EFFECT OF VISUALIZATION OF NEWS ARTICLES IN DATA DRIVEN GAMES

A Thesis

by

AKSHAY KULKARNI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Frank Shipman
Committee Members,	Richard Furuta
	Ergun Akleman
Head of Department,	Dilma Da Silva

May 2016

Major Subject: Computer Science

ABSTRACT

The popularity of prediction games such as fantasy sports has been on the rise and the amount of data available for the players to make the prediction in such games is growing rapidly. Prediction games, an area of data driven games require the user to interpret archival data along with real-time data from a domain to make a prediction about a future event. This work is being done in the context of a prediction game where players select geographic locations as part the game. News articles can serve as a source of information and have the potential to improve engagement and learning. To make sense of the millions of news artifacts published online is not possible manually. Moreover, keyword-based search is very limited when it comes to exploring data rather than just searching for something particular. The proposed work will develop a visualization and user interface to represent the news articles in an activity-appropriate manner and allow the users to explore news data using approximate search along with keyword-based search. The first component of this system extracts news articles related to the game, then geotags and clusters them based on the geographic references in the articles. Displaying these clusters on a map takes advantage of the spatially referenced news data. This thesis will compare alternative visualizations of the geo-tagged and clustered news articles and their value to players of the game.

A user study was conducted to evaluate the visualizations and their effect on the engagement of the players with a data driven game. The results show that the map visualization is very effective in engaging players with the game when compared to the

regular list form of news representation. Moreover, the overall performance of the players who used the map visualization was better than the performance of the users who used the list visualization. Future work will explore more on fine-tuning data sources which provide the input to the map visualization as well as variations in the display and accessible features on the map interface to enable users to control data visualization according to their imagination and preference.

DEDICATION

To my late grandfather who was my greatest mentor and to my parents, grandmother and brother for all their love and support.

ACKNOWLEDGEMENTS

I would like to take this opportunity to show my appreciation to everyone who has supported, supervised and provided assistance towards the development of the Fantasy Climate application. I would like to express my deepest gratitude and appreciation to my advisor and mentor, Dr. Frank Shipman. This thesis would not have been possible without his continuous guidance and support.

I would also like to thank my thesis committee members, Dr. Richard Furuta and Dr. Ergun Akleman and also to Dr. James Caverlee for helping me with my schedule.

A special mention to Gabriel Dzodom, for his constant help and feedback with the development of the system and the visualizations. Finally, I would like to thank all my fellow lab mates Meghanath Reddy, Sampath Jayarathna, Caio Duarte Monteiro, Christy Maria, Joshua Sheehy, Saketh Ram and Malavika Madhukumar for their insights and encouragement.

NOMENCLATURE

FC	Fantasy Climate
NER	Named Entity Recognize
NLP	Natural Language Processing
SD	Standard Deviation

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Kenneth L. Simons claimed that "If video games can be transformed so that their users learn, a great many people may come to understand and control dynamic systems" (Simons 1993, Garris, Ahlers et al. 2002). Since then there has been substantial progress in identifying which characteristics of games engage the user from an instructional and educational perspective (Garris, Ahlers et al. 2002). Additionally, a number of empirical studies have indicated that instruction incorporating game features leads to improved learning and understanding of complex topics (Whitehill and McDonald 1993) (Ricci, Salas et al. 1996).

Incorporating game design features such as contextualization, personalization and the provision of choices, produces a dramatic increase in learning motivation as well as the depth of engagement with the domain (Cordova and Lepper 1996). The U.S Navy conducted a similar training exercise for the submarine periscope handlers. The application, called Bottom Gun, was developed to examine the effects of the game-based training approach. Students using the Bottom Gun trainer were able to make better estimates of critical visual variables and showed greater improvement in periscope performance (Garris and Ahlers 2001, Prensky 2001).

Gamification is the usage of game-design elements, principles and characteristics in traditionally non-game contexts. The idea behind the process of gamification is to motivate and improve engagement in such tasks (Deterding, Dixon et al. 2011). Game

design elements such as check-ins and badges have been used in marketing strategies, where badges and check-ins give users a discount and increases customer loyalty (Kleinberg 2011). Gamification is widely used to motivate users to exercise and improve quality of health. User exercise statistics are clocked and timed, points are awarded based on these statistics and leaderboards are generated with the points earned. Users also receive badges for completing milestones and goals (Jeffries 2011). Gamification also can occur in the workplace. In a survey conducted in 2011, out of the 2000 people who participated 55% said that they would like to work for a company which uses games to increase productivity (Pleno 2011).Similar ideas of gamification are use in education, politics, hacker groups, training etc.

1.1 Prediction Games

Prediction games represent a category of gameplay where users predict future events based on their understanding of the domain and prior outcomes. Prediction games, of which fantasy sports are the best known category, are generally played online and have been rapidly gaining popularity. Players interpret historic and real time data and predict an event using the knowledge gained. These games are a subcategory of data driven games, as the foundation to understanding and interpreting the domain lies in the available data.

1.2 Fantasy Sports

Fantasy sports are the most games which follow the model of prediction games and have a history going back up to 50 years (Shipman 2001). Fantasy sports provide users with the opportunity to experience being in the shoes of a manager/coach and running their own virtual franchise sports team. The users manage their team and decide which players to draft in the team, which players to start each game period, and which players to trade with other players. The performance of the user's virtual team is computed based on the real-life performance of the athletes drafted by them. As such, roster decisions are predictions made before the actual sporting event happens.

According to Fantasy Sport Trade Association, as of September 2015 there are 56.8 million people above 12 years of age who have played fantasy sports in the US and Canada. Moreover, the average spending of a fantasy sports player (above 18 years of age) considering league-related costs, single player decisions and league-related materials in 2015 was $465 (FantasySportsTradeAssociation 2015).

One reason for such a high level of player motivation and engagement is the emphasis on competition, empowerment and participation. "The factors of competition, achievement, and surveillance had significant positive correlations with overall satisfaction while the factors of competition and camaraderie had significant positive correlations with future intentions" (Ruihley and Hardin 2011). To stay ahead in the competition, the engagement of the player with the actual sport also increases. According a survey conducted by Fantasy Sport Trade Association, 61% said that they watched more live sports because of fantasy sports and 60% said that they read more

about sports (FantasySportsTradeAssociation 2015). Research suggests that the continued involvement of the players with the activity leads to an increase in their level of domain knowledge (Garris, Ahlers et al. 2002, Davies and Graff 2005).

1.3 Models Based on Fantasy Games

Fantasy sports are a good example of gamification, where designs and characteristics of a game are applied in a non-game context. Fantasy sports follow guidelines from game design to get users more involved and engaged with the sport:

- They have a fixed goal to compete with other players and emerge victorious.
- They are easily accessible to a wide range of population over the internet.
- They require the users to draft a team for a future event.
- They warrant domain knowledge and enforce the users to learn more about the sport and the athletes.
- Leaderboards are created and users are awarded with points, achievements and badges.
- The game forces users to interpret historic data such as team and athlete statistics, fixtures, athlete form etc. as well as real-time data such as news articles, player interviews etc.

Fantasy sports educate users about the sport and athletes, while competing in a game. In such contexts, players learn about more than the sport. A middle school curriculum was designed to teach students about mathematical concepts while playing fantasy sports. The students had to follow all the normal rules for participating in the game such

as drafting a team and competing with other classmates. Students started to track the progress by interpreting and understanding the performance statistics of athletes, as well as reading newspapers and websites to gain more domain knowledge. They had to calculate the number of points each team received and implement the scoring using mathematical equations and calculations (Barr 2006).

Prediction games based on the fantasy sports model have been used in domains apart from sports to educate users about that domain while involving gameplay. An example of such prediction games was developed by the Harlan Law Institute and is called FantasySCOTUS. The game is a fantasy league for law students where given a particular case in the Supreme Court, players had to predict the verdict of each of the judges. Points were awarded for correct ruling predictions (Blackman, Aft et al. 2012).

1.4 Importance of News Articles in Prediction Games

The most important resource of a fantasy sports player is information. There are two types of information relevant to playing fantasy sports. The first type is historical and statistical data, which provides information about the various parameters useful in making a decision to draft an athlete. The second type of information is the near real-time data such as articles about a player's fitness, an expert's analysis of big game, weather conditions on the day of the game etc. The historical data can be gathered by mining data available from post-game statistics and records. Moreover data about statistics for fantasy sports are widely available online and are also provided by the fantasy sports website themselves (FantasyData , RotoGuru 2016).

This thesis aims at taking a deeper look into the information available to users from magazines, media, and the internet. I first discuss the growth in the audience for online news and then examine the relationship between news and prediction games in the form of fantasy sports.

The traditional newspaper was one of the first sources to convey information via mass media. On-line news sites vary their presentation, but the agenda is the still the same, i.e. to communicate textual information sometimes accompanied by graphical visualizations. A study conducted by Golovchinsky and Chignell, 1997, concluded that, "The newspaper represents a mature information presentation medium that is well suited to the display of relatively short, loosely related pieces of text" (Golovchinsky and Chignell 1997).

Statistics shows that it took only 7 years for the internet to reach 50 million users, whereas the telephone needed 50 years and the television needed 20 years (InternetWorldStats 2006). Since then, the rise and popularity of the internet is undeniable. "The dominant characteristic of the Internet is its ability to deliver information quickly and directly" (Straubhaar 1999). The first billion internet users were reached in 2005 and as of today there are 3.3 billion users on the internet (InternetLiveStats 2016). Figure 1.1 depicts the internet user growth since the start of the internet. For the first time in 2008, the number of users who rely on the internet (40%) overtook the number of users who rely on newspapers (35%) (Clark 2008, PewResearchCenter 2008). Figure 1.2 depicts the media share for news sources.

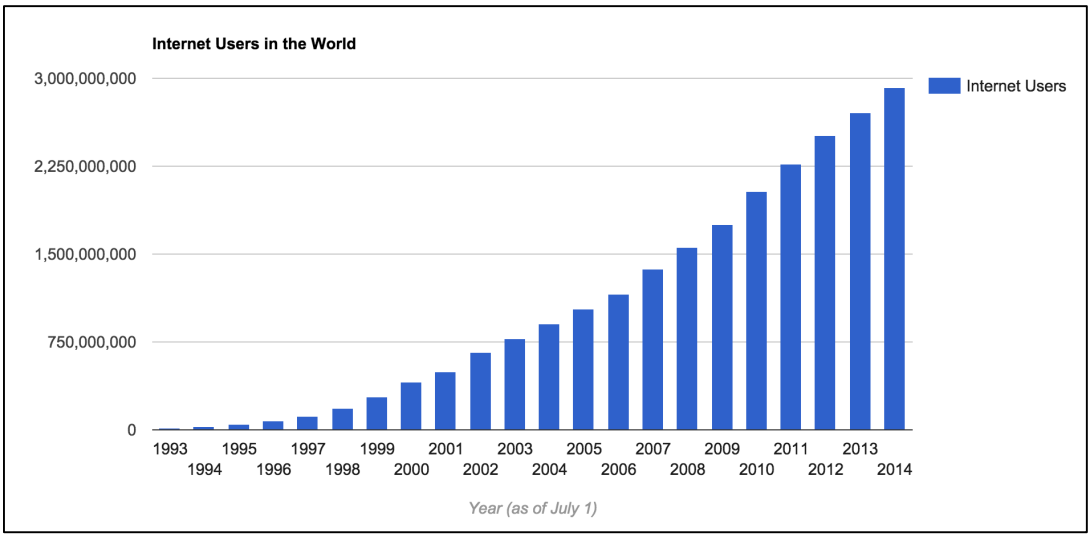

Figure 1.1 Internet user growth statistics (InternetLiveStats 2016)

Moreover, with an increasing number of people favoring mobile tablets and smartphones, the number of users relying on the internet as their primary source for news and information is rising. According to a survey conducted by NBC News (2,251 Americans, age 18 and older participated), 42 percent of those who own cell phones or tablet computers use those devices to check local weather reports and 24 percent check local sports scores and get updates. And when asked about the value of the local newspaper, 39% said that the loss of the newspaper would have no impact (Choney 2011).

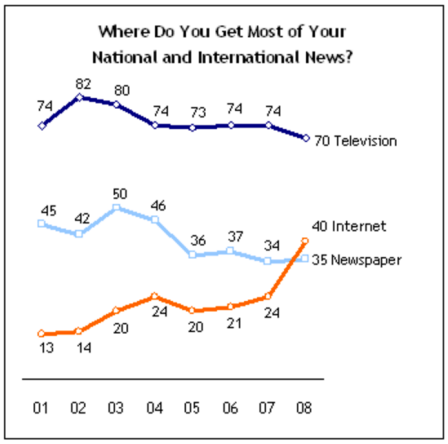

Figure 1.2 Percentage of people using different types of media as their news source (PewResearchCenter 2008)

Steven Whitlam and Hugh Preston, state that the relationship between organized sport and the media is symbiotic. The media provides information about the sport which makes the sport more attractive (Whitlam and Preston 1998). The popularity of fantasy sports has sparked an increase in the media, newspaper and television usage (Randle and Nyland 2008). Articles written about athlete trading conjectures, game result hunches, playing condition effects on the athletes, etc. along with articles providing statistical analyses and athlete comparisons are common and of value to fantasy sports players as well as other sports followers. 85% of the fantasy sports players frequently visit sports-related Internet sites and have readily available Web access (Felps 2000). A similar symbiosis between news and gaming may well be found in other domains modelled on prediction games. This thesis explores one such domain, using news sources to provide information about weather and climate.

There has been research done on examining the role of news articles on predicting domain information (Woodward 2005). Predictive machine learning and data mining techniques have been applied to investigate the correlation between financial news articles and the changes of stock trend (Fung, Yu et al. 2002). News articles were discovered to serve as the major source of non-quantifiable information on which predictions can be made. The strategies developed performed extremely well in predicting stock trends and showed encouraging results on measures of closeness, directional accuracy and simulated trading (Schumaker and Chen 2009).

1.5 Impact of Visualization

A challenge for providing news valuable to prediction games include the quantity of content available. There are 992 million sites present on the internet and all of these sites together cater to about 2.3 billion gigabytes of internet traffic on a daily basis (InternetLiveStats 2016). News articles are being published on media websites, blogs, social media, mobile apps, syndicated feeds (RSS) etc. There are 1331 daily newspapers published in the US (Statista 2014). "Fifty percent of social network users share or repost news stories, images, or videos, while nearly as many (46 percent) discuss news issues or events on social network sites" (Abramovich 2014). Figure 1.3 depicts the top 5 activities of users' on social media.

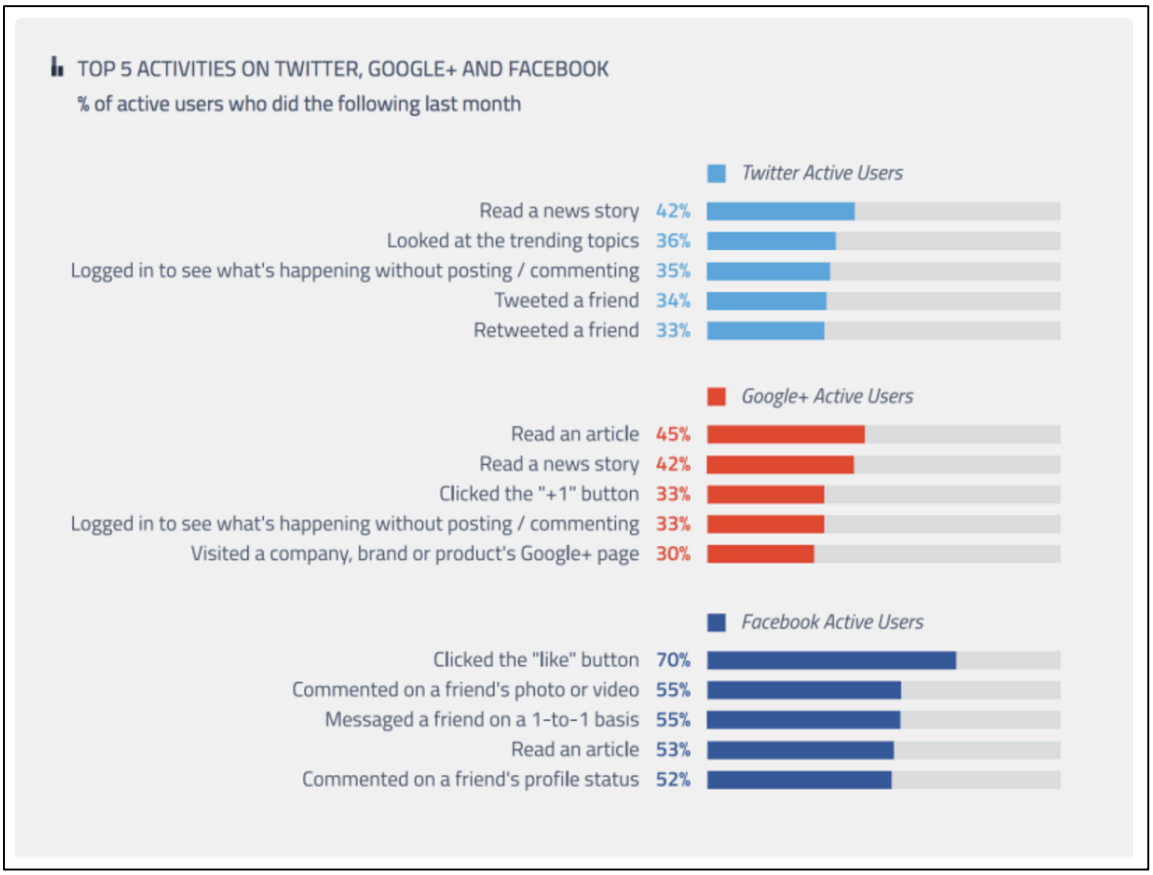

Figure 1.3 Social media as a platform for reading news articles (Abramovich 2014)

To analyze and find relevant information from such a big corpus of news articles is difficult. Data visualization is becoming an increasingly important component of analytics in the age of big data (SAS 2015). Visualization is an important criterion in helping users get a complete view of data and discover relevant items. The power of big data visualization can be harnessed by a number of smaller approaches such as dynamic changes in factors such as filtering and views of display (Wang, Wang et al. 2015). To help aid data analytics of a huge dataset, various data visualization methods have been developed such as treemaps, sunburst charts, streamgraph etc. The basic idea of

enhanced visualization techniques is to make it easier to glean intelligence from a mass of information (DeGeer 2014).

NewsMap is one such tool developed at trying to solve the data visualization problem of the enormous set of news articles present on the internet (Newsmap 2016). Figure 1.4 depicts the news map visualization. The application aggregates news articles from various sources and display them as a treemap. The size of each cell which represents a news article is directly proportional to the reporting volume of the topic. The news can be filtered by category (World, Business, Technology, Sports, Weather etc.) as well as by country (US, UK, Spain etc.). Clicking on a news article takes the user to the actual URL where the article is published (Fitzpatrick 2010).

Figure 1.4 The NewsMap visualization interface (Newsmap 2016)

From the previous discussion, it is evident that news articles play a significant role in fantasy sports and are an essential part of the decision making process of players. With the increase in the number of news articles being published over the internet, it has become an increasingly pain-staking process to search and find articles which pertain to a particular topic. The current size of the Google index is about 100 million gigabytes and the index consists of about 30,000 billion websites and pages across the internet (StatisticalBrainReserchInstitute 2015). Besides using search, another alternative is to go to a particular website or use a particular app to read news, but this limits the choice and assortment of content the user is exposed to. The ideal situation would be to read articles aggregated from different news sources and displayed on a single interface which provides a wide range of filtering and searching functionalities. Such sites exist, e.g. Google News, but they are not tailored with an understanding of the information needs of prediction game players.

