psychology would help to bring the team together on working more efficiently and effectively. Through his experience in presentation, communication and people skills from previous projects he hopes to help the team achieve the goals and objectives of this semester.

Jason Chun
Year: Fourth Year
Major: Computer Engineering
Description: Jason's computer background and analytical skills allow him to see straight to the root of a problem. He will use these skills to help keep the team on track and quickly work through what may seem like difficult issues. His off the wall thinking will help to bring new and unique ideas to the table and his natural tendency for simplicity will help to stream line those ideas.

Team Information

John Jewell
Year: Fourth Year
Major: Architecture
Description: John's experience with hand drawing allows him to communicate with the group in a visual manner. As well, he is strong in art work and layout design and will be helping the team in the presentation boards and possibly schematic drawings.

Kalman Varga
Year: Fourth Year
Major: Architecture
Description: Kalman's background in architecture & design brings a different eye to this team. He will be contributing his design efforts in the overall appeal of his teams work. He is very personable, communicates well with others and wants to use these strengths to give presentations, organize team members, and contribute to public speaking of this team.

Mark Michael
Year: Fourth Year
Major: Computer Information Systems
Minor: Business Management and Information Security
Description: Mark has a strong background with technology, and learning social media development. Also this is his third time working with NAVTEQ and he helped developed the objectives of this semester. After the success of the previous semesters he is using the skills and knowledge he gained from those projects to provide the perfect solution to the project at hand.

Team Information

Melanie Koto
Year: Third Year
Major: Biology
Minor: Math and Science Education
Description: As a future teacher, Melanie has a great interest in serving people. That desire leads her to this IPRO group with the expectation that she will help propose new ideas and programs that anyone, even the "average Joe", may be able to utilize in travel. The development of using data from individuals' everyday information is also appealing. As a user of social networking sources, Melanie feels that she can provide helpful input in researching and gathering data for the project.

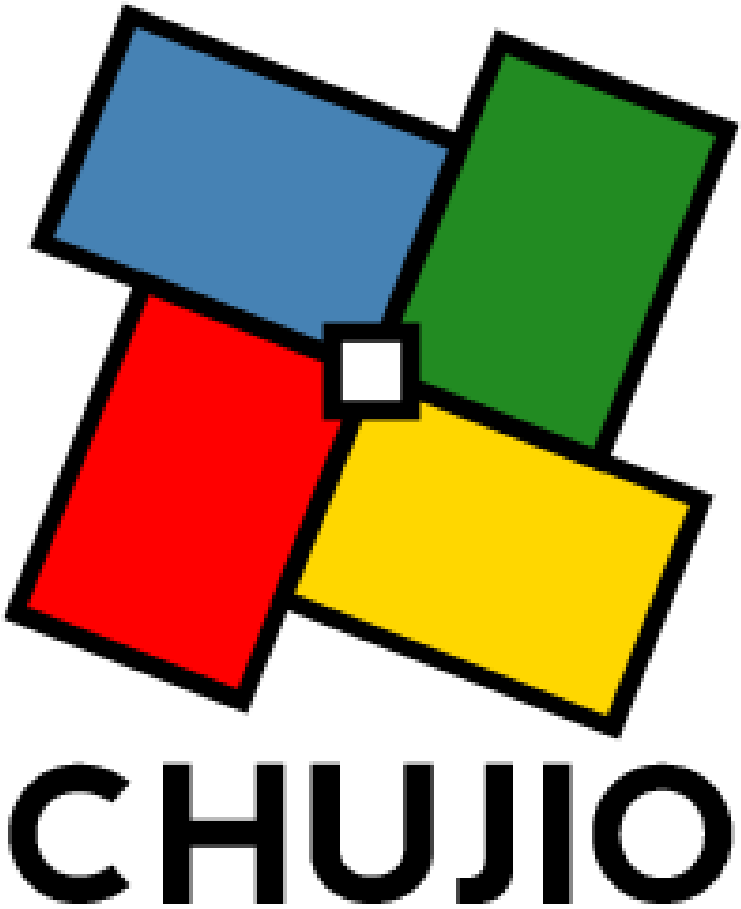
Patrick Tagny
Year: Fourth Year
Major: Computer Information Systems and Applied Mathematics
Description: Patrick is particularly interested in Software Engineering and ways to turn those products into revenues for companies. Regarding IPRO303, he is mostly excited about the combination of available social tools with consumer digital maps to get, in the end, a useful tool for consumer as well as a profitable product for the sponsor, NAVTEQ.

