CSCC01 Fall 2014

Phase 1 Project Report

Team Chihuahuas (16)
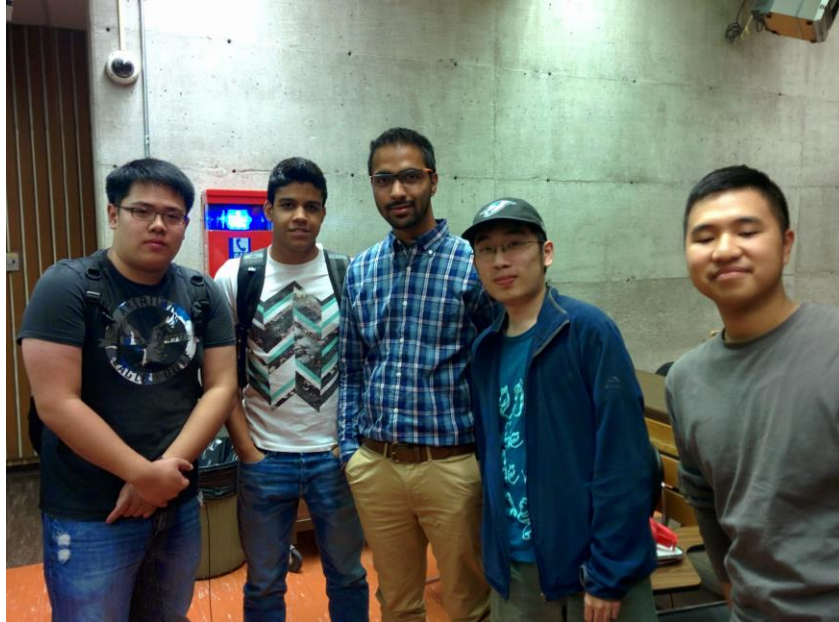
Table of Contents

Introduction

Team Name: *Team Chihuahuas*

Team Picture

Sharing Meal Picture: Sharing pizza!

Team Members

1. Kanwar Gill

Hello! My name is Kanwar Gill. I'm a fourth year computer science specialist (comprehensive) co-op student. I did six work terms in total; the majority of which were at IBM Canada. Coming back to UTSC, one of my main goals in addition to academics is to be involved in campus life. This year I'm the President of Co-op Students' Association and VP of Membership of Scarborough Toastmasters.

Having finished so many work terms, I have been introduced to both the agile as well as the traditional software development methodologies. I am quite comfortable in languages such as C and Python. I also have experience in scripting languages like PHP and Perl. I have learnt a fair amount of jQuery and additional plugins such as SlickGrid during my time at IBM. I'm looking forward to CSCC01 and I'm positive the methodologies that we will learn from this course will be long lasting.

2. Samson Ma

Samson Ma is a 3rd year student in the Specialist (Co-op) Computer Science program. His previous work terms were at the Ministry of Health and Long-Term Care and at Rouge Valley Health System, both as a Junior Developer. He was involved with a Java application project that assists health professionals in matching/consolidating student records for immunizations and a web application project that manages the professional medical staff for administration. At the University of Toronto Scarborough, Samson is in the Comprehensive Stream, and enjoys taking a variety of courses/topics, including databases, machine learning, embedded systems and network security. Outside of school, he enjoys swimming, reading and camping; his last camping trip was in Newfoundland for 10 days.

3. Carlos Ho

I am a student at University of Toronto Scarborough, in the computer science Comprehensive stream.

I started an interest in programming since before high school and have learned C, Java, and Python over the years of learning both on my own time and through academic courses. I've also picked up techniques in program efficiency, the software development lifecycle, databases, embedded systems, mobile development, and version control. Out of class, hobbies include coding personal projects to explore, badminton, and origami

4. Vinicius Barbosa

Vinicius is a Brazilian 20-year-old computer science undergraduate student. He's an adventurous guy and has an easy personality. He likes sports such as soccer, ice hockey and football.

Vinicius is at University of Toronto through a scholarship program named Science without Borders, sponsored by Brazil's government. During this year, he has completed the winter term at UTSC and a work term in a start-up called Wattpad, the world's largest community of readers and writers. He had a great experience to work as a QA developer, being able to work with tools as GitHub, Selenium, Appium and Xcode during the summer.

Vinicius has experience with programming languages as Python, Java, C++ and some knowledge of mobile programming in iOS.

As an international student, Vinicius is always seeking for new experiences and people during his stay in Canada. He's willing to help and provide any information about Brazil and Sao Paulo, his hometown.

5. Kenny Tsao

My name is Kenny Tsao. I am a computer science specialist in the software engineering stream. This is my third year at UTSC although officially I am a second year student as I currently am short half a credit from the third year threshold.

Growing up in the digital age I was always exposed to technology and naturally I developed a love for this technology and the programs that power them. In high school when I was selecting my grade 10 courses I'd come across a course named "computer science".