1.6 The Hypothesis

Our research group is developing a prediction game engine to create informative and educational games in a variety of domains. The domain which we have chosen to test our research idea is climate science. The application which we have designed and developed to test our game is called “Fantasy Climate”. The intention of this game is to improve the player’s understanding of the climate domain by motivating them to interpret historical data and news articles concerning the domain. The hypothesis of this thesis is that a superior visualization of the domain news articles will lead to an increase

in user engagement and satisfy the information need of the users. The evaluation of this hypothesis also assesses the importance of news articles in the prediction making process.

1.7 The Proposed Visualization

Visualization technology, no matter how well designed, is of little educational value unless it engages users with the activity (Naps, Rößling et al. 2002). In a study conducted by visualizing algorithms to Computer Science students, it was observed that the learning increased as the level of engagement of the student with the experiment increased. But the most important outcome was that engagement increased with the additional tasks structured around the visualization (Grissom, McNally et al. 2003). This result indicates than having a good visualization for news is not enough; it must be an appropriate visualization to support the users' tasks. In our case, this means the visualization of news articles interface should aid in making the prediction decisions.

In general, to develop a tool which helps in discovering necessary pieces of data from a huge dataset, there are two issues which influence the design. The first is information visualization, where the emphasis is on giving users and insight into data distributions, and the second is data mining, where emphasis is on statistical algorithms and machine learning (Shneiderman 2001).

Climate change differs in different regions of the world, depending on geographic features such as the amount of sunlight a region receives, the ozone cover, proximity to oceans and water bodies, cloud cover, volcanic eruptions, altitude and

finally people themselves (The-National-Center-for-Atmospheric-Research). What better visualization to represent geographic factors other than a map? Maps visualize geospatial data in the most effective format and help their users understand geospatial data (Menno-Jan Kraak and Ormeling).

Access to interactive maps is not limited to professionals as it once was. Products such as Google Maps/Earth are customizable and easily available and to developers. Visualizing unknown and new data such as news articles by using the geospatial analysis and spatial searching provided by a map is very effectual. In human cognition understanding and memorability are intertwined (Borkin, Vo et al. 2013). A visualization is most memorable when it includes embellishments, more color, higher visual densities and recognizable human elements (such as a map) (Chun 2015).

There have been projects which utilize the map interface for displaying news articles such as http://liveuamap.com/, http://newspapermap.com/ and www.maplandia.com. But the most significant research project which studies the advantages of spatial searching enabled by a map over the conventional keyword searching is called NewsStand (Teitler, Lieberman et al. 2008, Samet, Sankaranarayanan et al. 2014). The study concluded that the implicit geographic content exposed to readers by a map improves their understanding of today's news (Teitler, Lieberman et al. 2008). But these projects consider news articles in a broader context and do not study their effect on engagement and their influence in improving learning of a particular domain.

The particular visualization developed, GeoNews, combines known techniques from natural language processing and visualization. The news articles are first

aggregated from different reliable sources available on the internet, pre-processed to extract the geographic location of focus and visualized on the map. The GeoNews tool is also equipped with a variety of features to filter and customize the news articles displayed.

The following chapters provide an overview of the Fantasy Climate application, then describe the GeoNews tool and its implementation, the methodology used to test the hypothesis by conducting a user study, and the results from the study. I conclude with a summary of the concepts and results with ideas for future work.

1.8 Why Climate Change?

Climate change is one of the most important domain which has a widespread impact on the entire world. Climate change is defined as a change in statistical distribution of weather patterns when the change lasts for an extended period of time (Wikipedia 2016). Climate change results from a plethora of causes, both natural and human, such as solar radiation, volcanic eruptions, global warming etc. Historic evidence such as the melting of glaciers, change in average weather conditions, sea level change and many other such indications supports the claim of climate change (Gargate 2013).

Climate change has a lot of impact directly and indirectly on human and animal life. Temperature records for the warmest year are being broken frequently, the sea level keeps rising constantly and glaciers retreat on a regular basis leading to disruptions in aquatic life which in turn disrupt economy. Floods, wildfires and draughts due to climate

change affect human life immensely and cause concerns in national security. Greenhouse gases deplete the ozone layer and have harmful effects on plant life and agriculture leading to a decline in world food supply (Rosenzweig and Parry 1994).

Countries and institutions treat global warming with utmost importance. NASA has a dedicated webpage to make people aware of climate change and displays statistics to epitomize climate change (NASA 2016). The United States Environment Protection Agency also has a similar initiative by detailing the impacts of climate change on the place where you live. It also has literature which explains the causes and ways to help tackle the problem (EPA 2016).

The following sections talk about the system and application overview of the Fantasy Climate application, the implementation details of the GeoNews tool, the methodology used to test the hypothesis by conducting a user study followed by the results from the study, conclusion and ideas for future work.

CHAPTER II

FANTASY CLIMATE

The Fantasy Climate prediction game has been developed at Texas A&M and is the context for this thesis. Here I briefly review the design of the Fantasy Climate interface and the prediction game architecture.

2.1 Fantasy Climate Overview

Once the user logs in to the application using his/her credentials, the activity homepage is presented. The homepage conveys the state of the activity and available options, data and help required to play the game. The five main components of the homepage are the activity widget, the message board, the news articles section, the header menu and the chat system. This is shown in Figure 2.1.

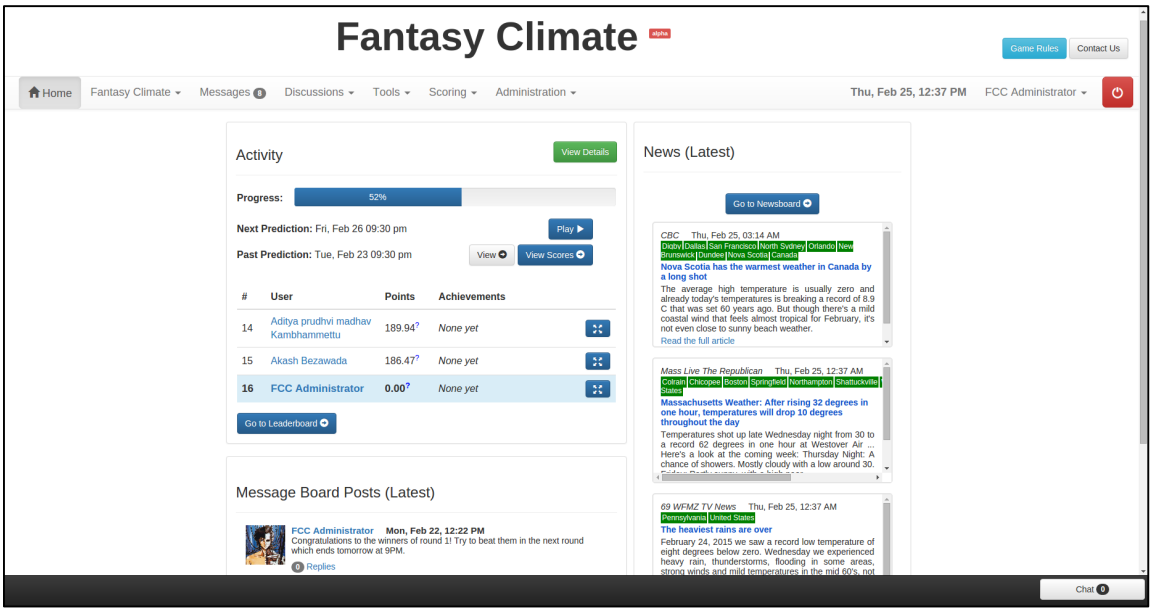

Figure 2.1 Home page of Fantasy Climate

The activity widget displays the current state of the activity. It shows the percentage of activity completed, the time of the next prediction, the time of the most recent past prediction and the leaderboard of the game. It also provides a button to go to the prediction page to make the next prediction.

The message board shows all the recent messages posted by users in the activity and provides a form of communication on the home page. Users can reply to messages from the homepage.

The news articles section displayed on the right hand side of the homepage consists of the recent news articles about the domain and provides a short summary which can be read on the homepage along with the option to read the full news article. This list only displays a few of the available articles, the entire list of news articles

which help in making predictions can be accessed using the button present above the list. It is the selection and visualization of news that is the focus of this thesis and will be discussed in more detail in the next chapter.

The Fantasy Climate header menu gives a wide variety of options to help the user navigate the system. It consists of the following sections:

- The home page link returns the player to the home page from any other screen.
- The activity dropdown provides options to view the game details, member directory (i.e. the list of users playing the game), and the prediction page to make predictions.
- The messages section allows users to send private messages to other users.
- The discussions dropdown lets the users access the forums and message board to engage in asynchronous communication.
- The tools dropdown provides access to the tools available in the game that aid interpreting historical and real time data sources.

Finally, the chat system on the bottom right of the screen enables synchronous communication with other online users in the application.

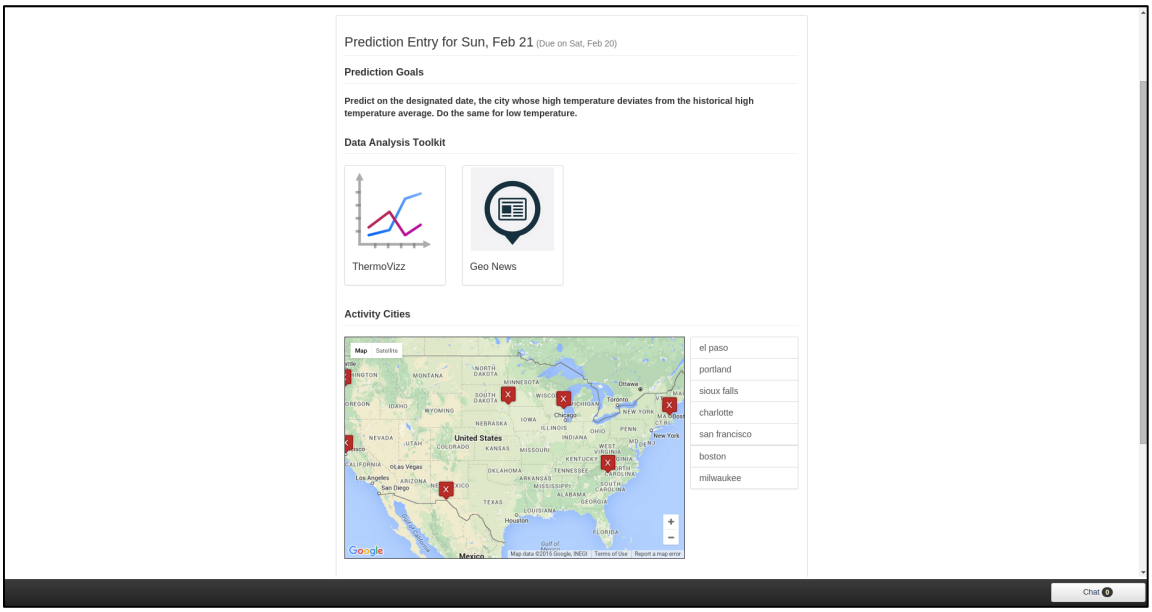

Figure 2.2 Prediction page in the Fantasy Climate application

2.1.1 Making Predictions

The prediction page (Figure 2.2) is where users make predictions in each stage of the activity. In Fantasy Climate, users need to predict two cities from a given list of cities:

1. The city whose high temperature on a specified upcoming day is expected to be greater than the historical average high temperature by the most.
2. And similarly the city whose low temperature on the specified day is expected to be lower than the historical average low temperature by the most.

Also available to the user on this page are data analysis tools to help users access and interpret historical data and news articles about the cities in consideration. The user can also see the scores calculated for prior predictions at the bottom of the page.

2.1.2 Scoring

Scores are calculated for the predictions made by the user for each stage of the fantasy climate activity. The formula used to calculate the scores is as follows:

overall score = base score + prediction score + achievements score

The base score is a fixed amount of points each user gets for each submitted prediction. The prediction score is calculated using the following formulas:

1. prediction score = predicted warmest city score + predicted coolest city score
2. predicted warmest city score = absolute maximum temperature – historical average maximum temperature
3. predicted coolest city score = historical average minimum temperature – absolute minimum temperature

Achievements are incentives provided to users who perform well in the game, for example the user with the highest score for a particular stage of the prediction activity receives 5 bonus points.

The scoring page (figure 2.3) displays the leaderboard consisting of all the users in the activity. Users can filter the scores based on a particular stage of the game or the overall points in the activity.

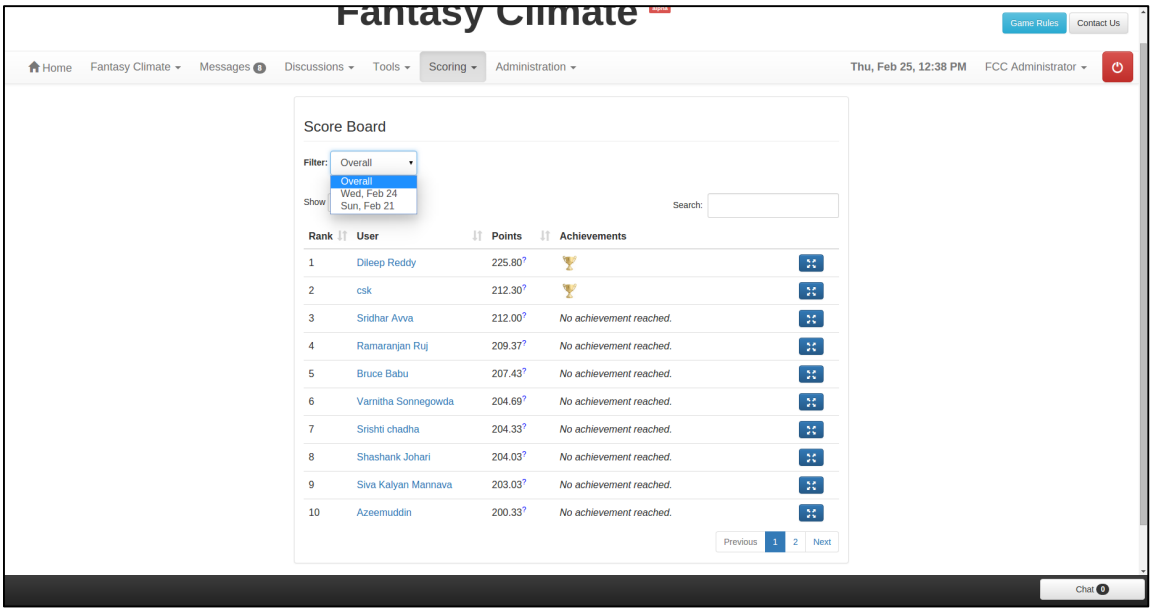

Figure 2.3 Scoring page in Fantasy Climate

2.1.3 Tools

The fantasy climate application has two main tools which help in making predictions: a historical data visualization tool and a related news tool. Here I describe Thermovizz, the historical data visualization tool. The related news tool, the focus of this thesis, is described in the next chapter.

Thermovizz provides visualizations of historic weather data for the activity cities as a graph. Each activity has a set of cities assigned to each prediction stage. Thermovizz gathers historic weather data (the high and low temperatures) for each of these cities over a period of time (starting from 1950 to the present). This data is then visualized as a graph to help users identify long-term trends in observed weather. There are two types of graph visualizations the users can choose from, the spline line and the

regression line. Figure 2.4 displays the spline line and figure 2.5 displays the regression line of the Thermovizz tool.

Figure 2.4 The thermovizz tool (Spline Line) inside Fantasy Climate

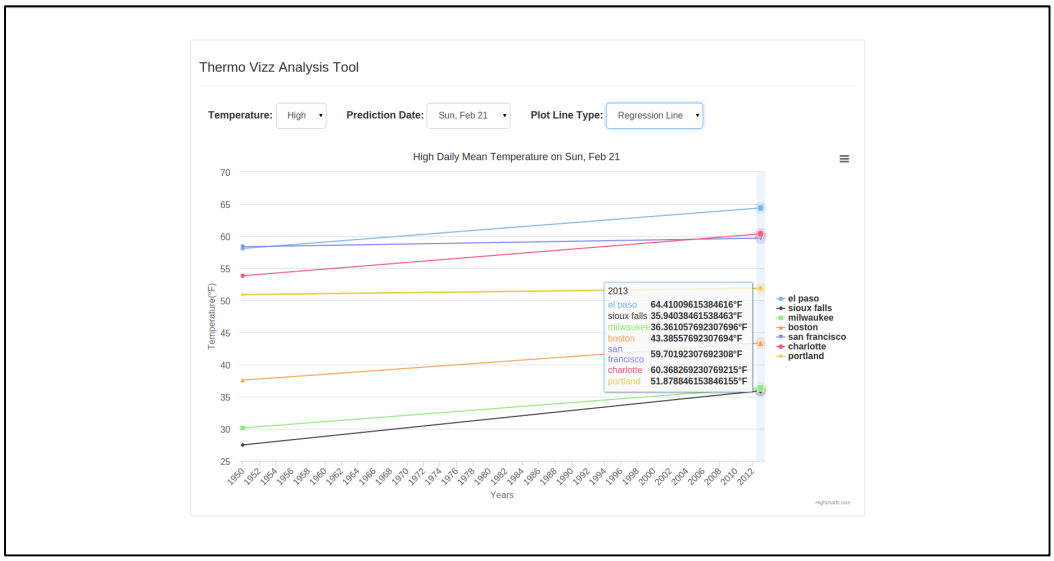

Figure 2.5 The thermovizz tool (Regression Line) inside Fantasy Climate

2.2 System Architecture and Implementation

Fantasy Climate has been designed as a web application. Modern web applications are conveniently described as a hybrid between a hypermedia and an information system (Stanley , Nielsen 1990) and have a lot of advantages, the major being the 'zero install' concept (Stanley). Today all computers and mobile devices have a web browser, so users can access the game from anywhere and at anytime without having to install it. Web apps can be maintained and updated without causing any major disruption to the users. Fantasy climate harnesses the power of open source assets present on the web such as binaries, libraries, and APIs.

The Fantasy Climate web application is comprised of three broad categories or layers:

- Data storage – This layer stores all the data used by the system. The data stored includes:
    a. data used by the system to run the application, and
    b. data provided by users during their interaction with the game.
- User interface – This layer supports interaction with and visualization of content and game activities. Interfaces to login to the system, setup the activity, accept predictions, visualize statistics and news articles, and display results are part of this layer.
- Backend – This is the controller between the data storage and user interface layers. It includes most of the logic involved in successfully running the system.

It persists and retrieves data from data storage, manipulates entities and serves requests from the user interfaces.