Rishi Kushaare
Year: Fourth Year
Major: Rishi's talents and strengths include mind Mapping, Intersecting area between technology, psychology, and business development; communication / keynote presentation, user interface optimization, and projection analysis. Also he has experience in business development, analysis, product research and development at HotSet Corp and Newell Rubbermaid (fortune 500) company.

Team Information

Team Identity

Name: Chujio
Name History: IPRO 303 previously known as Innovative Mapping has been assigned different goals for the semester and through those goals a new team name was selected to fit this current IPRO. Chujio is a Swahili word that means filter, this IPRO is based on filtering out any inadequate information from the social networking websites.
Motto: Solutions through Coal**iit**ion
Logo:



Team Purpose

Team Vision

To come together, synergize, and construct a practical, innovative, and simple solution to the problem at hand.

Team Goals

"Over the last few years there has been explosion of different websites that allow average person to contribute different content such as pictures, comments, stories and reviews about places around them. While there are many companies that focus on aggregating such data into common stream - complexity of interpreting such data in a meaningful actionable way remains a challenge for many companies including NAVTEQ." [Appendix B] Our goals for this semester are to create and demo a concept that accomplishes the following, using content sources such as Facebook, Twitter, and Flickr:

1) Identifies new points of interest from real time streaming content providers
   a) Twitter, Facebook, Flickr, and Fwix
   b) Two other social networking outlets to be determined by the group
2) Gather new and updated attribution regarding those places
3) Generates a time/space "mood-graph" describing each of these places in a relevant and contextual way

Team Objectives

- Aggregate and integrate meaningful data from real time streaming content providers and social networks (for example – Facebook, Twitter, Flickr, etc).
- Design a dynamic and easy-to-use algorithm that distinguishes the "noise" from the meaningful data while at the same time gathers updated information regarding the acquired data.
- Generate a "mood-graph" describing important details of the acquired data in a substantial manner.
- Build on the previous semester's developments.
- Work as a team: with honesty, integrity, and passion to develop the best solution to the project problem.

Background Information

Company History

NAVTEQ is a "world leader in premium-quality digital map data and content,"[1] and we are therefore very fortunate that NAVTEQ has decided to sponsor IPRO 303 this term. NAVTEQ supplies information used on in-vehicle navigation systems, mobile navigation systems, navigation programs used in cell phones, and internet based navigation websites. The data that NAVTEQ provides is extensive, covering 77 countries and spanning 6 continents. Formed in 1985 in Silicon Valley California, it is one of the founders of digital mapping data and map routing.[2] With the help of a number of investors and owners, NAVTEQ was able to grow and establish offices in countries outside the United States by the early 1990's. It eventually moved its headquarters to the current location in the heart of Chicago. NAVTEQ currently has approximately 4,000 employees spread over 190 offices in 39 different countries, and was recently acquired by Nokia, the popular phone and electronics maker.[3]

Company Challenges

NAVTEQ has presented this group with the challenge of finding, updating, and relating point of interest locations in relevant and contextual ways using automatic, live data streams from social networks. These points of interest locations are to be found though real time streaming content providers such as Facebook and Twitter. NAVTEQ plans to find new places and update existing places based on what people have to say about them. Then, given the kinds of ratings people provide, a mood graph is created to give other viewers an idea of what the place is like.

Technology and the Solutions

The technology will involve manipulating current open methods for gathering public data from a multitude of sources. This will include work with social networks current API as well as the development of an algorithm to filter and organize the data. The algorithm to be developed needs to be able to sift through large amounts of data quickly and efficiently. It needs to be able to identify new and recognize already recorded points of interest.

Background Information

Previous Work

This project is a continuation of the Fall 2009 and Spring 2010 semesters of IPRO 303 Innovative Mapping. Previously, the team worked to create a way for end users to send feedback to NAVTEQ in a fashion that was simple, accurate, and timely. The research conducted during the fall semester suggested that incorporating micro-landmarks with NAVTEQ databases would be better accomplished if the data came from parties with a vested interest, rather than coming from typical end users. Building off of that project, NAVTEQ now wants to automate the process of updating their current databases of Points of Interest (POI) using a live data stream from popular social sites. Since the groups focus has shifted from the end user to the backbone of the company, IPRO 303 has decided to change the name that has been used to better reflect the changes that have taken place.

Ethical Dilemmas in Research

The ethical issues that may result from investigating the problem include a potential invasion of privacy involving the observation of other people's conversations or posts. NAVTEQ may face confidential dilemmas on where they are data mining and how they go about obtaining it.