After reading the course description I'd decided to take this course as it seemed interesting and was related to something I had great interest in. In grade 11 and 12 when university applications were around the corner it did not take long to decide what program I would like to apply to in university. Here I am now, studying computer science at UTSC and I hope to, one day in the future, hold a career in this field for the majority of my working life.

Personas

Dr. Alejandro Paz - Researcher

Dr. Alejandro Paz, 30 years, is an Assistant Professor of Anthropology at the University of Toronto Scarborough. He is a calm and serene man, with a relaxed personality. He doesn't have a strict routine since he believes that life may be enjoyed along with work duties. Besides his teaching and researching duties, he usually plays soccer and baseball in his neighbourhood at Scarborough. Dr. Paz is single, but has one son from a previous marriage.

Dr. Paz has graduate affiliations in the departments of Anthropology and Linguistics, teaching courses in both departments. His researches include a variety of topics, including Language in Transnational and Global Contexts, Media and Publics, Translation and Commensuration, Pragmatics and Interaction and Israel in the Middle East. His current works are based on extensive ethnographic, archival and linguistic fieldwork in Israel and Palestine.

Since Dr. Paz doesn't have a strict routine as a professor and researcher, when he is not teaching he can be either in his research lab in the Anthropology departments where he manages a research team with a graduate student or at his house, where he has a huge office to work on his researches from home.

Dr. Paz has great computer skills, being able to easily answer emails and write notes through his iPhone, and use a variety of computer software packages to collect, analyse and organize information and data in his MacBook, which he then present to other members of his department, either in a written report or oral presentation.

One of Dr. Paz's main needs now is to be able to stay updated in the world's ethnographic and transnational news and conflicts. Not only to be aware of what is happening but also to have reliable sources so that he can work through some of them in order to develop his researches. He needs software that connects news sources in such a way that both the connection between the news and the sources are strictly reliable, coming from well-known news agencies and serious journalists. According to Dr. Paz, that would increase the productivity of his researches and be a valuable tool for his department in the future.

Alex Ferguson - Graduate Student

Alex Ferguson, 25 years, is a Graduate Student of the Anthropology Department at the University of Toronto Scarborough. He has a serious personality and is very hard working and energetic. He always wants to be on schedule and get things done, no matter how hard he has to work on them. Mr. Ferguson is engaged and expecting to get married within the next months. One of his main hobbies is fishing, which he believes is a good way to free his mind from his intense work scheduling.

Mr. Ferguson got a BSc in Anthropology at University of British Columbia. He was awarded as Alumni student for being outstanding scholarly during his undergraduate years. He specialized in Socio-Cultural Anthropology, emphasizing his interests in Middle East region.

Currently working with Dr. Alejandro Paz at the Anthropology Department at UTSC, Alex's main responsibilities involve defining research parameters and expanding the knowledge of that research field by discovering a unique approach to different global conflicts/news to help Dr. Paz in his projects. That requires a lot of time looking for trustable sources and reliable connections between them to make resourceful articles.

With good skills in computer usage, being able to use applications to manage his researches such as word processors and web browsers, Alex's work would be facilitated by software that could make those connections between news sources and help him build solid articles for Dr. Paz to work in.

Andrew Stanley - Librarian

Andrew is a 33-year-old librarian. He is a very systematic person and can seem aggressive to any person or object not following his methods of organization. Though uptight about procedures, he does not mind not having a set routine of tasks, as long as everything he does is neat, efficient, and can be documented up to his standards. Andrew is single, and whenever he isn't working at the university library, he is usually keeping up with the news and researching new occurrences brought up in them.

Andrew works in a library at a university, ensuring everything is in its proper space, which is very important in his setting. All materials must be in its proper section to better cater to the students, and he must also know this system completely, in order to answer questions from them quickly and accurately.

While the majority of the books in his current workplace are in English and of purely academic nature, Andrew previously worked in a public library, where he was exposed to many books of a variety of backgrounds and languages while he was returning them to their proper shelves, and read any he found interesting. This majorly contributed to the many languages Andrew can communicate in, some very fluently. It also contributed to the many cultures and activities in the world, sparking his interest to how they form or link together.

The library systems are mainly on a computer, which Andrew has learned to be very skilled with. He is always at one of the administrative library computers during work unless he is returning books, even staying at his post while eating lunch. He has also taken to using his computer skills off work to read the news and research on the web, which he finds fascinating at the volume of information on it.

User Stories

Registration
1. As a user, I want to be able to create an account, so that I can access the program.
2. As an admin, I want different permissions levels for the users, so that they only have access to the functions they require for the execution of their duties.

Login/Logout
3. As a user, I want to be able to log in and out of my account, so that I can access the program.

Adding/Defining news sources
4. As a researcher, I want to be able to add sources to the program, so that I can accommodate the expanding needs of my research.