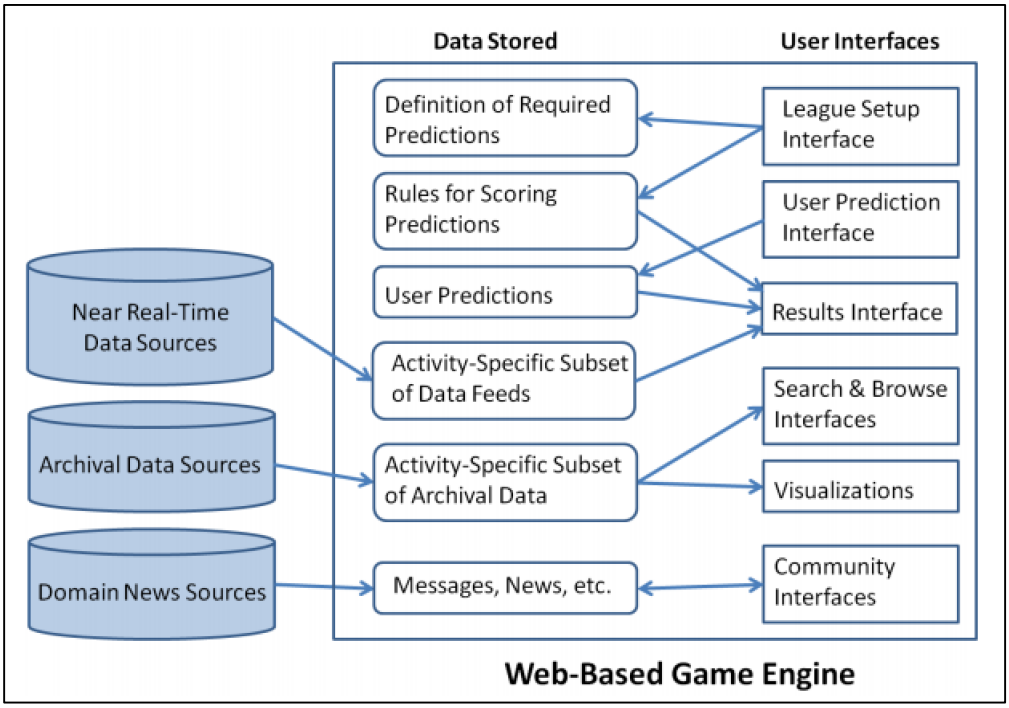

Figure 2.6 System level architecture of Fantasy Climate

Figure 2.6 shows the system level architecture of the Fantasy Climate system. The data storage layer uses both table-oriented storage and document-oriented storage. A MySQL database is used for the table storage and JSON files are used for document-oriented storage. All the game activity related data such as the list of users, their predictions and results, social communications in the game, notifications etc. is stored in the database. Statistical data such as historical high and low temperatures, precipitation and so on for all the cities displayed in the game are served using the document storage

(Atkinson, Bancilhon et al. 1989, Han, Haihong et al. 2011). The same idea also applies to the news articles and other related information.

The user interface is designed using programming languages which are system independent and can run on any browser. Thus HTML, CSS, Javascript and jQuery are the preferred technologies, which work on every browser irrespective of the platform. Any user with an internet connection has access to the game and does not have to install anything in order to run it, the browser takes care of all these background tasks. AJAX calls are used to communicate with the backend server to request or submit data. Since the interfaces need to be dynamic and handle interactions as well as in-place data displays, PHP is used to handle them (Garrett 2005, Paulson 2005). Bootstrap and other open source libraries are used to construct the web page structures and designs. All these assets and resources which form the front-end layer of the application are hosted on the Apache web server.

The backend is entirely built using Java. The core functionality libraries, the logic handling packages and the web request handling services form a major part of the backend. The functionality to access the data storage, retrieve and persist data is developed as a separate library which is incorporated in this layer. The scheduler services to handle recurring functionality such as data retrieval, news retrieval, score calculation, reminder notifications etc. is also part of this layer. The backend is hosted on a Tomcat server which exposes the request handling components as web services (Benslimane, Dustdar et al. 2008). The server listens for calls made to these services from the front end layer and delegates the request to the appropriate class. Once the

request is served, the web service responds to the front-end layer with the generated response. This mechanism uses the Representational State Transfer (REST) architecture.

CHAPTER III

RELATED NEWS TOOLS

This thesis explores how different methods of providing news articles related to weather and climate affect player engagement with and performance in Fantasy Climate.

3.1 News Tool Overview

Real time and recent data plays an important role in predicting locations that will be warmer or cooler than their historic norms. One important source of up to date information is news. News articles are available through a variety of digital and print media. The large quantity of news articles can lead to difficulty for users trying to locate articles for the particular locations and regions being considered in a Fantasy Climate prediction round.

To support the location and browsing of relevant news, Fantasy Climate provides a visualization of a corpus of weather and climate news articles gathered from sources available on the internet. Each news article is processed to extract the locations about which the news article is written. Additionally, a summary is generated for each of the articles. The articles are then presented to the user via one of two visualizations.

3.1.1 GeoNews

There has been a lot of research concerning the identification and visualizations of geo-locations for content. From an interface perspective, interactive map-based

exploration of retrieved data is of value (Rauber and Merkl 1999) during analysis. Map-based interfaces to news articles enable browsing content via location. Keyword or query-term access requires users to identify the terms they care about while there may be many different terms relevant for the same geographic location. In models where the user is not sure about what keyword to search for, keyword-based searches are not very effective (Samet, Sankaranarayanan et al. 2014). Maps also provide users the ability to explore data they are not exactly looking for (Teitler, Lieberman et al. 2008).

GeoNews presents news article on a map as shown in Figure 3.1 to allow cartographic exploration and spatial search. In addition to the pins on the map, the user can search for particular topics (Figure 3.1 top right) and see the title and other metadata about the articles on the right side of the interface, along with the functionality to read the full article in place using a reader present in the interface. At the bottom of the map are the locations part of the current prediction round. Clicking on one of these filters the articles to show articles related to that location.

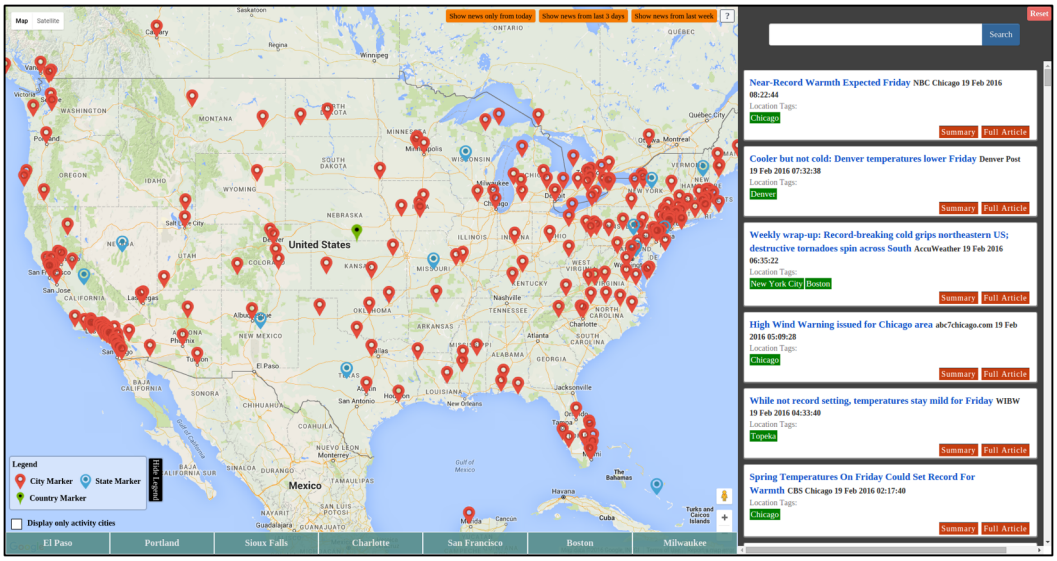

Figure 3.1 The GeoNews tool interface

3.1.2 News Board

The News Board visualization presents a list view of the news articles much like the interfaces found in news aggregators such as Google News. The news articles and their summary are displayed as a list ordered on the basis of their publishing dates. Searching for particular terms (top of Figure 3.2) filters the list to matching articles. Selecting a location from the current prediction round from above the list filters the set of articles to show those related to that location.

The main difference between GeoNews and the News Board is that the News Board provides the summary of each article while GeoNews presents a map view of the locations discussed in the articles.

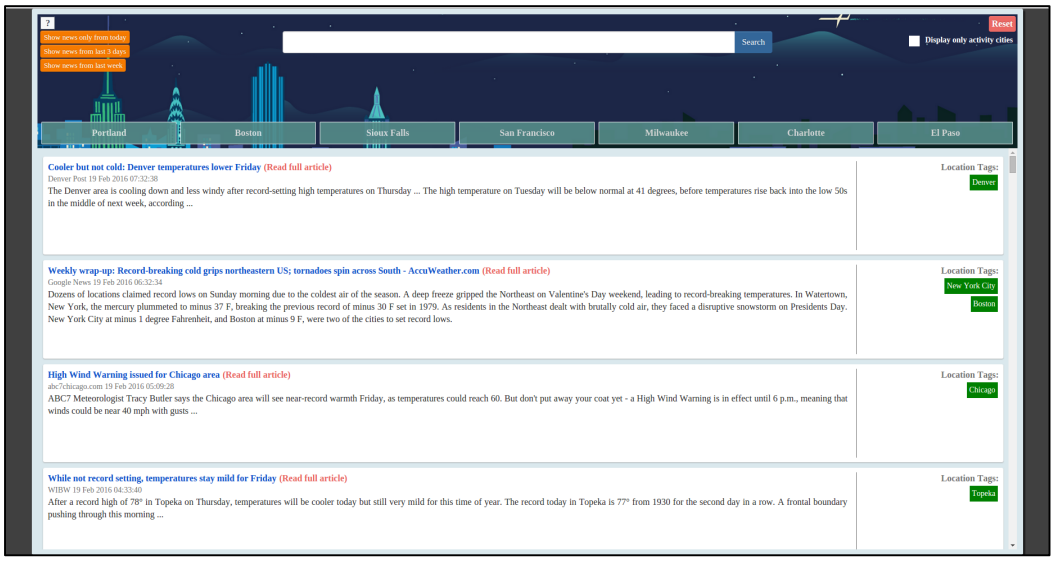

Figure 3.2 The NewsBoard (List News) tool interface

3.2 Implementation Details

The following are the steps involved in selecting, geolocating, and visualizing news articles in Fantasy Climate.

3.2.1 Gathering News Articles

With the abundance of articles present in digital media from sources such as news publishers, blogs, social media, RSS feeds etc. the amount of information available is ever increasing. Fantasy Climate gathers articles from RSS feeds and APIs provided by Google News and Bing. The articles are requested by supplying keywords only related to weather and climate. These news articles are stored in the database along with related fields such as publishing date, source, URL etc. Some news articles also contain

the summary of the artifact; those are stored as well. The extraction of news articles is done by a scheduler on a regular basis to update the news available.

3.2.2 Extraction of Entities

Each news article gathered from the sources is then processed to extract the entities present inside them. The natural language processing library provided by Stanford's NLP group is used to extract entities (Manning, Surdeanu et al. 2014). The Named Entity Recognizer (NER) present inside the library labels words or sequences of words which correspond to either a person, an organization or a location. The locations identified by the NER are stored in a list for further processing.

3.2.3 Identifying Locations of Focus

From the list of locations identified by the NER, the next step is to identify those that are the focus of the article. In a news article, a lot of locations may be mentioned either as a reference or as a point of comparison. There are a number of approaches to select the focus location(s) from among the mentioned locations. The highest frequency algorithm uses a count of the number of times each location is mentioned and selects those with the highest number of mentions. Similarly, the weighted frequency algorithm adjusts the number of mentions by weighted their likelihood based on their distance from the location mentioned at the start of the article. For this work, the maximum frequency algorithm was used.

3.2.4 Generating Summaries

A summary is valuable to quickly assess whether an article will be valuable to the current task. Some news sources provide summaries of their articles, but some do not. The summaries provided by sources are used without processing them further, but for articles provided by sources which do not provide summaries, a shorter version of the article is generated.

To generate a summary of the news article, the text from the article is extracted using an open source project called BoilerPipe (BoilerPipe). BoilerPipe provides algorithms to detect and remove surplus clutter from the HTML format available through the news links. After extracting the pure text form of the news article, it is used to generate the summary. TextSummarization.net has an online API which provides text summaries of given web pages or text based on natural language processing and machine learning technologies (TextSummarization). The pure text of the article is sent as a request to the TextSummarization API and the response returned is stored with the news article as its summary.

3.2.5 Visualization

GeoNews: Visualization on the Map. The map used in the Fantasy Climate interface is provided by the Google Maps API. The maps API provides enables panning, zooming, map/satellite view, adding markers, info windows etc.

The GeoNews interface requests for the geo-tagged set of news articles already pre-processed by the backend. The articles are then grouped by their locations of focus.

Each location may contain one or more news articles and a marker is placed on the map with the location's latitude and longitude. Each of the three geographic categories – city, state and country has a different color marker associated and a legend is provided to explain this. After this step the map contains markers for each location that was identified as a focus in the set of available news articles.

Along with visualization on the map, all the news articles are displayed in a list in the right hand panel of the GeoNews interface. These listings are ordered by their publishing date, with the most recent article on top. These entries display the geo-location tags along with the option to read the summary or the full article.

Implementation of News Board: List of Summaries Interface. In order find the effectiveness of the map interface over the traditional list interface, the split testing technique is used (Deng, Xu et al. 2013). All the data processing stages and user interface features (such as filters and keyword-searching) are the same in both interfaces (Kohavi, Deng et al. 2012).

Each entry in the news list displays the title, source of the article, publishing date and the summary. The geo-tags associated with each news article are displayed along with the article. The option to read the full article is also displayed with the entry. The NewsBoard also supports similar filters to manipulate the list of articles such as filtering based on publishing date, search terms and activity cities.

CHAPTER IV

GAMEPLAY

The gameplay of Fantasy Climate follows that of fantasy sports in general. Players join the activity, groups and leagues are created, predictions are made and the player with the best predictions wins. There are three stages in the gameplay: Pre-activity, Activity, and Post-activity.

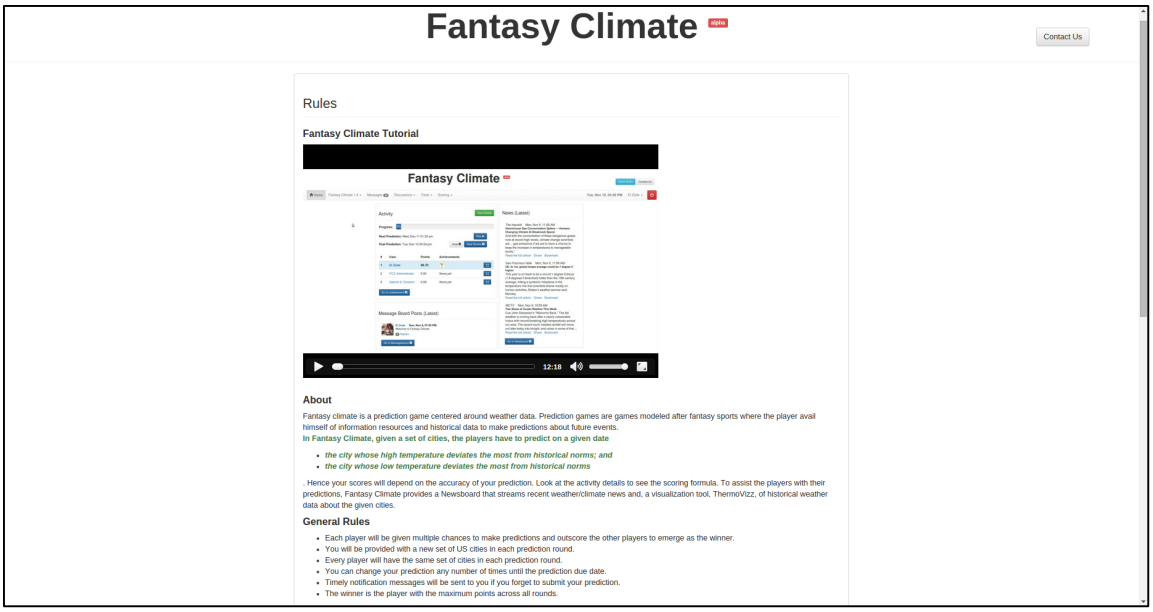

Figure 4.1 The rules page which trains users before the start of the game

4.1 Pre-Activity

During this phase of the game, the login page contains the link to the rules page, where rules of the game and training data are available. The rules page (Figure 4.1) gives

a general overview of the activity along with the code of conduct. It also contains a training video which provides a detailed walkthrough of the application, explaining all the features and options.

Users can log in to the system, and have a look at the homepage (figure 4.2). The activity details panel only displays the time at which the activity is scheduled to start. The news articles on the homepage are displayed, as well as with option to open the NewsBoard/GeoNews tool. The chat system at the bottom of the homepage is disabled, and is activated only when the game starts.

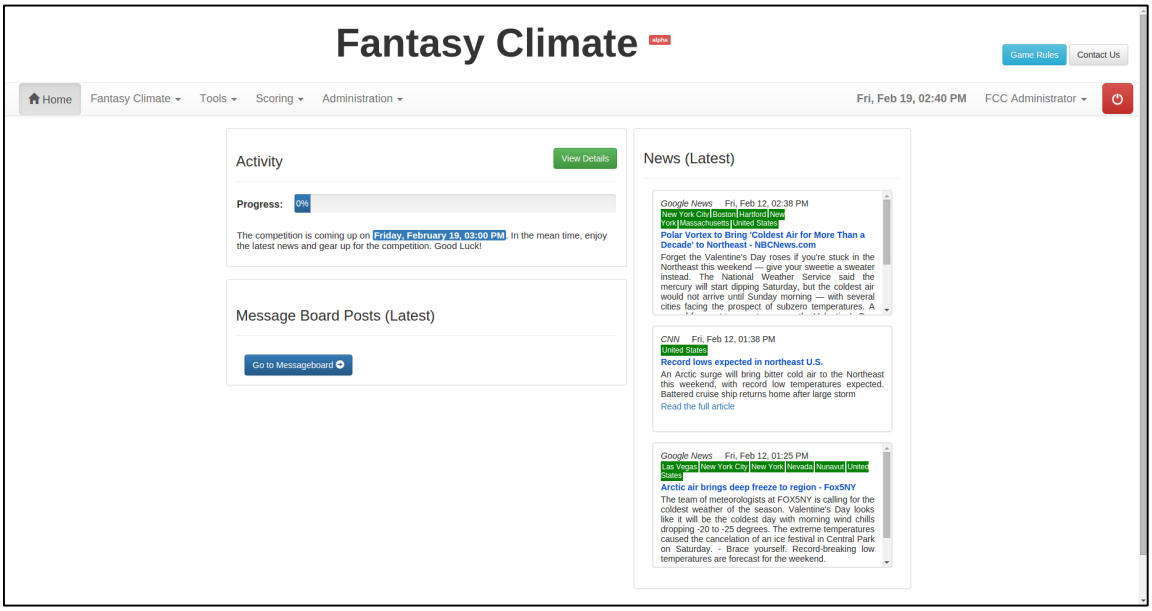

Figure 4.2 The user home page before the start of the activity

Users can also view the game details (Figure 4.3) which gives a brief overview of the game details and schedule. The member directory is also available to see who the other users in the activity are.

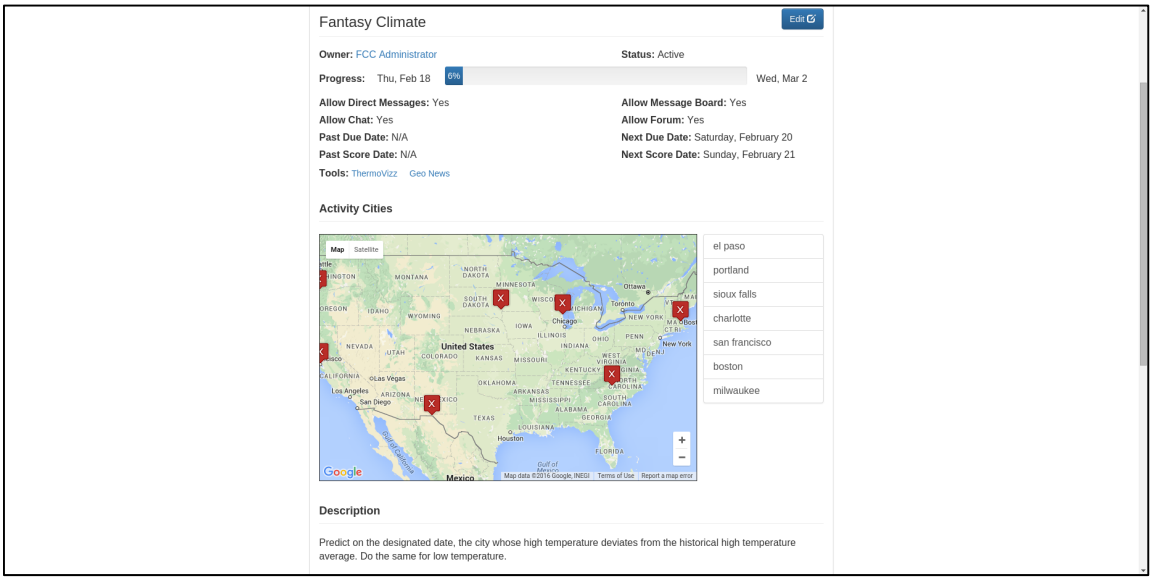

Figure 4.3 The game details page which users can access throughout the game

4.2 Activity

After the activity starts, users can start making predictions using the submission page, where the city going warmer and the city going cooler need to be selected. Also, all the tools and features available in the game are enabled. This enables users to develop better interpretations about the weather and climate of each city.