Sociological Effects and Costs

This will provide users with another way to find reviews and information about places they wish to go but have yet to see. It provides a more relaxed sense of interest as opposed to a firm review by a journalist or critic. From colloquial conversation people find the best information on a subject rather than a biased source. Thus, this project will help to bypass those who are paid to promote and those who are completely opposed and reach a more objective assessment from others who may better explain their experiences. Furthermore this project has the potential to change the public view of NAVTEQ, depending on how the solution is drawn out. It will be the responsibility of IPRO 303 to make sure that the data and or methods that are used and recommended to NAVTEQ will be ethical and will not harm the company’s image anyway.

Background Information

Proposed Solution

This semester NAVTEQ would like to see a working algorithm that filters raw data into meaningful information that can be automatically added to their database to update the current set of POIs in a fast and efficient method. IPRO 303 will work to provide NAVTEQ with an algorithm that filters a raw data into information that can be readily used. To do this the team plans on taking these steps:

1. Define a certain global format for raw data
2. Define the type of data that is useful
3. Aggregate the data in a meaningful way in a database

Previous Research

The previous research that was done with this IPRO is very helpful in identifying the main groups that use the websites we are focusing on, however the research does not provide the required knowledge for where this project is headed. According to our sponsor's goals and objectives we are no longer focusing directly on the end-users direct interaction with NAVTEQ and instead will be focusing on automatically updating NAVTEQ's database.

Team Values

Team Ethics

Chujio must follow all laws and regulations addressed by the Illinois Institute of Technology and national procedure. Each team member is held responsible for attending scheduled meetings and fulfilling scheduled requirements for all work that is assigned. A professional attitude is to be demonstrated at all times.

Conflict Resolution

The following are procedures provided to help address any problems:

- Each team member is required to participate in all team activities
- All members are required to communicate via iGroups and e-mail on a daily basis
- All work will be evenly distributed among the individual team members
- Each team member is responsible for submitting documented progress and any assigned work in a timely matter
- All team members will be respected and treated equally

Work Breakdown Structure

Problem solving process

The main purpose of our project is to help NAVTEQ's Location Content Management System (LCMS) team to investigate a new way to gather meaningful information about Points Of Interest (POI). Through already existing ways of communication like publicly on different social media websites, we will focus on the source and the mining of useful information. Some of the points we will try to solve this semester are:

1. Brainstorm internally in order to come up with meaningful sources of information about POI, not only in the US but possibly in the entire market covered by our sponsor NAVTEQ.
2. Identify different patterns socialites use in their posts to describe information that matters to us.
3. Seek support from researchers and professors at IIT's Information Retrieval Lab as NAVTEQ expects us to mine a very important chunk of raw data to make a marketable product out of it.
4. Write an algorithm to filter the any input data and extract the information we are looking for in a very efficient and valuable way.
5. Test our algorithm first on the quality of its results, then last but not least on the quantity it will be able to handle.

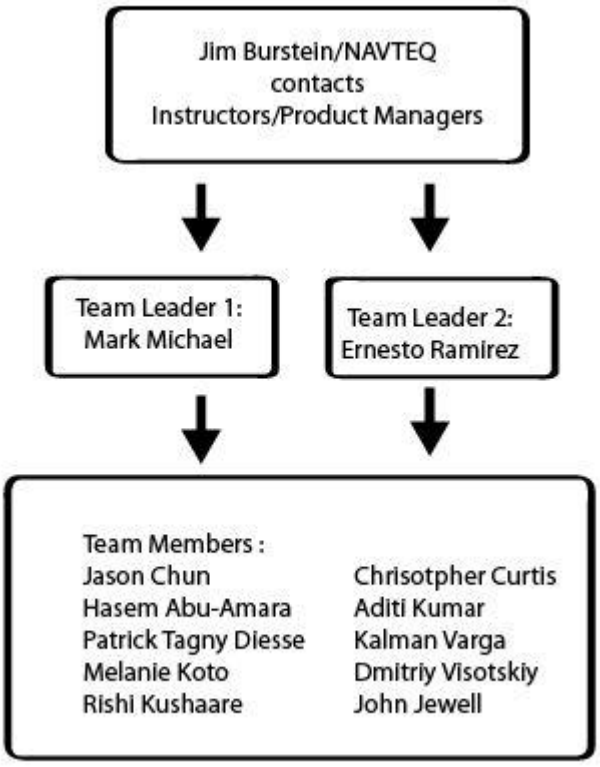
Team Structure

Despite the diversity of majors in our team, the fact that all members use social media and will be able to contribute greatly to our research led us to work as one team without any sub-team. This approach will, at the same time, get everybody involved in the project, and also guarantee perfect communication among members.