Modifying news sources
5. As a researcher, I want to be able to modify the sources monitored by the program, so that I can add/remove sources as the focus of my research changes.

Web Crawl
6. As a researcher, I want to be able to find relationships between news source, so that I can better quantify my research
7. As a researcher, I want to find all the links to other news sources for a particular topic, so that I can better understand the relationships.

Exporting results
8. As a librarian, I want to be able to export the data to storage, so that it can be preserved for future reference.

Displaying information
9. As a researcher, I want to display the result of the web crawl in a table, so that I obtain an overall status/general report.
10. As a researcher, I want to filter the result of the table, so that I can be more efficient in my analysis/research.
   - filter by news source
   - filter by author
   - filter by keywords

11. As a researcher, I want to search the result of the table, so that I can be more efficient in my analysis/research.
   - search by article name
   - search by keywords

12. As a researcher, I want to sort the result of the table, so that I can be more efficient in my analysis/research.
    - sort by number of times it references articles from monitored sources
    - sort by number of times it is being referenced
    - sort by date article created

13. As a researcher, I want to view each article being referenced over time in a line graph, so that I am more efficient in my analysis of the data.

14. As a researcher, I want to display how sources are being referenced in a directed graph, where the nodes are the articles, and the edge representing an article referencing another article, so that I can analyze the data and understand how sources are being referenced.

Saving Information
15. As a researcher, I want to save the results of the web crawl, so that I can access the data at a later date and provide more data for the librarian to extract.
    - news source
    - article name
    - number of times it references articles
    - number of times it is being referenced
    - date article created
    - author
    - any relevant keywords

System Design

System Metaphor
The best way to describe the application is a journalist.
- The journalist will be assigned a topic and possibly people/sources to look for information.
- For the facts to be reliable they need to come directly from the source. The journalist will continue to search for other people until the original source is found.
- Once all the facts are gathered, the journalist will create a report.
- The report will be published and also archived for future use.

Release Plan

Planning Game
See video in the GitHub repository.

Release Plan

As a researcher, I want to be able to add sources to the program, so that I can accommodate the expanding needs of my research.		
Priority: High, Must	Estimate: 2 Developer Days	Release: First

As a researcher, I want to be able to modify the sources monitored by the program, so that I can add/remove sources as the focus of my research changes.		
Priority: High, Must	Estimate: 2 Developer Days	Release: First

As a user, I want to be able to create an account, so that I can access the program.		
Priority: High, Must	Estimate: 1 Developer Day	Release: First

As a user, I want to be able to log in and out of my account, so that I can access the program.		
Priority: High, Must	Estimate: 1 Developer Day	Release: First

As a researcher, I want to save the results of the web crawl, so that I can access the data at a later date and provide more data for the librarian to extract. - news source - article name - number of times it references articles - number of times it is being referenced - date article created - author - any relevant keywords		
Priority: Medium, Must	Estimate: 2 Developer Days	Release: Second

11

As a researcher, I want to find all the links to other news sources for a particular topic, so that I can better understand the relationships.		
Priority: Low, Could	Estimate: 3 Developer Days	Release: Second

As a researcher, I want to display the result of the web crawl in a table, so that I obtain an overall status/general report.		
Priority: High, Should	Estimate: 4 Developer Days	Release: Second

As an admin, I want different permissions levels for the users, so that they only have access to the functions they require for the execution of their duties.		
Priority: High, Must	Estimate: 1 Developer Day	Release: Second

As a researcher, I want to sort the result of the table, so that I can be more efficient in my analysis/research.
 	
   - sort by number of times it references articles from monitored sources
   - sort by number of times it is being referenced
   - sort by date article created		

Priority: Medium, Must	Estimate: 2 Developer Days	Release: Second

As a researcher, I want to be able to find relationships between news source, so that I can better quantify my research		
Priority: Low, Could	Estimate: 6 Developer Days	Release: Second

As a researcher, I want to view each article being referenced over time in a line graph, so that I am more efficient in my analysis of the data.		
Priority: Medium, Could	Estimate: 3 Developer Days	Release: Third

As a researcher, I want to search the result of the table, so that I can be more efficient in my analysis/research.
 	
   - search by article name
   - search by keywords		

Priority: Medium, Must	Estimate: 3 Developer Days	Release: Third

As a researcher, I want to display how sources are being referenced in a directed graph, where the nodes are the articles, and the edge representing an article referencing another article, so that I can analyze the data and understand how sources are being referenced.		
Priority: Medium, Must	Estimate: 5 Developer Days	Release: Third

As a librarian, I want to be able to export the data to storage, so that it can be preserved for future reference.		
Priority: High, Should	Estimate: 3 Developer Days	Release: Third

As a researcher, I want to filter the result of the table, so that I can be more efficient in my analysis/research. - filter by news source - filter by author - filter by keywords		
Priority: High, Must	Estimate: 3 Developer Days	Release: Third