4.2.1 Prediction Submission

For a given activity, the number of stages of predictions is decided beforehand. Each stage contains a different cluster of cities. On the submission stage (Figure 4.4), there are four components.

- The first section is the list of data analysis tools available for the players in the game. These are displayed on the submission screen for easier access. They can also be accessed from the header menu present on the screen.
- The set of cities for the present activity stage are displayed on a map, accompanies by a list on the side. Clicking on a location, displays the maximum and minimum temperatures for the current day.

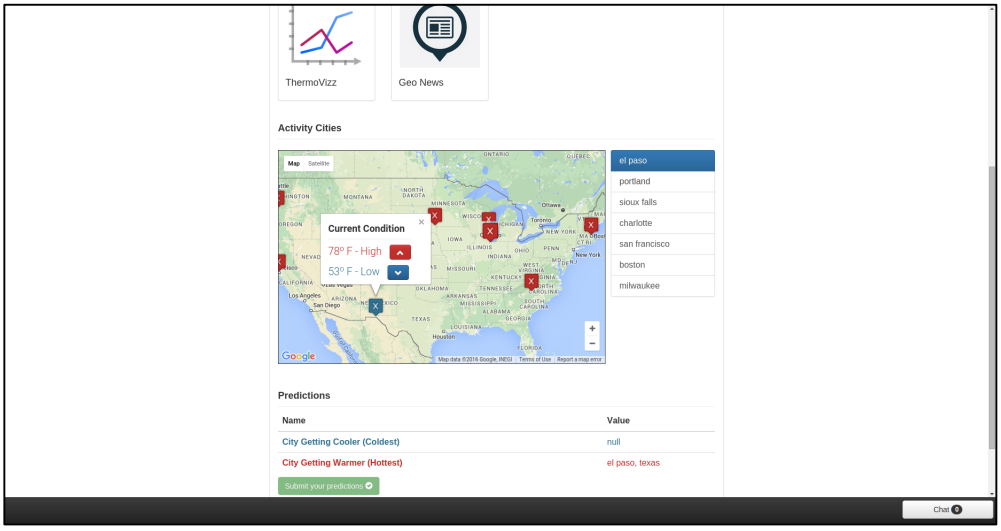

Figure 4.4 Selecting a city as part of the prediction

- The next section displays the user selections for the city expected to be warmest relative to its historical average and the city expected to be coolest relative to its historical average.
- The final section is only displayed if one stage of predictions has passed. This section displays the cities selected in the previous round and the score calculated by the system for these cities (Figure 4.5).

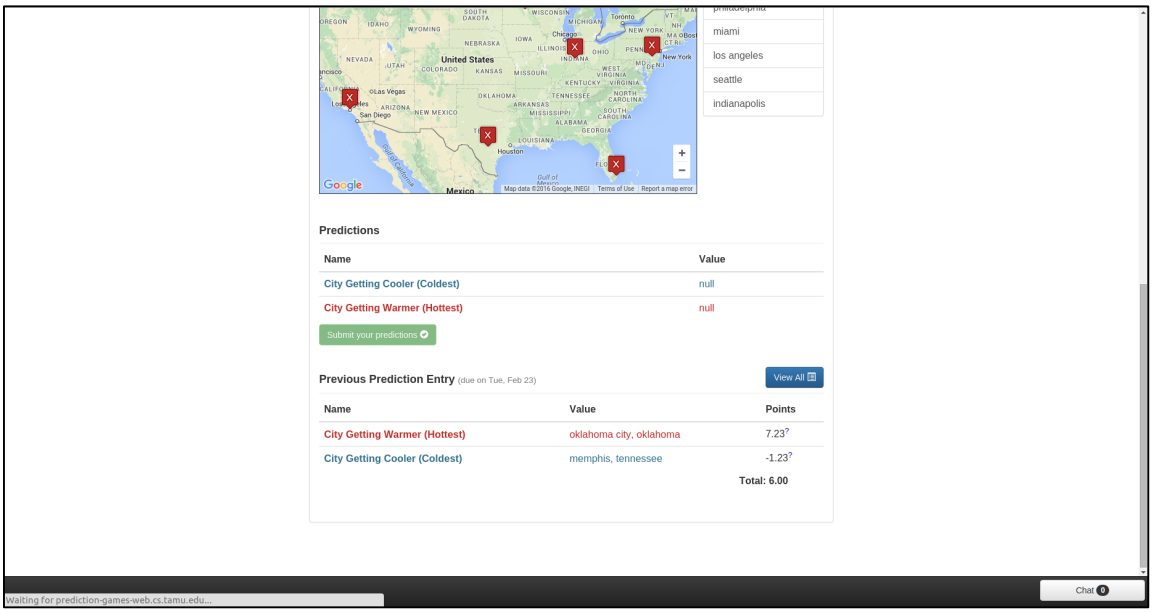

Figure 4.5 The points section which shows points earned in the previous round

4.2.2 Communication

Communication is a vital part of social interaction among people. Games have been identified to act as a social activity and help build social connections. They act as a medium which fosters togetherness, collaboration, competition and interaction (Kappen,

Gregory et al. 2013). Communication has also been shown to enhance player experience and engagement with the game. According to the study conducted by Schoenau-Fog, communication within a game motivates players to be more involved and return to keep playing. Chatting or writing about experiences can encourage continued play (Consalvo 2011, Schoenau-Fog 2011).

On the basis of this evidence, Fantasy Climate includes both asynchronous and synchronous communication channels for players. It includes a message board that displays messages in temporal order and is visible on the homepage, a forum for topic-based discussions, and direct player-to-player in-game messaging. There is also a chat system where players can engage in real-time conversations with other players currently logged into the game.

4.2.3 Scoring and Achievements

A key to prediction games is competition among players. Additionally, evaluation and feedback motivate players to continue playing a game (Garris and Ahlers 2001). The score board (Figure 4.6) provides feedback on prior predictions in Fantasy Climate, enabling players to view scores for a particular stage of the competition or the overall score.

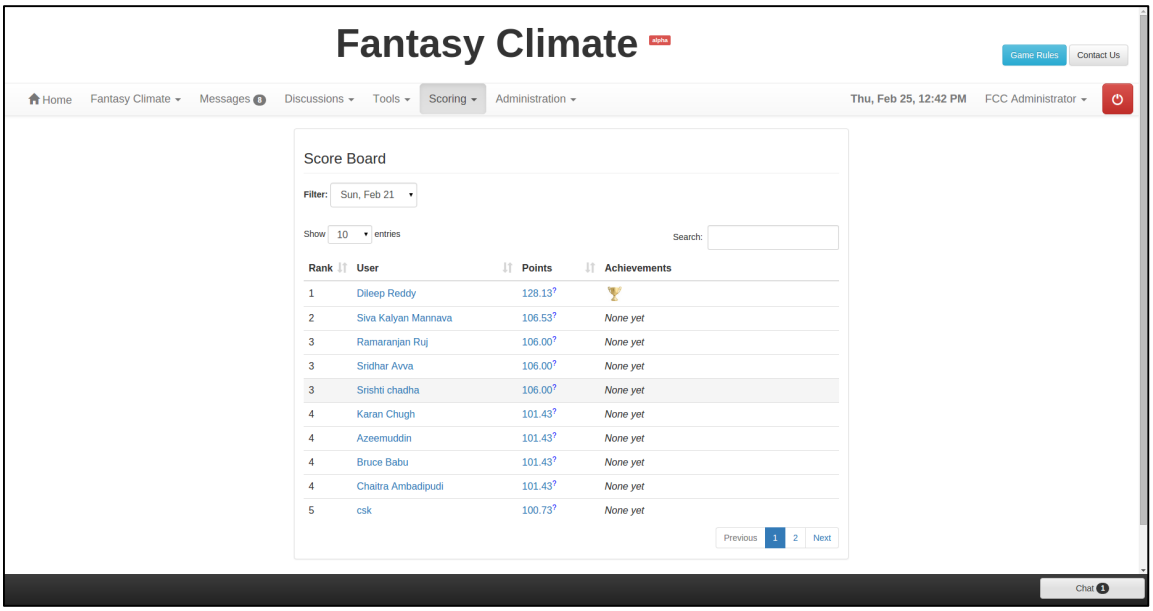

Figure 4.6 The leaderboard table in the scoring page which shows the points earned by each user

Along with viewing the points for each player in the score board, users can also view the score breakdown and the reasons for someone performing better and someone performing poorly. The score breakdown (Figure 4.7) displays the following statistics for each user in the scoreboard:

- Prediction Stage/Round
- Prediction Date
- City selected by the user as the warmest and coldest for the prediction date
- Absolute high/low temperature observed for the selected cities
- Historical average high/low temperature norms for the selected cities
- Points calculated for the predicted city

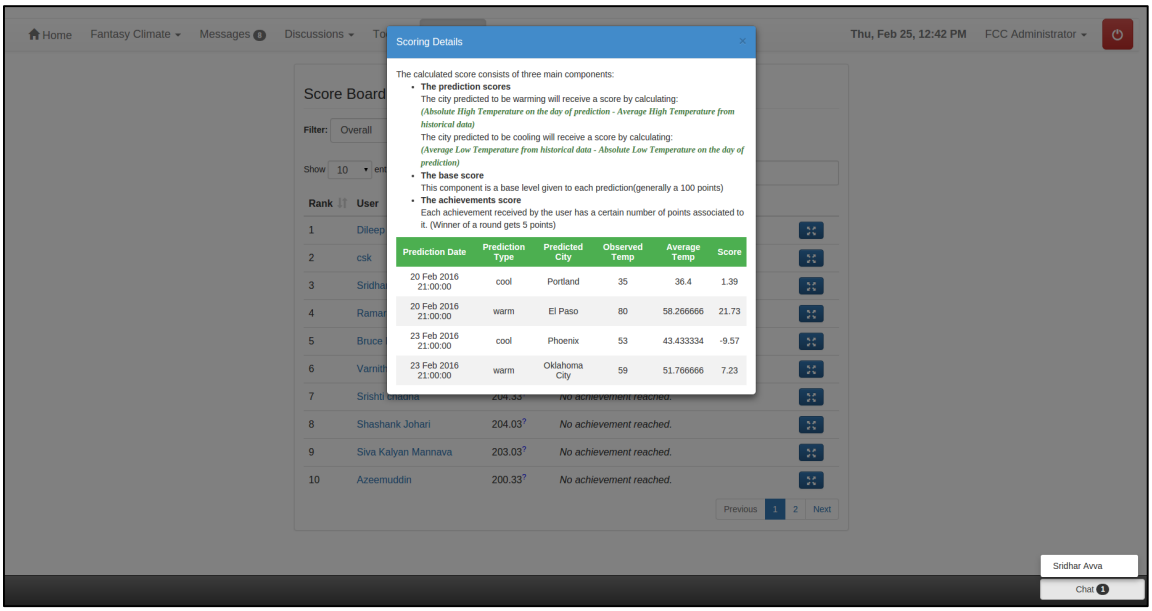

Figure 4.7 The score breakdown dialog which gives detailed analysis on how the points were allocated

4.2.4 News Visualizations

4.2.4.1 GeoNews

Users can interact with the map by panning and zooming to find content as well use of the many custom features provided on top:

1. Filtering news articles - There are a number of options provided to filter the news articles such as:
    - Filtering based on activity cities – display news articles only for the set of cities from which the predictions need to be made for the current stage.
    - Filtering based on search terms – display articles which contain the search term either in the article, the title of the article or the geo-tags.

- Filtering based on dates – provides options to display articles either from the last week, the last three days or the present day.

2. User interactions and available options - The most important part of any data visualization is the interaction it provides, allowing the user to manipulate views and display what the user exactly wants to see. The map is directly linked to the list on the right panel. Each article on the right correlates to the markers displayed on the map. The following features emphasize the integration of the two components:
    - When a marker is clicked on the map, the map opens an info window (Figure 4.8) showing the news articles one at a time which the users can browse through. This click also triggers the right panel list to display only those articles which pertain to that marker. Closing the info window, or clicking anywhere else on the map closes the info window on the map and resets the right panel to display the list which was previously being displayed.

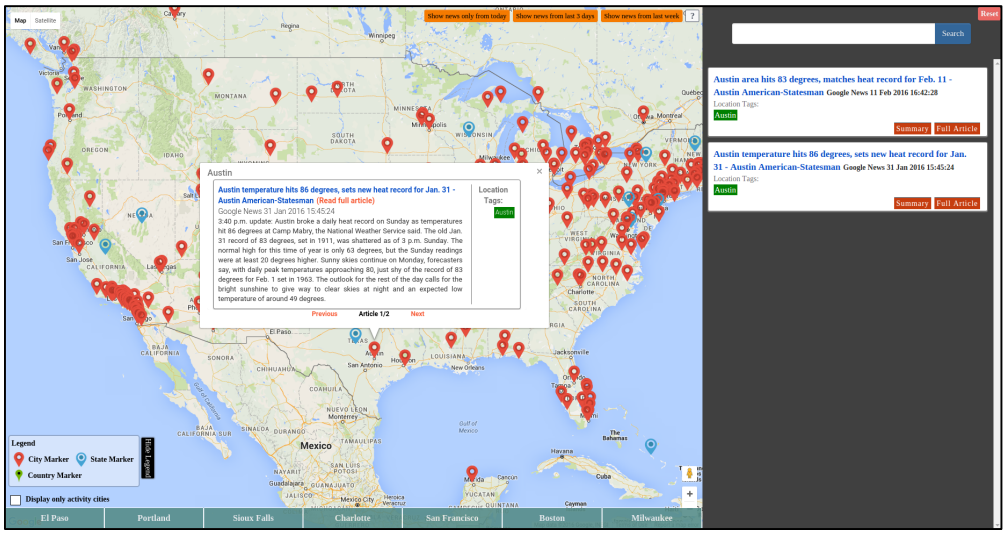

Figure 4.8 Clicking on a marker on the GeoNews interface

- The filtering options available in the tool (filtering based on activity cities, search terms and activity cities) also affect both the map and the right panel list. The filters can be enable/disabled individually or can together form a complex filter, the news articles are filtered based on all the filters currently enabled by the user. The map and the right panel list display only this subset of articles (Figure 4.9).

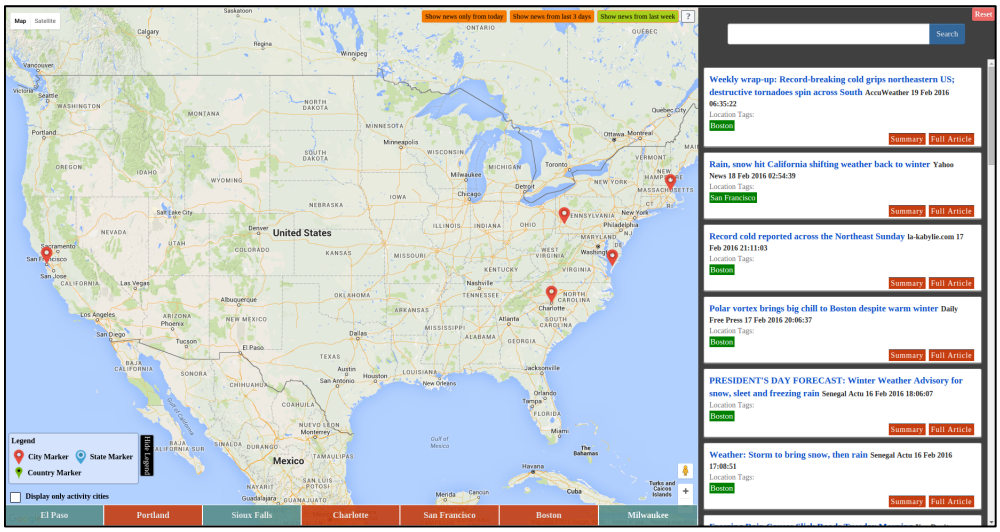

Figure 4.9 Filtering the news articles based on time frame and activity cities

- On panning across the map, or when zooming in/out of the map, the area on the map keeps changing and the markers present in the view also change. The right panel list follows the exact the same behavior. As the view on the map keeps changing, the right panel filters news articles to display only those which represent the markers in the current view. This provides an additional filtering option based on spatial searching and discovery. This is shown in Figure 4.10.

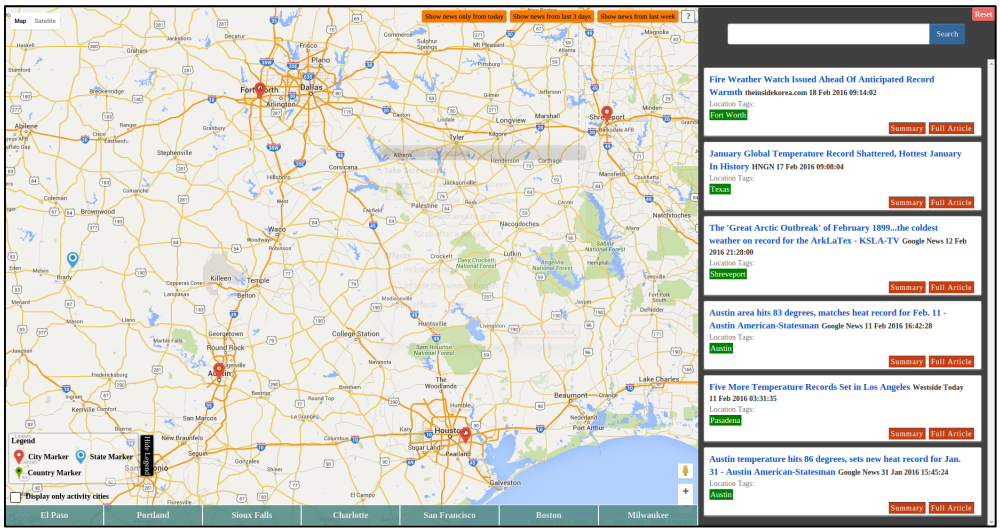

Figure 4.10 Filtering news articles based on zooming and panning on the map in the GeoNews tool

- The reset button present in the tool, clears all filters and restores the map's position and zoom to the initial value.

4.2.4.2 NewsBoard

The news board enables the same set of features as those provided in the map interface, such as filtering based on the date, the activity cities for the current prediction round, and searching using keywords. These can be used together to create a complex filter as well. See Figure 4.11.