Taking a different approach from the previous terms, this year's team will not be split into subgroups at the beginning of the term but rather throughout the semester in order to complete the task at hand. Throughout the lifecycle of our project however, we will use every single team member based on his/her individual strengths to accomplish subtask. This evidently will make documentation of all work done the responsibility for us all.

Work Breakdown Structure

The established teams are organized as follows:

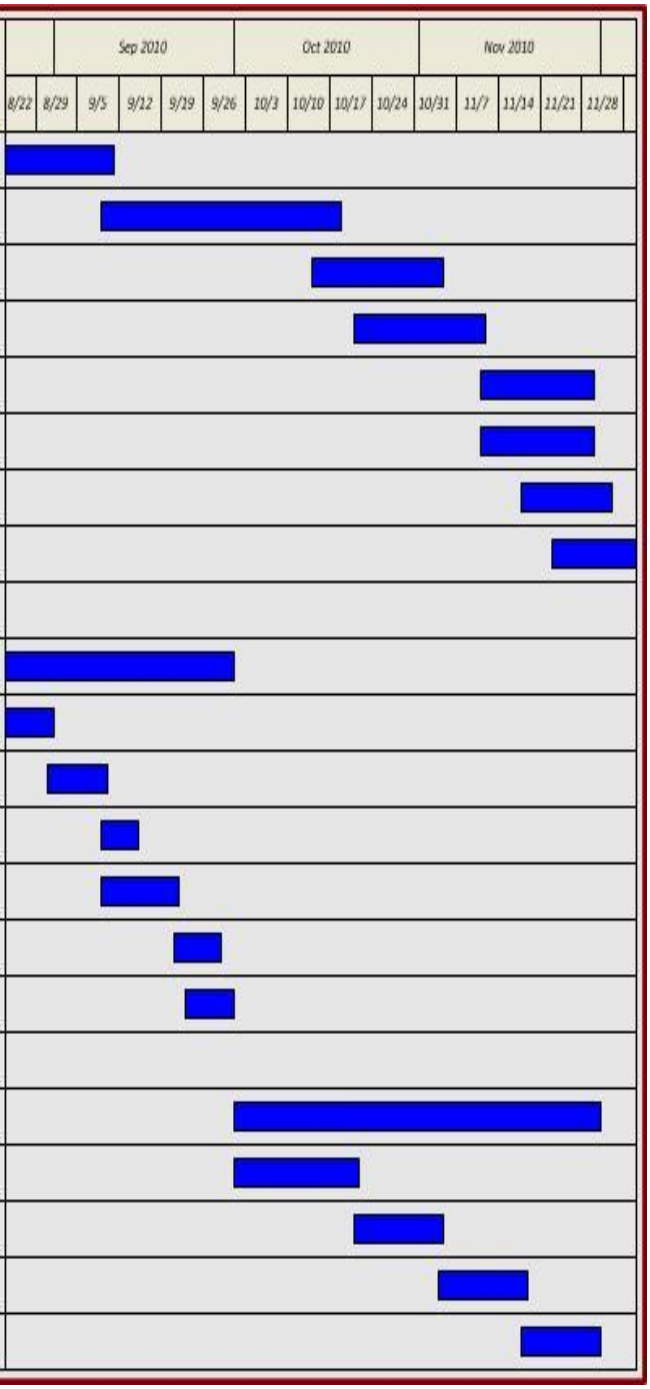


Meeting Times

The Chujio team is scheduled to meet every Tuesday and Thursday afternoon. Meetings are utilized as a forum for members of the Chujio project to present reports on recent developments, to address obstacles faced within the team, and to discuss what these mean for the entire group. Solutions are then brainstormed and the process is repeated. Additionally, time in class is spent reviewing and revising the set schedule of events to make sure the issues outlined are relevant and realistic based on current progress. The next page contains a basic breakdown of the work process structure and the dates that Chujio plans to complete each step.