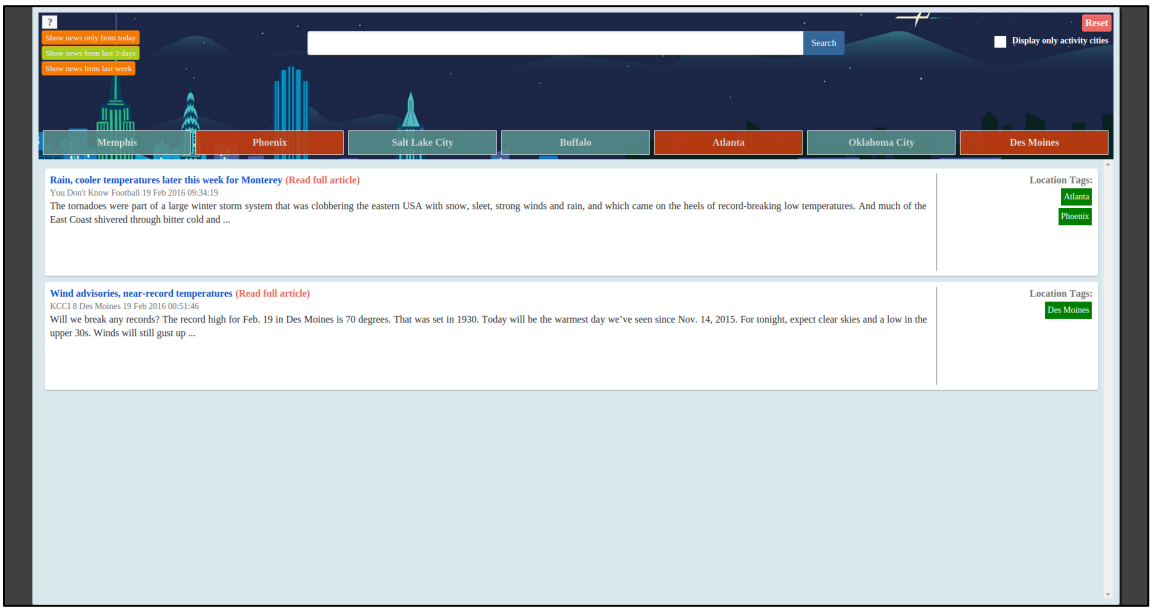

Figure 4.11 Figure displaying the NewsBoard (List News) interface when multiple filters such as time frame and activity cities have been selected

CHAPTER V

USER STUDY

To evaluate the effect of visualization of news articles, a user study was conducted using the Fantasy Climate prediction game. User studies are important for evaluating particular features of designs and they provide insight into overall interface usability and experience (Kittur, Chi et al. 2008). The user study aims at distinguishing the effects of the two interfaces, GeoNews and NewsBoard, within the domain of prediction games. Students from Texas A&M University were invited to participate in the user study. None of the participants had any kind of involvement in the design, development, testing or deployment of Fantasy Climate.

5.1 Methodology

The use on online controlled experiments has grown over time. Web facing companies such as Facebook, Google, Amazon, LinkedIn use controlled experiments to guide product development and gain valuable feedback (Kohavi, Deng et al. 2013). A particular category of controlled experiments is known as A/B tests, split tests or randomized experiments. They have had profound influence in multiple fields, including software development, medicine, agriculture and advertising (Kohavi and Longbotham 2011). The basic idea behind A/B testing is simple, given a base system and system with a new feature/change, split testing establishes a relationship between the changes and their influence on user-observable behavior (Kohavi, Henne et al. 2007).

Following the directions of the split testing mechanism, the participants were divided into two groups. All the features and tools provided to both groups were identical except that one of the groups was provided the GeoNews tool to explore news articles whereas the other group was provided the traditional NewsBoard tool. All the steps in the news gathering and processing stages were exactly identical except for the interfaces. The participants had no information about the other group or the alternate tool design.

5.2 Protocol

The Fantasy Climate game was hosted on a public server so the only requirement for participants was to have a web browser and an internet connection. Users had continuous access to the application throughout the period of the user study. There were two separate activities created for the two groups on the same server. All the participants in a group were oblivious to the activity being conducted with the other group. The user study consisted of three stages:

1. Pre-Task Questionnaire – After agreeing to participate in the user study, they were asked to sign a consent form followed by a pre-questionnaire form. The pre-task questionnaire asked for demographic information such as age, gender and ethnicity. These questions were followed by questions which gauged the participants' prior experience with fantasy sports. Finally, questions regarding participants' interest in climate data and weather news were presented.

2. Prediction Activity – After all the users had filled the consent and pre-questionnaire, the participants were randomly divided into two groups. The first group received the version of the fantasy climate game which was equipped with GeoNews, and the second group received the version of the game with the NewsBoard. Users were registered into their respective game and the activity was started. More details about the prediction activity are described later.
3. Post-Task Questionnaire – During the final stage of the prediction game, users were given the post-task questionnaire to summarize their experiences with news article visualization and also about the prediction game in general. The post-task questionnaire was designed so that the research hypothesis was not reflected in the questions (O'Brien and Toms 2010).

5.3 Activity Timeline

The user study was conducted over a period of 13 days. There were three stages of predictions in the activity for three different city clusters. The timeline for the user study is depicted in figure 5.1. The following gives a detailed timeline of the activities required of the participants:

1. The first two days of the user study did not have a stage of prediction. This time was given to the participants to familiarize themselves with the fantasy climate system and explore the features present in the game. Participants could go through the rules page which has a detailed walkthrough video of the entire

fantasy climate system. They could also explore weather-related news with the corresponding tool available to their group.

2. The next phase in the timeline consisted of three sets of predictions for the three city clusters. Both participant groups followed the same timeline and schedule, with the city clusters also given in the same order. Each prediction stage was scheduled to last three days. The predictions made were scored on the next day after observing the absolute temperatures in the predicted cities.
3. After the three stages of predictions, the system was kept online for two more days so that participants could come back to the game and explore news articles, analyze scores and individual prediction breakdown, view achievements and communicate with other participants.

5.4 City Clusters

The city clusters, i.e. the set of cities that were being chosen from, were different for each prediction stage to make it more challenging to the participants. Each cluster included 7 cities from which to choose. The cities in each cluster were spread across the entire country of the United States in order to encourage players to analyze weather conditions of different regions of the country.

Figure 5.1 User study schedule and timeline

CHAPTER VI

EVALUATION AND RESULTS

6.1 Pre-Task Questionnaire

The pre-questionnaire provides information about the demographics of the players along with their prior experience with prediction games and their interest with weather-related news articles. The pre-task questionnaire was filled out by all the participants before the start of the user study. The following sections provide detailed information about the participants.

6.1.1 Demographics

29 students currently enrolled in the Texas A&M University participated in the study. Out of the 29 students, 27 of them belonged to the age group of 23-27 and 2 participants were in the age group of 18-22. Also, 22 participants were male and 7 were female. The study consisted of participants from a diverse set of backgrounds varying from Engineering (Computer Science, Computer Engineering, Civil Engineering, Industrial Engineering, Petroleum Engineering, Construction Engineering and Management), Sciences, Architecture, Business, and Management.

6.1.2 Experience with Fantasy Sports

Figure 6.1 shows the participants' responses when asked about the number of years they have had played any form of fantasy sports. Out of the 29 participants, 16 had

never played fantasy sports before, 6 participants had 0-1 year of experience, 4 had 1-2 years, 2 participants had 2-3 years and 1 participant had over 4 years of experience.

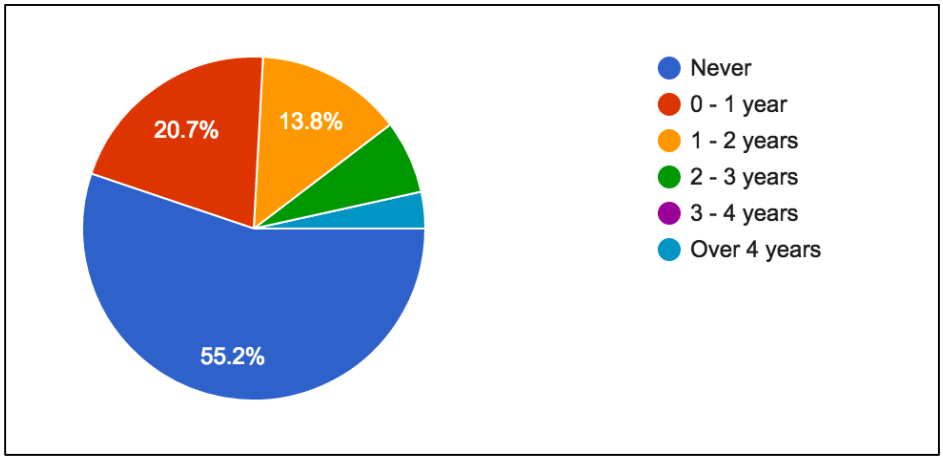

Figure 6.1 Participants' experience with fantasy sports

The next questions in the pre-task questionnaire asked about participants' most recent experiences playing fantasy sports. Participants who had not played before were expected to respond that the question was Not Applicable. Figure 6.2 shows their estimate of the average duration of their interaction in the previous or the current season. Out of the 14 participants who felt it was applicable to them, 4 participants had less than 15 minutes of interaction, 3 participants had 15-30 minutes of interaction and 7 participants had an average interaction of more than one hour.

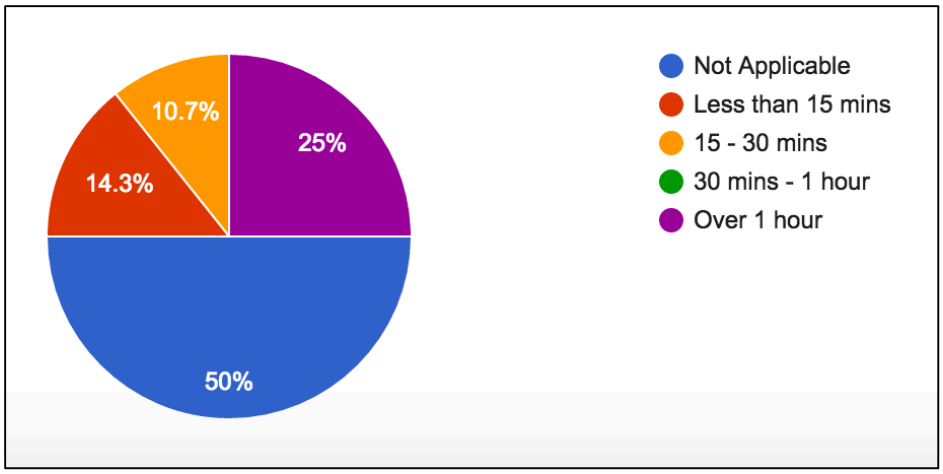

Figure 6.2 Participants' average duration length during their last interaction

The participants were also asked about the frequency of interactions they had with the fantasy game in their last season or current season. Out of the 12 participants responding to the question, 4 participants interacted once per week, 4 participants interacted 2-4 times per week, 3 participants interacted 4-7 times a week and one participant had more than 7 interactions per week. (See Figure 6.3)

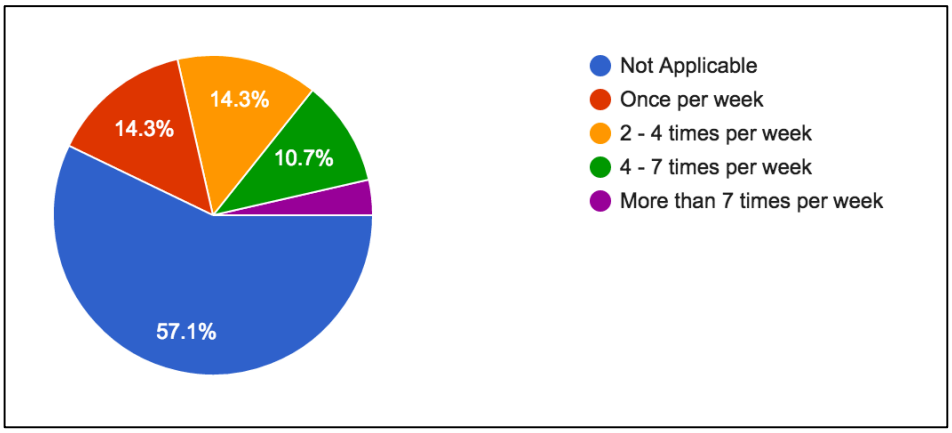

Figure 6.3 Frequency of interactions of the participants in their last or current fantasy sports season

To gauge the types of information resources participants used to make a prediction, they were asked to respond on their use of in-game resources, external online resources and offline on a scale from strongly agree to strongly disagree. Figure 6.4 shows the responses of the participants.

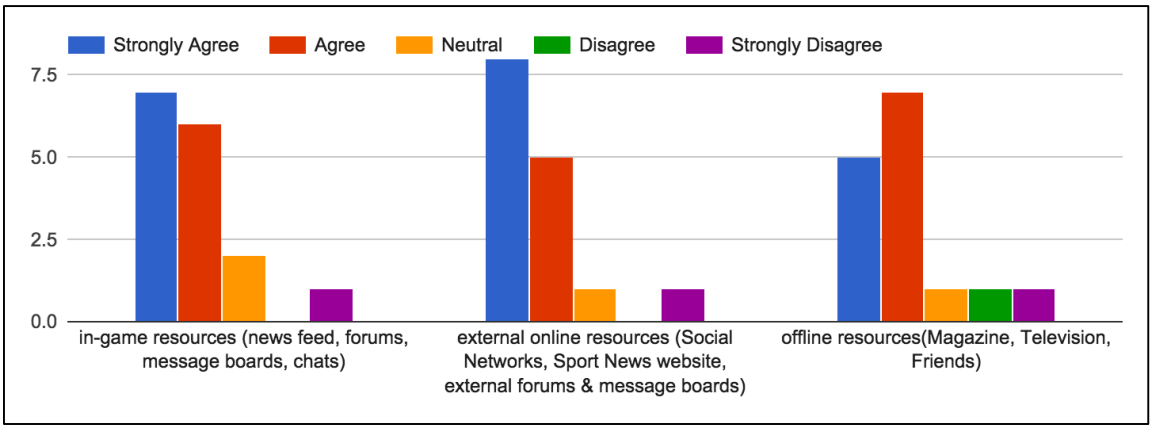

Figure 6.4 Participants' use of information resources while playing fantasy sports

6.1.3 Interest in Weather/Climate

In order to assess participants' interest in weather and climate-related news articles, a few questions were then asked related to these topics. First, participants were asked how often they followed climate and weather news articles. Figure 6.5 shows the results. Out of the 29 participants, 4 participants said that they always followed weather news articles, 10 participants responded that they often follow, 13 participants responded that they follow news articles sometimes, 1 participant said rarely and another participant said never.

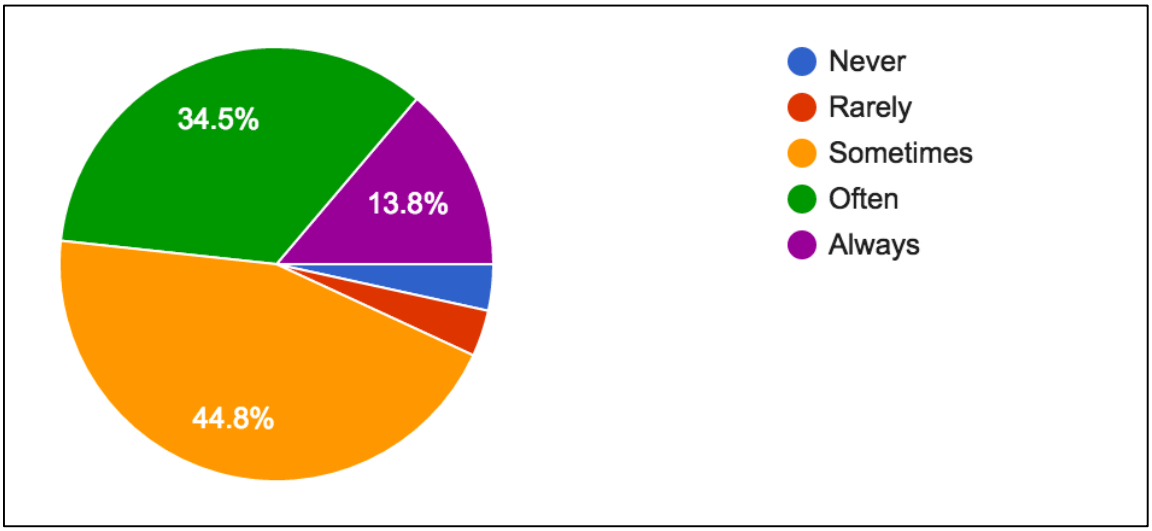

Figure 6.5 Participants' responses to how often they followed climate/weather related news articles

When asked about whether participants track weather and climate in cities other than where they reside, 55.2% of the participants responded that they do. Some of the responses from the participants' reasons to follow weather in other cities include:

- City they might travel to, or cities travelled in the past.
- Cities where relatives and friends reside.
- Weather conditions at their hometown.
- When planning for a trip.
- Curiosity.

The participants were then asked about their preferred method to read on-line news articles. The following graph (Figure 6.6) shows the participants' responses to this question. The participants who responded with the other option explained that their source was from social media and other news apps.

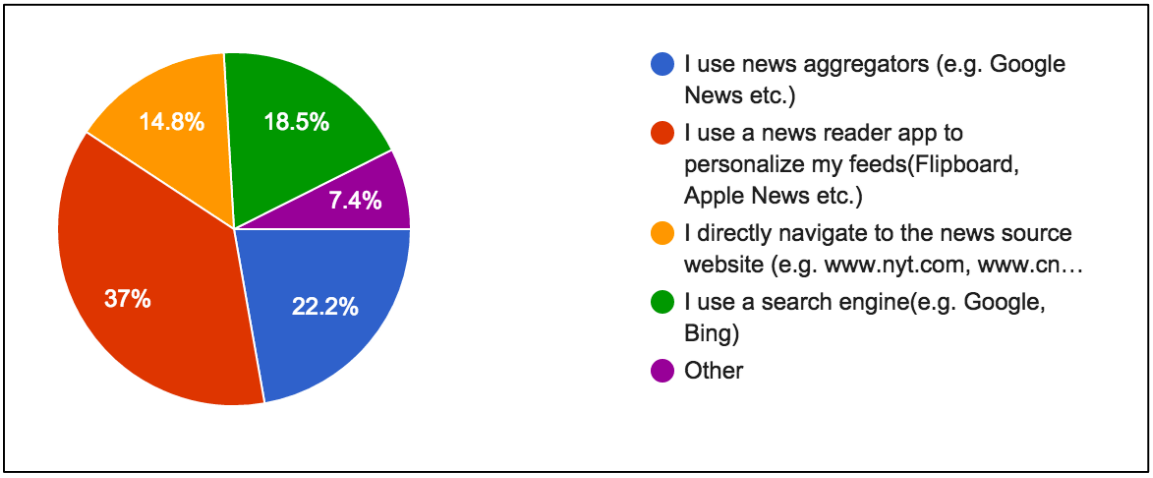

Figure 6.6 Participants' preferred method to read online news articles

6.2 Gameplay

Engagement has been a buzzword in various fields and domains across the internet. Everyone tries to measure users' interest in their online portfolio and understand the perception put forward by them. Web analytics helps understand and optimize web content by measuring, collecting and analyzing web activity data. The data collected from web analytics is used widely in internet marketing and advertising. The ideal outcome is to recognize visitor engagement and advance the understanding of the audience to design web content accordingly (Peterson and Carrabis 2008). Google Analytics is a "freemium" web analytics service which tracks and report website traffic. It is the most widely used web analytics service on the internet (W3Techs 2016). Various studies show how to use Google Analytics to measure performance and optimize or improve web content accordingly (Fang 2007, Plaza 2011).

Activity by participants during the user study was also explored via Google Analytics to measure web traffic and metrics which indicate user engagement with the system. Analytics was enabled on all the web pages present in Fantasy Climate, and click events were incorporated inside the GeoNews and NewsBoard tools. The participants were unaware of the analytics involved in the study.

The two groups in the user study had the exact same setups of the Fantasy Climate game with the only difference being the news tool. Details about the two groups are given below:

- Group 1 – Geo News – 14 participants
- Group 2 – News Board (List News) – 15 participants

Engagement is defined as an estimate of the degree and depth of visitor interaction on the site against a clearly defined set of goals (Peterson and Carrabis 2008). The following sections describe the parameters considered to gauge users' engagement in Fantasy Climate.