Gantt Chart

ID	Task Name	Start	Finish
1	Project Plan	8/24/2010	9/10/2010
2	Midterm Review Presentation	9/9/2010	10/18/2010
3	Ethics Reflective Report	10/14/2010	11/4/2010
4	Final Project Report (Draft)	10/21/2010	11/11/2010
5	IPRO Abstracts/Brochure	11/11/2010	11/29/2010
6	Poster	11/11/2010	11/29/2010
7	Final Presentation	11/18/2010	12/2/2010
8	Final Project Report (Final)	11/23/2010	12/6/2010
9			
10	Phase 1	8/24/2010	9/30/2010
11	Study and Analyze Requirements	8/24/2010	8/31/2010
12	Conduct Market research	8/31/2010	9/9/2010
13	Discuss research/surveys results	9/9/2010	9/14/2010
14	Reach Consensus	9/9/2010	9/21/2010
15	Define Requirements	9/21/2010	9/28/2010
16	Propose Achievable solution	9/23/2010	9/30/2010
17			
18	Phase 2	10/1/2010	11/30/2010
19	Develop an algorithm	10/1/2010	10/21/2010
20	Test Algorithm with raw data	10/21/2010	11/4/2010
21	Apply Improvements to prototype	11/4/2010	11/18/2010
22	Deliver working solution	11/18/2010	11/30/2010



Expected Results

Expected Activities

The team aims to deliver at least one working solution to NAVTEQ. A group of activities are involved in this process. These activities fit into two chronological phases:

- Phase I (August-September) - Research phase. During this phase the students will:
    - Study the requirements that were given by NAVTEQ
    - Reach a consensus of specific milestones for this term
    - Define the requirements
    - Propose achievable and realistic solutions
- Phase II (October-November) - Development and Testing phase. During this part of the project the students will:
    - Develop an algorithm for submission
    - Test the algorithm with test set of random raw data
    - Apply improvements to the prototype(s) according to test results
    - Deliver a working solution during a formal meeting with NAVTEQ

Research and Testing Results

The main focus of this term will be on data mining social networks and social websites. Research will be conducted to help define what meaningful data is and how it can be sorted out from an irrelevant pool of information. The study of NAVTEQ's current system will help Chujio better define and format the data that will need to be provided to NAVTEQ.

The second phase of this project will include tests on the proposed algorithm and whether or not it is a plausible solution. The tests will also allow Chujio to optimize the algorithm to produce an efficient and powerful tool for gathering POI data.

Expected Results

Potential Project Tasks' Outputs

This semester's focus will be primarily on "meaningful data mining from social networks like Facebook and Twitter."[Appendix B] The main goals for Fall 2010 are to define ways to gather useful and relevant data from a large pool of information which can be accessed from several social networks. While we hope that our tests will prove that our algorithm is correct, in the chance that it does not, the team will then present the necessary steps to correct the algorithm.

The Deliverables

The team anticipates two deliverables to NAVTEQ:

- The first deliverable is a document defining the required format of parsed raw data. This should result from the first phase of the project and have the approval of NAVTEQ and the team.
  Deliverable 1 Target : First half of October

- The last deliverable of this project is a presentation of the team's research and proposed algorithm to NAVTEQ. The sponsor's decision to approve the algorithm after this presentation is the team's ultimate goal.
  Deliverable 2 Target : Towards end of November

Project Challenges and Constraints

A few challenges have been identified so far by the team members. A significant challenge that we will face is the time limitations of one semester's worth of work. This issue has been a reoccurring theme when a project of this magnitude is presented and IPRO 303 is excited to meet the challenge. The issue of getting meaningful data from the social networks that we will connect to our mapping system might also prove to be challenging. The key is to find a useful way to use the data that we gather from these social networks and location based services.

Technical standards and norms create additional challenges to this project. Due to the multitude of sources that the data may come from, each source will have the data formatted in a certain way. It will be part of our project to generalize and transform that data into one main format that then can be parsed and analyzed.

Results as a Solution

Chujio will deliver to NAVTEQ one working algorithm to filter and sort the raw data into useful, pertinent POI information. This algorithm will be generic enough to work with multiple variations of data formats and specific enough to translate that raw information into something that can be added to NAVTEQ's current databases.

Project Budget

The following is a list of materials and expenses that will be required to properly complete this project. Supplies will be required in the research and development phases of this project. A software development kit may be required in the design and creation of a prototype. A mobile device will be essential during the testing phase of this project.

BUDGET IPRO303	
FALL 2010	
Research	
Printing	$150
Development	
Final Presentation Supplies	$150
Survey Supplies	$100
Team Building*****	$200
Total	**$600**

*"Team Building" is any light competition to promote fellowship. Event includes food.

Designation of Roles

Minute Taker: Melanie Koto
Responsible for preparing and recording the decisions made in every meeting. Shortly after the meetings are held, in charge of uploading meeting minutes to iGroups for the team to use as a reference.

Team Leaders: Mark Michael & Ernesto A. Ramirez
Responsible for providing guidance, instruction, and direction to the team. Monitor the progress of the team to ensure results that need to be obtained are successful. Also, ensure that the project proposed by NAVTEQ is accomplished.

Agenda Maker: Mark Michael
Responsible for making an agenda for every team meeting. Makes sure that meetings go as planned and run smoothly. The agenda must be as such that deadlines are met on a timely basis and that the meetings are as productive as possible.

iGroups Moderator: Patrick Tagny
Responsible for ensuring that the IPRO deliverables are completed on time and properly uploaded in an organized manner to the iGroups website. Ensures that major milestones are added to the iGroups site and that necessary information is delivered to the team. Responsible for creating necessary folders for the appropriate files for better accessibility. Maintains that the allotted amount of space on iGroups does not get exceeded.

Works Cited

1. *NAVTEQ Corporate - About Us*. Web. 6 Sept. 2010. <http://corporate.navteq.com/index.html>.

2. "NAVTEQ Corporate - Global Leadership." *NAVTEQ Corporate - About Us*. Web. 10 Sept. 2009. <http://corporate.navteq.com/global.html>.

3. "NAVTEQ Corporate - History." *NAVTEQ Corporate - About Us*. Web. 6 Sept. 2010. <http://corporate.navteq.com/history.html>.

4. "IPRO 303 Innovative Mapping Project Plan Spring 2010 ." *IPRO iGroups*. IPRO, 02 Feb 2010. Web. 6 Sep 2010. <http://igroups.iit.edu/download.php?id=75557>.

Appendix

Appendix A:Team Roster

Name	Major	E-Mail	Phone #
Abu-Amara, Hashem	Electrical Engineering	habuama@iit.edu	[redacted]
Chun, Jason	Computer Engineering	jchun4@iit.edu	[redacted]
Curtis, Christopher	Computer Science	ccurtis7@iit.edu	[redacted]
Jewell, John	Architecture	jjewell1@iit.edu	[redacted]
Koto, Melanie	Math and Science Education	mkoto@iit.edu	[redacted]
Kumar, Aditi	Computer Science	akumar14@iit.edu	[redacted]
Kushaare, Rishi	Business Administration and Applied Science	rkushare@iit.edu	[redacted]
Michael, Mark	Computer Information Systems	mmichae3@iit.edu	[redacted]
Ramirez, Ernesto	Computer Engineering and Computer Science	eramire1@iit.edu	[redacted]
Tagny Diesse, Patrick	Computer Information Systems and Applied Mathematics	ptagnyd1@iit.edu	[redacted]
Varga, Kalman	Architecture	kvarga@iit.edu	[redacted]
Vysotskiy, Dmitriy	Electrical and Computer Engineering	dvysotsk@iit.edu	[redacted]

Non-Student Members

Name	Position	E-Mail
Burstein, Jim	Advisor	burstein@iit.edu
Ko, David	Sponsor Representative	david.ko@navteq.com
Ramic, Haris	Sponsor Representative	haris.ramic@navteq.com
Redzic, Gordan	Sponsor Representative	gordan.redzic@navteq.com
Holsinger, David	Sponsor Representative	holsinger.david@navteq.com

Appendix B: NAVTEQ Requirements

(see attached)

IIT IPRO 303 Project

Project Summary:

Over last few years there has been explosion of different websites that allow average person to contribute different content such as pictures, comments, stories and reviews about places around them. While there are many companies that focus on aggregating such data into common stream - complexity of interpreting such data in a meaningful actionable way remains a challenge for many companies including NAVTEQ. Hence, NAVTEQ is challenging bright students such as yourself to help it solve the problem of interpreting data in a meaningful way as described in project goals.

Project Goals

Create and demo a concept that accomplishes the following:

1. Identifies new points of interest from real time streaming content providers (e.g. Twitter, Facebook, etc...)
2. Gather new and updated attribution regarding those places
3. Generates a time/space "mood-graph" describing each of these places in a relevant and contextual way.

Content Sources to be used to

1. Facebook
2. Twitter
3. Fwix
4. ***Flickr***
5. ***Others to be determined by students (minimum of 2)***

Initial Dataset

NAVTEQ team is going to provide initial sample data in order to help students focus on interpreting the data instead of gathering the data. However algorithm should work on any dataset.

Potential skill set required

Information Retrieval, Data mining, Clustering, Understanding of Social Graph, Real time data processing