6.2.1 Page Views and Page Views Per Session

The Web Analytics Association Standards (2006) committee defined page views and page views per session as important metrics for web analytics (Omidvar, Mirabi et al. 2011). Page views represent the total number of pages seen by visitors to the application. It is the summation of the number of pages visited in every session connected to the system. Given the similar interaction design experienced by the two groups of participants, the number of page views is directly proportional to the engagement of users with the game.

The number of pages visited by users in group 1 (GeoNews) was 1528 whereas the number of pages visited by users in group 2 was 1151 (NewsBoard). Moreover, the number of unique page views was 861 in group 1 and 650 in group 2. This is shown in Figure 6.7.

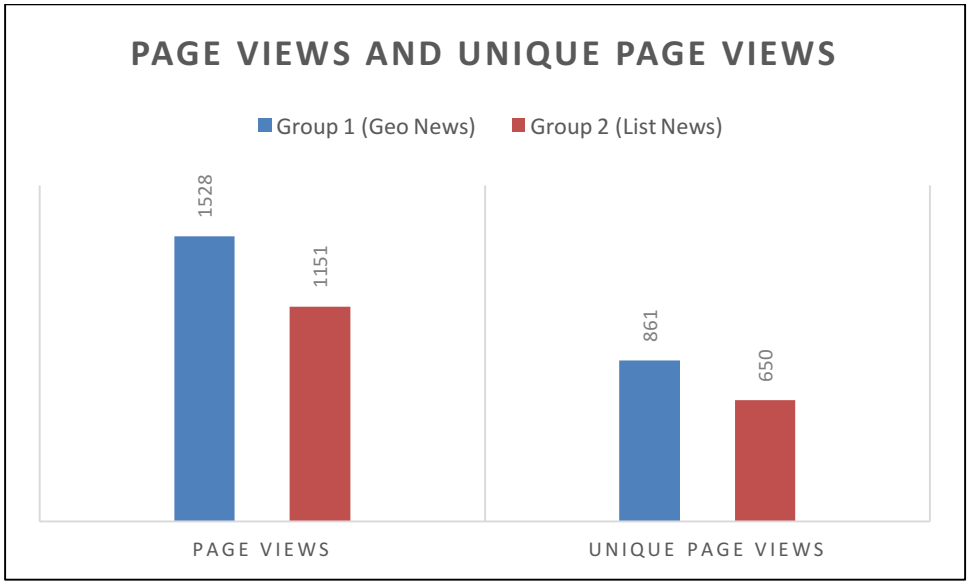

Figure 6.7 Graph depicting the number of page news and unique page views in the Fantasy Climate game during the user study

The number of sessions is another measure of engagement – measuring the number of times players visit the game. As shown in Figure 6.8, the number of sessions recorded for group 1, i.e., the group with GeoNews enabled was 170 over the entire length of the game, and the number of sessions in group 2 was 138. The number of sessions is higher for GeoNews users than NewsBoard users. But this metric by itself does not indicate the length of sessions. These values were not adjusted for the fact that the NewsBoard group had one more participant and thus should have had slightly higher number of sessions and page views than the GeoNews group assuming equivalent engagement.

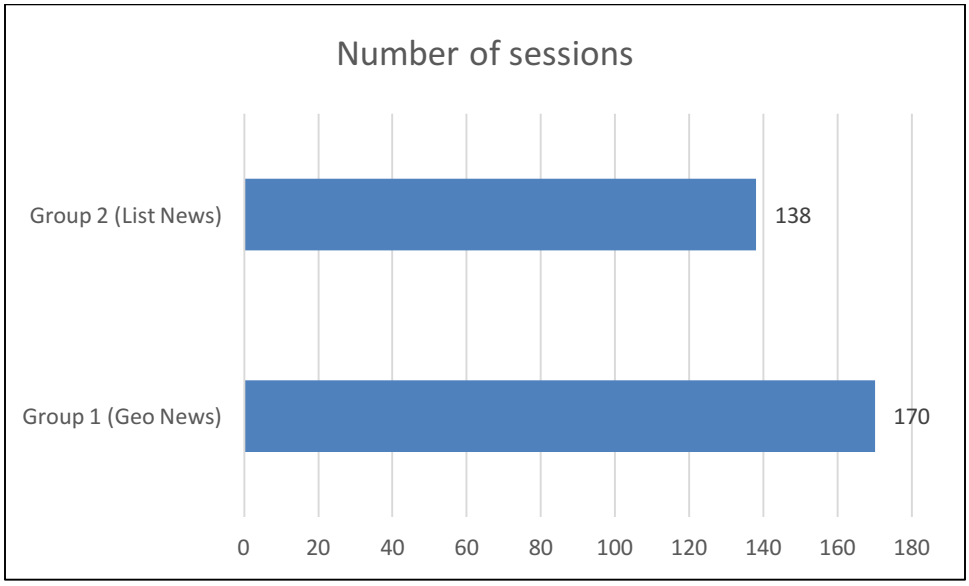

Figure 6.8 Number of sessions per group

Page views per session is also a very important metric as it provides the average number of pages visited in a session. A higher number of pages visited in a session indicated higher engagement with the system. The number of page views per session for the GeoNews users was 8.99 while the number of page views per session for the NewsBoard users was 8.34. Similarly, there were an average of 5.06 unique page views for the GeoNews users and 4.71 unique page views for NewsBoard users.

In all of these metrics, the GeoNews group was more engaged in the activity than the NewsBoard group.

6.2.2 Bounce Rate

The bounce rate is another indicator of how engaging a site is. Bounce rate is defined as the number of single page view sessions to the system (Waisberg and Kaushik

2009). It gives an indication of the quality of traffic to the game and its appeal on the visitors. The higher the bounce rate of the system, the lower is the appeal to the user and vice versa. The bounce rate for group 1 was 11.83% whereas the bounce rate of group 2 was 21.01% (Figure 6.9).

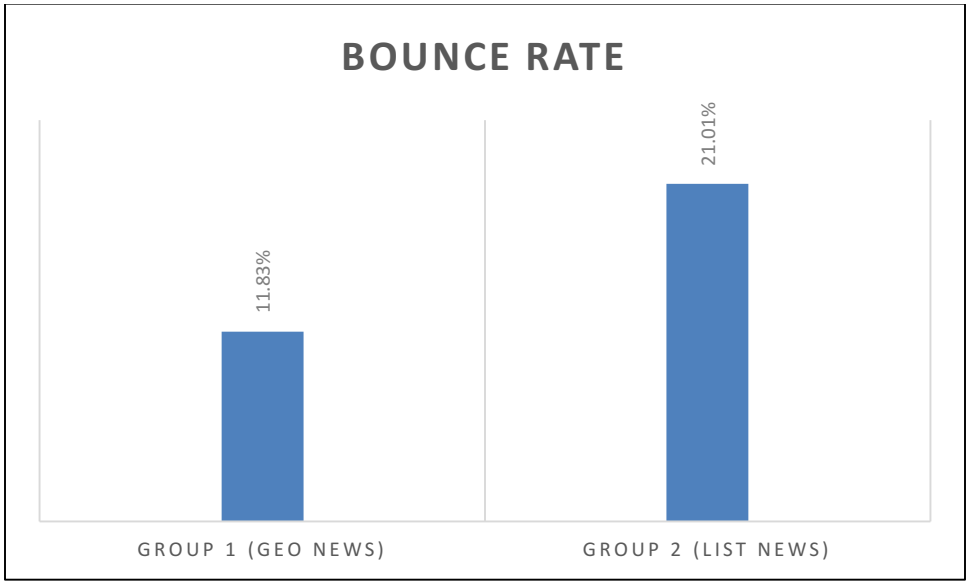

Figure 6.9 Bounce rate of the two groups in the user study

6.2.3 Average Session Duration and Time Spent Over the Entire Game

The amount of time spent by the user during a particular visit to the system is called session duration. Figure 6.10 shows the average session duration for users in group 1 (Geo News) was 9 minutes and 29 seconds, whereas for users in group 2 it was 6 minutes and 28 seconds.

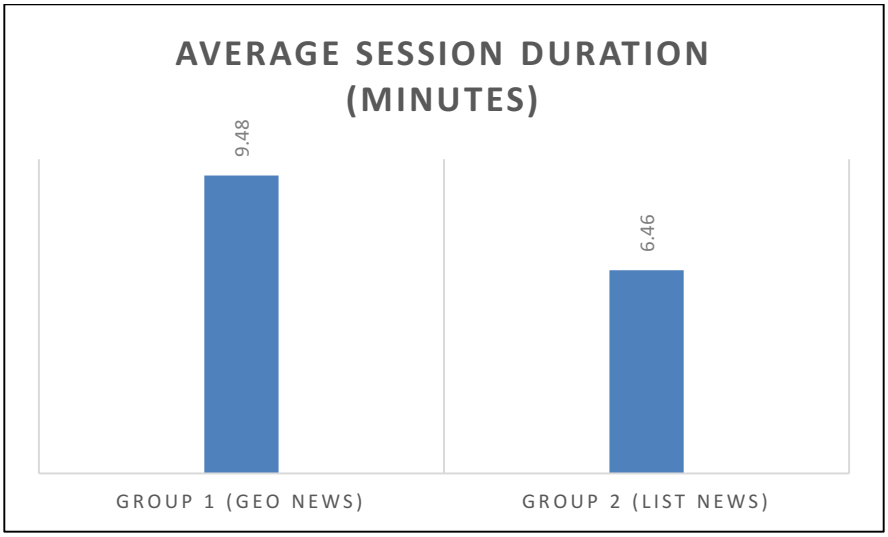

Figure 6.10 The average session duration for the two user study groups

The average overall amount of time spent on the game by each participant provides a useful metric to indicate information consumption effects of GeoNews or ListNews. The amount of time spent by a user is calculated from the total amount of time spent by the participants in a group (number of sessions * average session duration) divided by the number of users in the group. Figure 6.11 shows that the amount of time spent per participant was almost double in the GeoNews condition when compared to the NewsBoard condition.

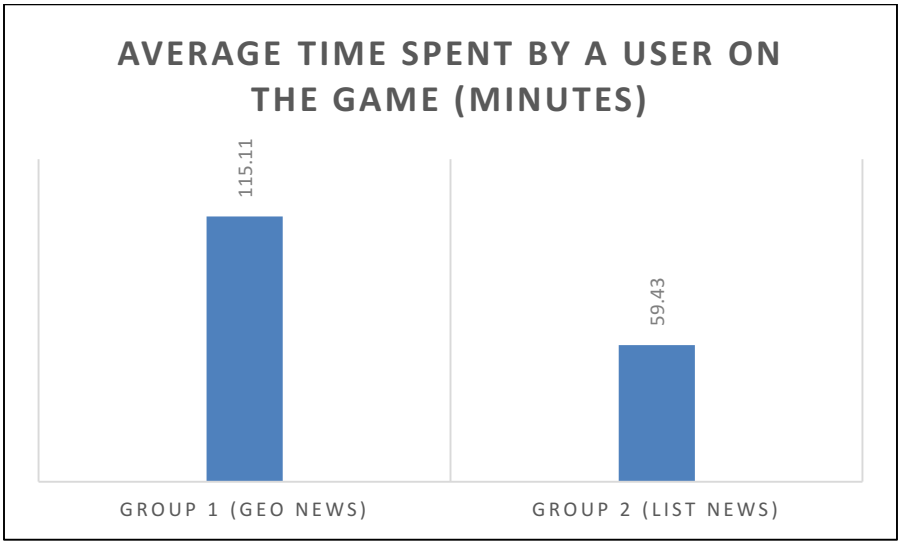

Figure 6.11 The average time spent by a user on the game for the two groups

6.2.4 Session Duration Distribution

Google Analytics session duration analysis provides the distribution of the sessions and page views based on the amount of time spent in the session.

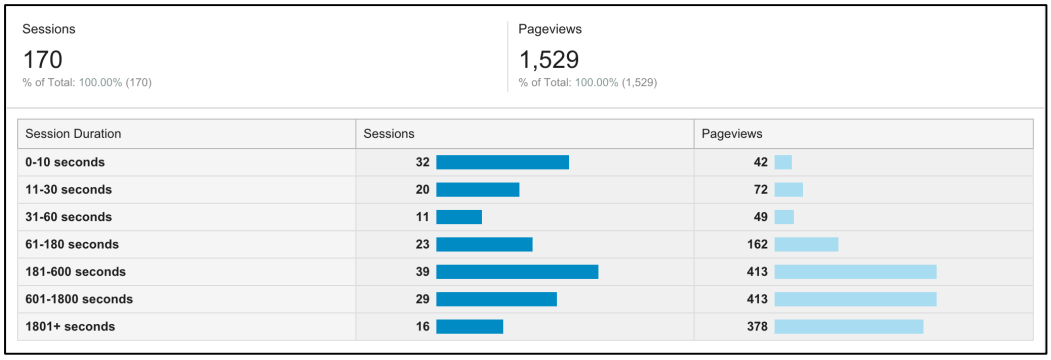

Figure 6.12 Distribution of length of sessions in terms of time for GeoNews

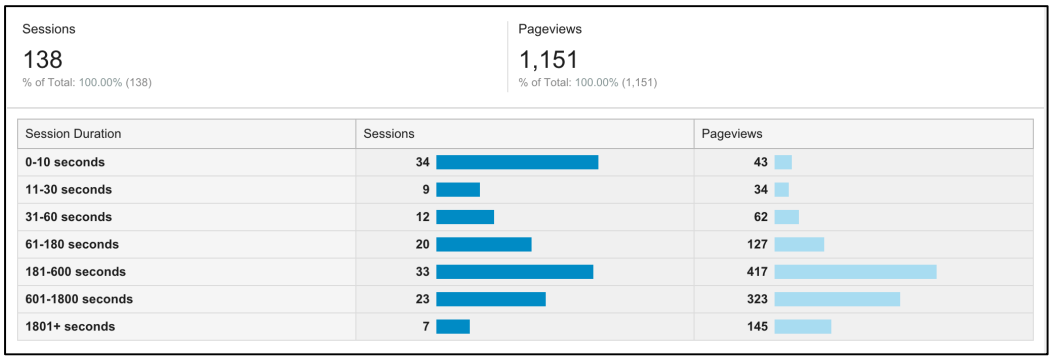

Figure 6.13 Distribution of length of sessions in terms of time for NewsBoard

Figures 6.12 and 6.13 show that the number of sessions with higher session durations are greater for GeoNews than it is for NewsBoard. The number of sessions with a session duration over 3 minutes is 84 for GeoNews and 63 for NewsBoard. As expected, the number of page views is higher in sessions with higher session durations.

6.2.5 Average Time Spent on GeoNews vs NewsBoard Over the Entire Game

The metrics so far have aggregated activity across the whole Fantasy Climate game. To get greater insight into the differences in use between GeoNews and NewsBoard, we examine the data just for the news portions of the interface.

The amount of time spent by users on the GeoNews interface and the NewsBoard interface gives a direct indication into differences in how the alternative presentations motivated news consumption. The number of page views of the GeoNews tool in group 1 was 101, and the number of page views of the NewsBoard tool in group 2 was 54. This is shown in Figure 6.14.

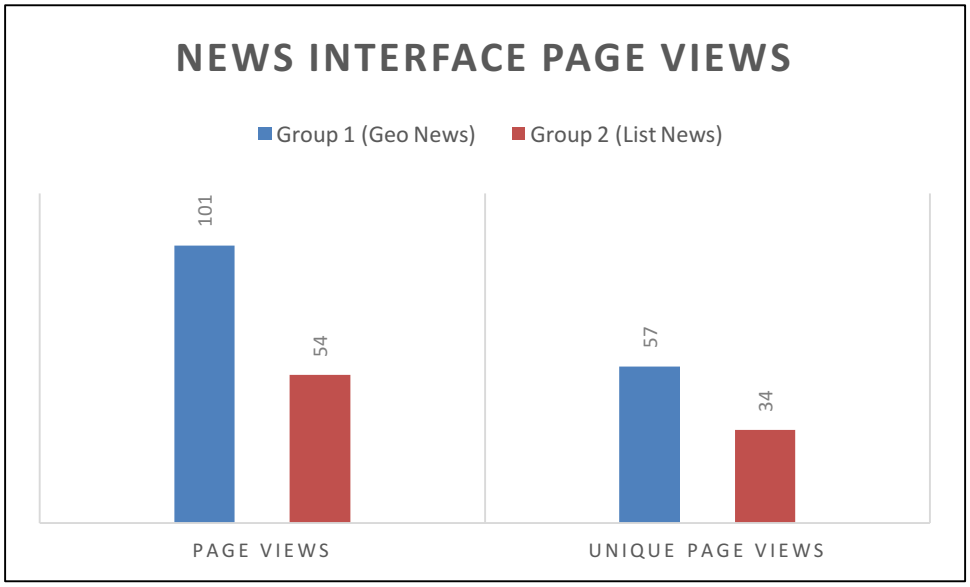

Figure 6.14 Page views and unique page views of the news interface in the two user study groups

The average time spent in the GeoNews tool and the NewsBoard tool is calculated by dividing the total time spent by all users on each tool (news related page views * time spent on each view) divided by the number of users in the group. The result indicates that the average amount of time spent by a user on the GeoNews tool was 34.12 minutes whereas the average time spent by a user on the NewsBoard tool was 11.7 minutes. The amount of time spent during each page view on GeoNews was 4 minutes and 44 seconds compared to 3 minutes and 15 seconds on the NewsBoard. See Figure 6.15.

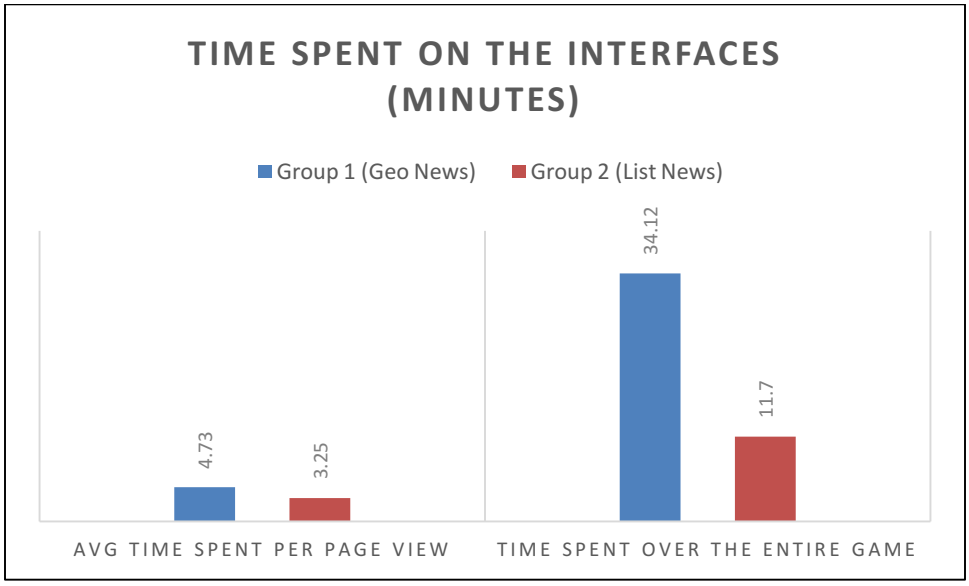

Figure 6.15 Average time spent by users on the news visualizations in the two groups

Overall, these analyses show that participants in the GeoNews tool condition had a higher level of engagement than the participants in the NewsBoard tool condition. Engagement is important for web sites maximizing traffic and attention but in our context, it is only meaningful if the increased engagement resulted in improved understanding and decision making.

6.3 Performance Evaluation

All the metrics discussed in the previous section strongly indicate that the GeoNews tool made the fantasy climate game more engaging than did the NewsBoard tool. The GeoNews tool increased the number of sessions, page views, time spent on the game etc., while decreasing the bounce rate. Moreover, the average amount of time spent

by a user on the GeoNews tool is thrice the time spent on the NewsBoard tool. This indicates that users engage with the GeoNews tool better.

But this time spent can mean two things, the first is that the users found the tool interesting and explored news articles on the interface to make better predictions, or the second is that the Geo news tool was harder to interpret and it took users longer than the NewsBoard tool to read the news articles. To find out which of the two possibilities is the actual case, the performance of the players in each group needs to be evaluated.

6.3.1 Methodology

During each stage of prediction, players had to make two predictions, one for the city whose upcoming high temperature would be the most above the city's historic average high temperature and one for the city whose upcoming low temperature would be the most below the historic average low temperature. Points in the game are awarded based on the difference between the observed temperature and historical average temperature.

At each stage of prediction, all the cities were ranked on two scales based on their value to the players. So, at each stage of prediction, each available city was awarded two ranks, the warming scale rank and the cooling scale rank. Rank 1 on the warming scale means that the city's observed temperature was the warmest compared to the historic average of the cities available for selection. Similarly, rank 1 on the cooling scale means that the city was the best available option for that selection as well, even if the city's observed temperature was actually warmer than the historic average low.

The most points for each stage of the game can thus be acquired by predicting the rank 1 cities on the warming and cooling scales. Thus, the rank of the users' prediction cities is a good measure of their performance.

6.3.2 Results

After the end of each stage, the rank of the two prediction cities of each player were determined, and the average of all these ranks for each stage for the participants in the two conditions were computed. There were 7 cities to choose from in each round so the maximum ranking was 7 and a random selection would average to a ranking of 4.0. The following table shows the average performance of each group over the three stages of prediction.

Prediction Stage	Group 1 (GeoNews)	Group 2 (NewsBoard)
Round 1	3.53	3.73
Round 2	3.5	3.96
Round 3	2.82	4.2
Overall	3.28	3.96

Table 6.1 Table depicting the performance of the users in each round of prediction in the user study

Table 6.1 shows that the participants in the GeoNews context made better predictions at every stage of the activity. Overall, the average performance of the players with GeoNews was 3.28 compared to the 3.96 average of the players with NewsBoard.

But did the availability of news help improve predictions over time? The data in Table 6.1 and graphed in Figure 6.16 shows that while the quality of the predictions by GeoNews participants improved over the three stages, the quality of the predictions of the NewsBoard participants actually got worse. Given the two groups were selecting among the same sets of cities for each round, the difficult of the selection task was identical across the conditions.

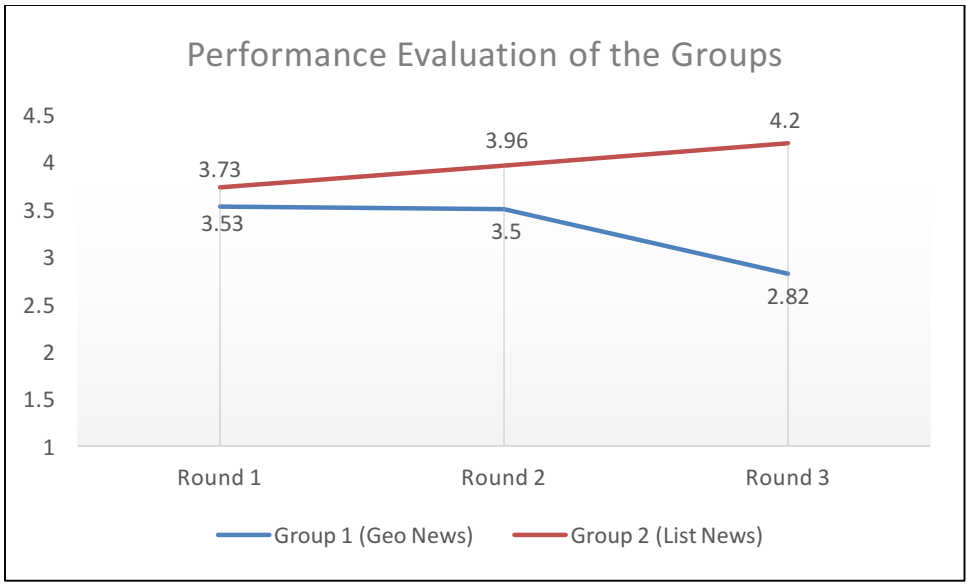

Figure 6.16 Average ranking of decisions at each stage of prediction in the game

It is evident from the above table and graph that the performance of the players in group 1 was better than the players in group 2. One interpretation of these results is that,

given the extra time spent playing the game and consuming news about weather and climate, the GeoNews participants developed a better mental model of the domain.

6.4 Post-Task Questionnaire

The post-task questionnaire was given to the participants after the completion of the game. The questionnaire consisted of questions evaluating the general features of the fantasy climate game along with focused questions concerning the news interfaces in the two groups.

Players were asked to described their strategies for predicting the warmest and coolest locations. A common answer was the usage of tools provided inside the game and analyzing data to make their predictions. Some of the responses as provided by the players are listed below:

1. "Checked the weather data provided in the application (mostly regression curves)"
2. "I would look at the list of cities mentioned and then go search in Geo news if there was any mention of any weather trends recently in those cities, and then predict according to what was being expected. "
3. "Used thermoviz and geonews to analyze historical averages and predict"
4. "I have utilized the given two tools. Mostly thermovizz."
5. "Using thermovizz and news"
6. "Accuweather predictions"

7. “Look at the news, estimate which city will have most deviation from mean hot/cold temperatures”
8. “The strategy was to use the new articles and graphs to predict a reasonable change.”

To further probe into the usage of tools in making predictions in the game, the users were asked whether they used any external sources. Six participants responded that they made use of external sources, some of their responses are listed below:

1. “Sometimes, I did check weather.com to see predicted weathers for certain cities.”
2. “Google weather”
3. “Weather channel”
4. “The first time I manually calculated the average from all the years provided. For the second two rounds I used http://www.usclimatedata.com/. Not every location has the required data but the website was pretty useful and made things easier.”
5. “Google and Wikipedia”

After learning about the resources used, the participants were asked to respond to a series of questions about their news interface. The participants were not made aware of the existence of the two interfaces.

Firstly, they were asked if the news articles in general were important to the prediction making strategy. Figure 6.17 shows the users responses on a Likert scale (1 to 5). The GeoNews users viewed news more important than did NewsBoard users.

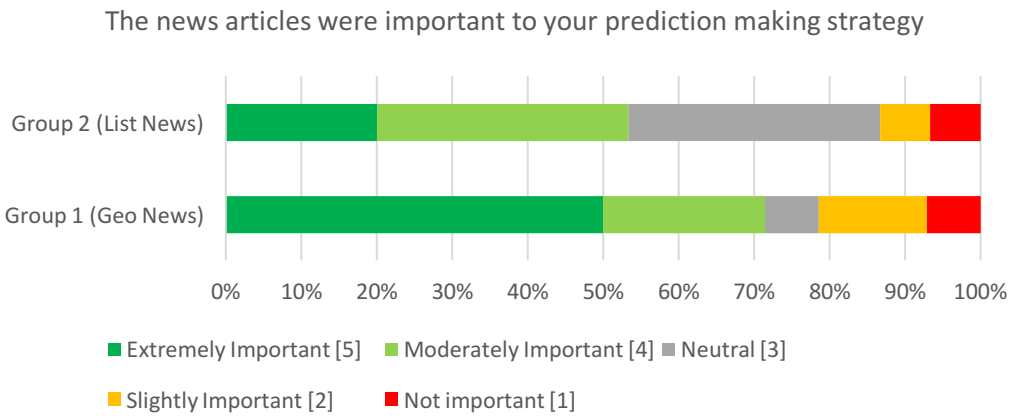

Figure 6.17 Participant assessments as to the importance of news articles to their prediction making strategy

Secondly, the participants were asked how often they read the full news article after viewing the summary of an article. As can be seen from Figure 6.18, most of the users did not read the full news article after reading the summary although more GeoNews users reported reading full articles.

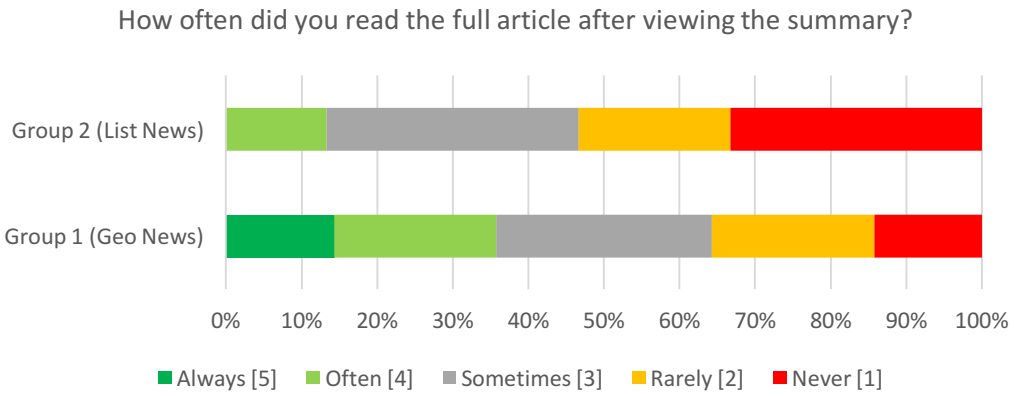

Figure 6.18 How often did users read the full article after viewing the summary

Following that were questions asking about the news interfaces. Figure 6.19 shows responses concerning whether the interfaces were intuitive and easy to navigate. Figure 6.20 shows responses to whether the interfaces were engaging and fun. The overall responses were similar, but GeoNews had an edge over NewsBoard in each.

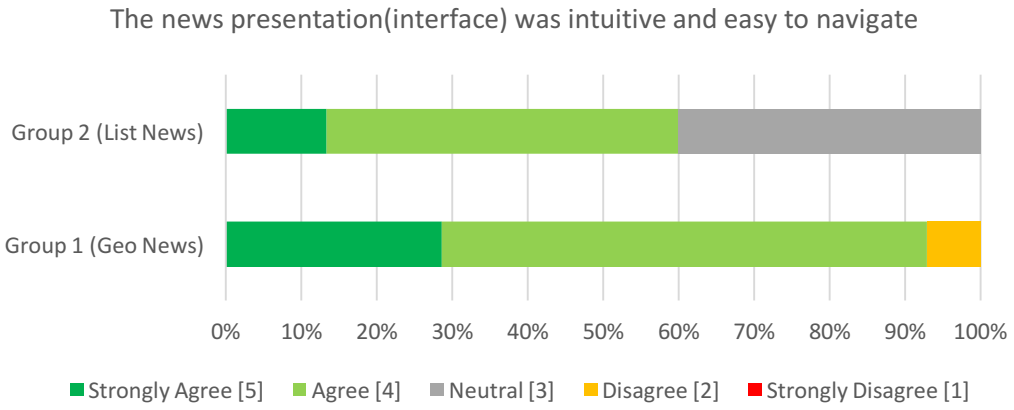

Figure 6.19 Whether the new interface was intuitive and easy to navigate

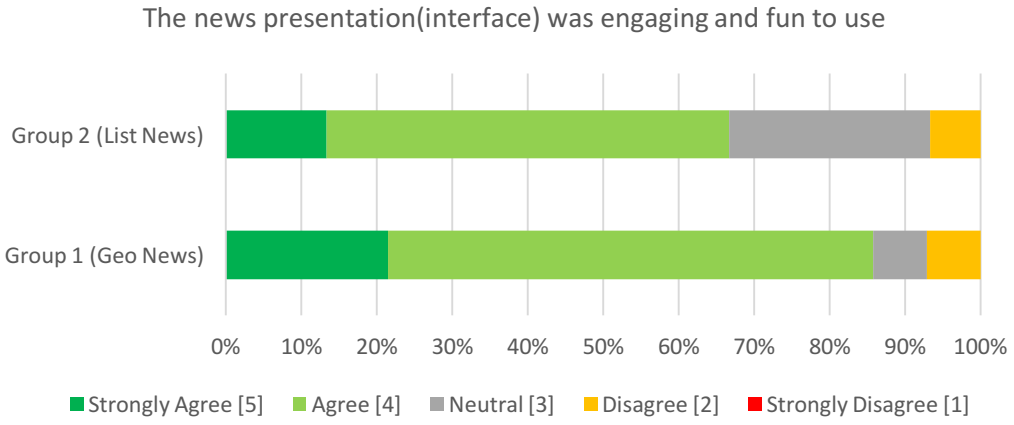

Figure 6.20 Whether the new interface was engaging and fun to use

The participants where then asked questions about the features of the news interfaces to gauge which of the features worked better in the context of exploring news articles easily. Figure 6.21 shows responses to whether the locations extracted from the articles and attached to the news in GeoNews and NewsBoard helped. Figure 6.22 shows responses to whether participants used the search bar when looking at news.

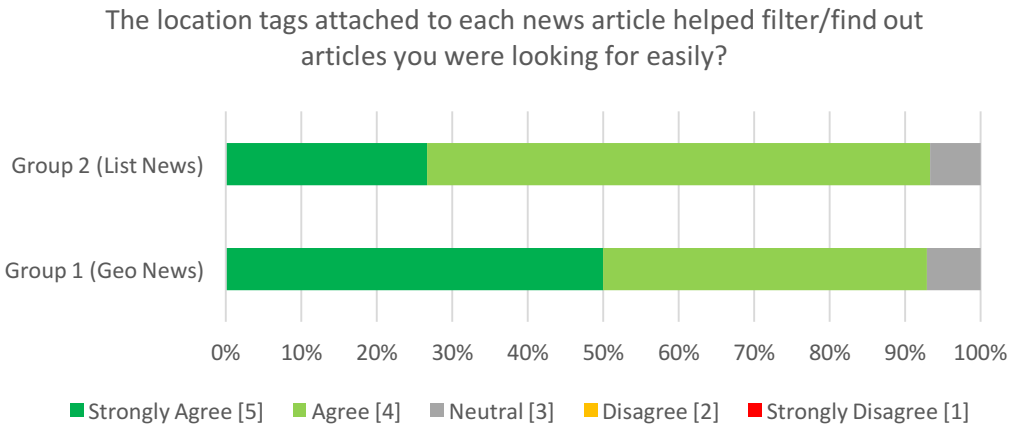

Figure 6.21 Whether the location tags helped find news articles easily

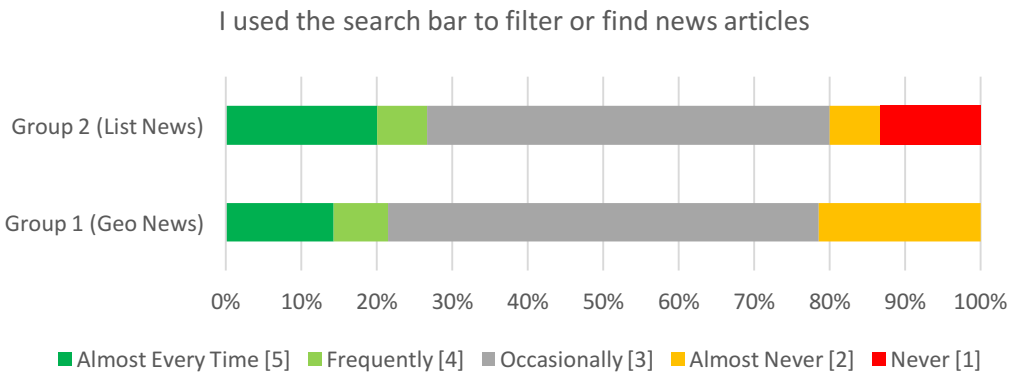

Figure 6.22 Whether the players used the search bar to find articles

Participants were also asked to list down the most and least favorite features about the news tool. Some of the responses recorded are listed below:

GeoNews most favorite features:

1. “Filter on zoom”

2. “The search fields and the option to filter news articles from either just the day, or the current week. The tag feature that gave information about the area to which the article was related to, also helped in viewing and shortlisting, sometimes even without the need for searching.”
3. “The tool was very interactive and dynamic”
4. “Filter, pan and zoom”
5. “Selecting the city to get related news articles.”
6. “The interface”
7. “Maps and news feeds were integrated together. I could select the location directly after reading the news feeds without navigating to other pages”

GeoNews least favorite features:

1. “Sometimes, the links to certain articles wouldn't work.”
2. “It became very cluttered at times that made it difficult to select locations.”
3. “Too many news articles displayed at once.”

NewsBoard most favorite features:

1. “The interface”
2. “Easy to use”
3. “The tabs. Sorting and filtering was easy.”
4. “City wise segregation.”
5. “Search”

NewsBoard least favorite features:

1. “Long summaries”

2. “Clicking on a city added its news to the list along with other cities as opposed to removing other cities from the list. This was not evident.”

The participants were also asked if they had any feedback and improvements for the news tools. The following are their responses:

GeoNews:

1. “Even though the Geo news tool was greatly helpful in making my predictions, it did take some effort to go search for cities and making choices. Probably a way to integrate could be to have the geo news search results pop up automatically when a city is chosen within the game.”
2. “Pop-ups of new articles and live feed”
3. “Make predictions directly from geonews”
4. “Have weather report YouTube videos instead of reports to read.”

NewsBoard:

1. “Better use of news reports”
2. “It could have the latest news.”
3. “Historical Data as well must be provided”
4. “Clicking on a city could bring up its related news”

Finally, participants were asked their experience and engagement with the news tool as well the whole game in general. Responses to how engrossing the game was

(Figure 6.23) and whether news made the game fun (Figure 6.24) were similar across the conditions.

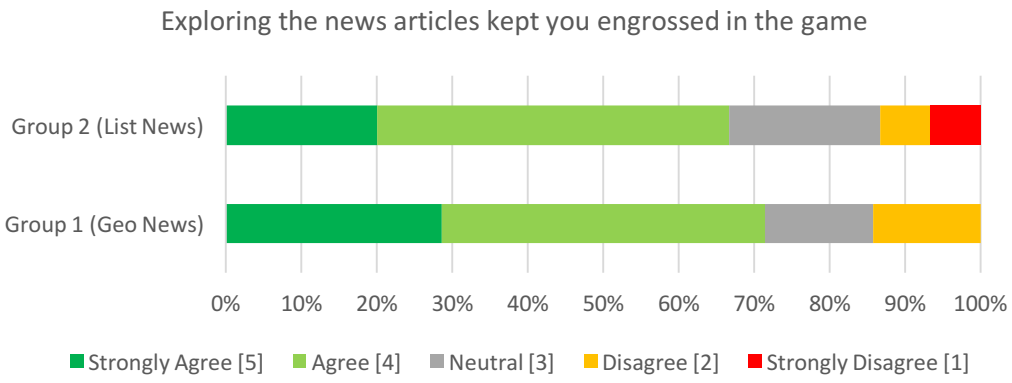

Figure 6.23 Figure depicting engrossment of players due to the news tool

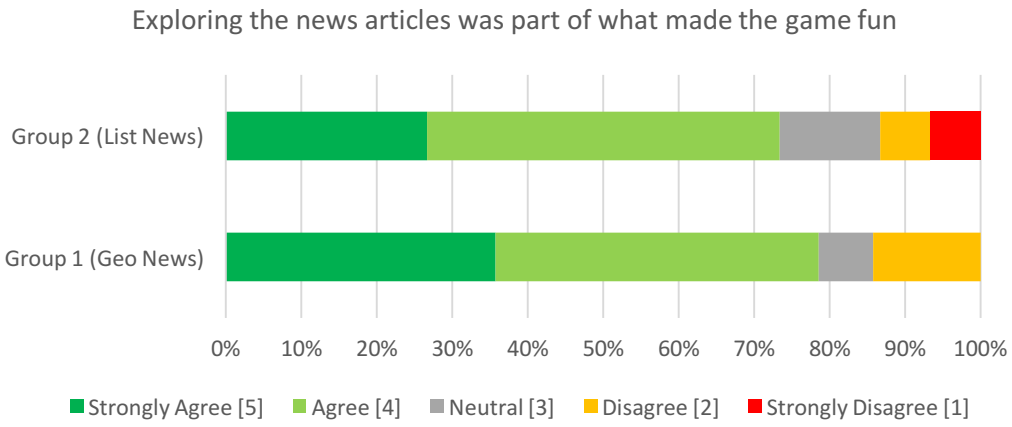

Figure 6.24 Figure depicting user engagement owing to news articles

Responses to whether the participant played the game longer than intended (Figure 6.25) and got really involved in the game (Figure 6.26) showed the increase in engagement for GeoNews users when compared to NewsBoard users that was also

observed in the web analytics data. Table 6.2 shows the mean and standard deviation data from the post-task questionnaire.

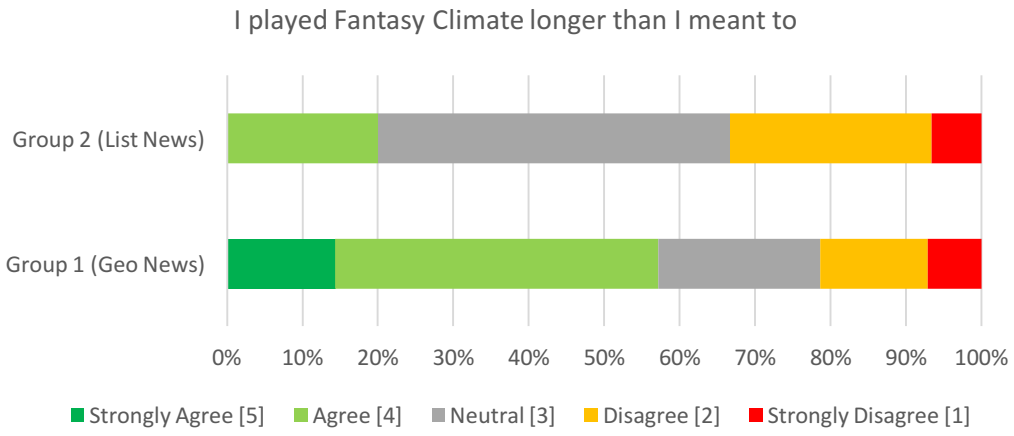

Figure 6.25 Figure depicting user engagement with the game

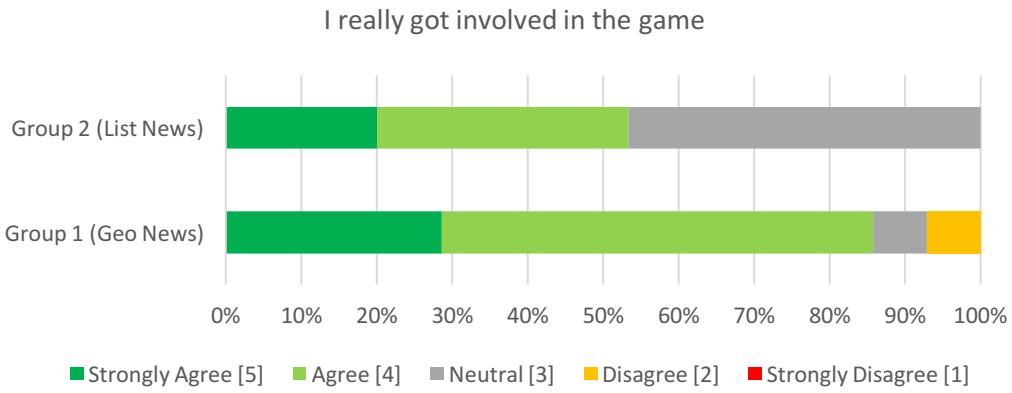

Figure 6.26 Figure depicting user involvement in the game

Question	Geo News		News Board	
	Mean	**SD**	**Mean**	**SD**
The news articles were important to your prediction making strategy	3.929	1.385	3.533	1.125
How often did you read the full article after viewing the summary?	3	1.301	2.267	1.1
The news presentation(interface) was intuitive and easy to navigate	4.143	0.77	3.733	0.704
The news presentation(interface) was engaging and fun to use	4	0.784	3.733	0.799
The location tags attached to each news article helped filter/find out articles you were looking for easily?	4.429	0.646	4.2	0.561
I used the search bar to filter or find news articles	3.143	0.949	3.133	1.246
Exploring the news articles kept you engrossed in the game	3.857	1.027	3.667	1.113
Exploring the news articles was part of what made the game fun	4	1.0378	3.8	1.146
I played Fantasy Climate longer than I meant to	3.429	1.158	2.8	0.862
I really got involved in the game	4.071	0.829	3.733	0.799

Table 6.2 Table depicting the means and standard deviations of the Likert scale responses in the post questionnaire

CHAPTER VII

CONCLUSION AND FUTURE WORK

Data driven prediction games have the potential to engage players in analyzing historical data and as a result improving domain knowledge and data interpretation skills. This thesis explores the potential for news articles as an alternate form of information in such games and explores how alternative interfaces to news affect player engagement and performance. In particular, news articles were presented in a traditional list view that included summaries (NewsBoard) or via a map view (GeoNews).

Results gathered from web analytics and the post questionnaire submitted by the players indicate that the amount of time spent was significantly higher in the GeoNews interface compared to the NewsBoard interface. The number of visits, the session durations, the overall time spent on the visualizations, the responses in the post questionnaire regarding the visualizations and the Likert scale analysis all suggest that the GeoNews tool is more engaging to users and increases their involvement in the game.

But the increased engagement developed by GeoNews does not necessarily imply that the interface is better than the list structure of displaying news. Performance assessments indicate that the additional time spent playing the game improved decisions made by the players and likely resulted in greater domain learning. GeoNews players not only consistently made better choices but their choices improved over the course of the game while NewsBoard player performance actually decreased over time.

To conclude, the hypothesis that news articles play a significant role in the prediction making strategy of the players, and that a superior news visualization leads to an increased engagement with the game, has been supported by the analyzing the web analytics data and the questionnaire responses of the user study.

The results of this study identify areas for future work. Most of this thesis was concentrated on the visualization part on the map interface, but more effort must be put into the data retrieval and the quality of news articles. The visualization is only as effective as the quality of information it displays and provides access to. In terms of interface, it would have been valuable to allow players to make predictions directly from the news interface rather than having to switch interfaces to submit predictions. Furthermore, feedback from the post-task questionnaires indicates that valuable additional features include alphabetic sorting based on city name, and reducing the number of articles displayed by default by allowing the users to control the quantity of articles they wish to see. Finally, many news events are being conveyed by social media in recent times, so incorporating social media data such as tweets or Facebook posts along with regular news articles provides additional opportunities.

The applicability of the results should transfer to other domains where there is a significant quantity of news. More abstract and less news-heavy domains, such as math, science etc., may be difficult to support as the quantity of timely news related to the current activity is likely a strong factor in how news encourages engagement and improves performance.

REFERENCES

Abramovich, G. (2014). "15 Mind-Blowing Stats About Digital Publishing." CMO. http://www.cmo.com/articles/2014/4/22/Mind_Blowing_Stats_Digital_Publishing.html. Retrieved April 26, 2016.

Atkinson, M. P., Bancilhon, F., DeWitt, D. J., Dittrich, K. R., Maier, D., & Zdonik, S. B. (1989, December). The Object-Oriented Database System Manifesto. In Deductive and Object-Oriented Databases (Vol. 89, pp. 40-57).

Barr, J. (2006). "Fantasy football adds up for students." ESPN, from http://sports.espn.go.com/espn/news/story?id=2680335/. Retrieved April 26, 2016.

Benslimane, D., Dustdar, S., & Sheth, A. (2008). Services mashups: The new generation of web applications. IEEE Internet Computing, 12(5), page 13.

Blackman, J., Aft, A., & Carpenter, C. M. (2012). Fantasyscotus: Crowdsourcing a prediction market for the supreme court. Northwestern Journal of Technology & Intellectual Property, Volume 10, Page 125.

BoilerPipe. Boilerplate Removal and Fulltext Extraction from HTML pages. From http://boilerpipe-web.appspot.com/. Retrieved April 26, 2016.

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What makes a visualization memorable? Visualization and Computer Graphics, IEEE Transactions on, 19(12), 2306-2315.

Choney, S. (2011). "Online news readership overtakes newspapers." from http://www.nbcnews.com/technology/online-news-readership-overtakes-newspapers-124383. Retrieved April 26, 2016.

Chun, R. (2015). "6 lessons academic research tells us about making data visualizations." Poynter, from http://www.poynter.org/2015/6-lessons-academic-research-tells-us-about-making-data-visualizations/381168/. Retrieved April 26, 2016.

Clark, A. (2008). "Web overtakes newspaper as source of news in US Survey." The Guardian, from http://www.theguardian.com/media/2008/dec/27/us-media-internet-newspapers-television. Retrieved April 26, 2016.

Consalvo, M. (2011, June). Using your friends: social mechanics in social games. In Proceedings of the 6th International Conference on Foundations of Digital Games (pp. 188-195). ACM.

Cordova, D. I. and M. R. Lepper (1996). "Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice." Journal of Educational Psychology 88(4): Page 715.

Davies, J. and M. Graff (2005). "Performance in e-learning: online participation and student grades." British Journal of Educational Technology 36(4): 657-663.

DeGeer, W. (2014). "What's Next in Big Data: Visualization That Works the Way the Eyes and Mind Work." Innovation Insights. Wired, from http://insights.wired.com/profiles/blogs/what-s-next-in-big-data-visualization-that-works-the-way-the-eyes.

Deng, A., Xu, Y., Kohavi, R., & Walker, T. (2013, February). "Improving the sensitivity of online controlled experiments by utilizing pre-experiment data." In Proceedings of the sixth ACM International Conference on Web search and Data Mining (pp. 123-132). ACM.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011, September). "From game design elements to gamefulness: defining gamification." In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (pp. 9-15). ACM.

EPA (2016). "Climate Change." United Stated Environmental Protection Agency, from http://www3.epa.gov/climatechange/. Retrieved April 26, 2016.

Fang, W. (2007). "Using Google Analytics for improving library website content and design: A case study." Library Philosophy and Practice 9(2): Page 22.

FantasyData. "Fantasy Data." 2016, http://fantasydata.com/. Retrieved April 26, 2016.

FantasySportsTradeAssociation (2015). "Industry Demographics." Fantasy Sports Trade Association, from http://fsta.org/research/industry-demographics/. Retrieved April 26, 2016.

Felps, P. (2000). " Fantasy sports players getting a big assist from the Internet." News on 6, from http://www.newson6.com/story/7677401/fantasy-sport-players-get-a-big-assist-from-the-internet-and-new-software. Retrieved April 26, 2016.

Fitzpatrick, J. (2010). "Newsmap Displays Local and World News as a Treemap Visualization." Life Hacker, from http://lifehacker.com/5668711/newsmap-displays-local-and-world-news-as-a-treemap-visualization. Retrieved April 26, 2016.

Fung, G. P. C., Yu, J. X., & Lam, W. (2002). "News sensitive stock trend prediction." In Advances in Knowledge Discovery and Data Mining (pp. 481-493). Springer Berlin Heidelberg.

Gargate, R. (2013). Weather Data Gamification. Computer Science & Engineering, Thesis, Texas A&M University.

Garrett, J. J. (2005). "Ajax: A new approach to web applications." University of Washington, https://courses.cs.washington.edu/courses/cse490h/07sp/readings/ajax_adaptive_path.pdf. Retrieved April 26, 2016.

Garris, R., & Ahlers, R. (2001, December). A game-based training model: Development, application, and evaluation. Paper presented at the 2001 Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL.

Garris, R., et al. (2002). "Games, motivation, and learning: A research and practice model." Simulation & Gaming 33(4): 441-467.

Golovchinsky, G. and M. H. Chignell (1997). "The newspaper as an information exploration metaphor." Information Processing & Management 33(5): 663-683.

Grissom, S., McNally, M. F., & Naps, T. (2003, June). Algorithm visualization in CS education: comparing levels of student engagement. In Proceedings of the 2003 ACM Symposium on Software Visualization (pp. 87-94). ACM.

Han, J., Haihong, E., Le, G., & Du, J. (2011, October). Survey on NoSQL Database. In Pervasive Computing and Applications (ICPCA), 2011 6th International Conference on IEEE. (pp. 363-366).

InternetLiveStats (2016). "Internet Statistics." 2016, from http://www.internetlivestats.com/. Retrieved April 26, 2016.

InternetWorldStats (2006). From http://www.internetworldstats.com/. Retrieved April 26, 2016.

Jeffries, A. (2011). "The Fitocrats: How Two Nerds Turned an Addiction to Videogames Into an Addiction to Fitness." The Observer, from http://observer.com/2011/09/the-body-hackers-behind-the-scenes-at-fitocracy-the-addictive-fitness-game-that-will-make-you-want-to-work-out/. Retrieved April 26, 2016.

Kappen, D. L., Gregory, J., Stepchenko, D., Wehbe, R. R., & Nacke, L. E. (2013, April). "Exploring Social Interaction in Co-located multiplayer games." In CHI'13 Extended Abstracts on Human Factors in Computing Systems (pp. 1119-1124). ACM.

Kittur, A., Chi, E. H., & Suh, B. (2008, April). "Crowdsourcing user studies with Mechanical Turk." In Proceedings Of The SIGCHI Conference On Human Factors In Computing Systems (pp. 453-456). ACM.

Kleinberg, A. (2011). "How To: Gamify your marketing." Mashable, Social Media. from http://mashable.com/2011/07/18/gamification-marketing. Retrieved April 26, 2016.

Kohavi, R., Deng, A., Frasca, B., Longbotham, R., Walker, T., & Xu, Y. (2012, August). "Trustworthy online controlled experiments: Five puzzling outcomes explained." In Proceedings Of The 18th ACM SIGKDD International Conference On Knowledge Discovery And Data Mining (pp. 786-794). ACM.

Kohavi, R., Deng, A., Frasca, B., Walker, T., Xu, Y., & Pohlmann, N. (2013, August). "Online controlled experiments at large scale." In Proceedings Of The 19th ACM SIGKDD International Conference On Knowledge Discovery And Data Mining (pp. 1168-1176). ACM.

Kohavi, R., Henne, R. M., & Sommerfield, D. (2007, August). "Practical guide to controlled experiments on the web: listen to your customers not to the hippo." In Proceedings Of The 13th ACM SIGKDD International Conference On Knowledge Discovery And Data Mining (pp. 959-967). ACM.

Kohavi, R. and R. Longbotham (2011). "Unexpected results in online controlled experiments." ACM SIGKDD Explorations Newsletter 12(2): 31-35.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014, June). The Stanford CoreNLP Natural Language Processing Toolkit. In ACL (System Demonstrations) (pp. 55-60).

Menno-Jan Kraak and F. Ormeling. (2010). "Cartography, Visualization of Spatial Data", Third Edition, Pearson Education Limited.

Naps, T. L., Rößling, G., Almstrum, V., Dann, W., Fleischer, R., Hundhausen, C., & Velázquez-Iturbide, J. Á. (2002, June). "Exploring the role of visualization and engagement in computer science education." In ACM SIGCSE Bulletin (Vol. 35, No. 2, pp. 131-152). ACM.

NASA (2016). "Global Climate Change - Vital signs of the planet." 2016, from http://climate.nasa.gov/. Retrieved April 26, 2016.

Newsmap (2016). from http://newsmap.jp/. Retrieved April 26, 2016.

Nielsen, J. (1990). "Hypertext and Hypermedia", Academic Press. Chapter 8.

O'Brien, H. L. and E. G. Toms (2010). "The development and evaluation of a survey to measure user engagement." Journal of the American Society for Information Science and Technology 61(1): 50-69.

Omidvar, M. A., Mirabi, V. R., & Shokry, N. (2011). "Analyzing the impact of visitors on page views with Google Analytics." arXiv preprint arXiv:1102.0735.

Paulson, L. D. (2005). "Building rich web applications with Ajax." Computer 38(10): 14-17.

Peterson, E. T. and J. Carrabis (2008). "Measuring the immeasurable: Visitor engagement." Web Analytics Demystified 14: Page 16.

PewResearchCenter (2008). "Internet overtakes newspaper as news outlet." People Press, from http://www.people-press.org/2008/12/23/internet-overtakes-newspapers-as-news-outlet/. Retrieved April 26, 2016.

Plaza, B. (2011). "Google Analytics for measuring website performance." Tourism Management 32(3): 477-481.

Pleno, I. (2011). "New study from Saatchi & Saatchi S shows majority of Americans want gamification at work and from brands." PR Newswire, from http://www.prnewswire.com/news-releases/new-study-from-saatchi--saatchi-s-shows-majority-of-americans-want-gamification-at-work-and-from-brands-123444939.html. Retrieved April 26, 2016.

Prensky, M. (2001). "True believers: Digital game-based learning in the military." Digital Game-Based Learning, Published by McGraw-Hill, Chapter 10.

Randle, Q. and R. Nyland (2008). "Participation in internet fantasy sports leagues and mass media use." Journal of Website Promotion 3(3-4): 143-152.

Rauber, A. and D. Merkl (1999). "Using self-organizing maps to organize document archives and to characterize subject matters: How to make a map tell the news of the world." Database and Expert Systems Applications, Springer.

Ricci, K. E., Salas, E., & Cannon-Bowers, J. A. (1996). "Do computer-based games facilitate knowledge acquisition and retention?". Military Psychology,8(4), p 295.

Rosenzweig, C. and M. L. Parry (1994). "Potential impact of climate change on world food supply." Nature 367(6459): 133-138.

RotoGuru (2016). "Exploring Fantasy Sports Strategy." 2016, from http://rotoguru.net/general/home.html. Retrieved April 26, 2016.

Ruihley, B. J. and R. L. Hardin (2011). "Beyond touchdowns, homeruns, and three-pointers: An examination of fantasy sport participation motivation." International Journal of Sport Management and Marketing 10(3-4): 232-256.

Samet, H., Sankaranarayanan, J., Lieberman, M.D., Adelfio, M.D., Fruin, B.C., Lotkowski, J.M., Panozzo, D., Sperling, J. and Teitler, B.E., 2014. "Reading news with maps by exploiting spatial synonyms." Communications of the ACM, 57(10), pp.64-77.

SAS (2015). Five big data challenges, SAS, https://www.sas.com/resources/asset/five-big-data-challenges-article.pdf. Retrieved April 26, 2016.

Schoenau-Fog, H. (2011). "The player engagement process–An exploration of continuation desire in digital games." Think Design Play: Digital Games Research Conference, Volume 6.

Schumaker, R. P. and H. Chen (2009). "Textual analysis of stock market prediction using breaking financial news: The AZFin text system." ACM Transactions on Information Systems (TOIS) 27(2): p 12.

Shipman, F. M. (2001). Blending the real and virtual: Activity and spectatorship in fantasy sports. In Proceedings from DAC'01: The Conference on Digital Arts and Culture. (Vol. 8, p. 2007).

Shneiderman, B. (2001). "Inventing discovery tools: Combining information visualization with data mining." Information Visualization Journal, Volume 1 Issue 1, Pages 5-12.

Simons, K. L. (1993). "New technologies in simulation games." System Dynamics Review 9(2): 135-152.

Stanley, P. "Advantages of Web Applications." 2016. Paul Stanley Software from http://www.pssuk.com/AdvantagesWebApplications.aspx. Retrieved April 26, 2016.

Statista (2014). "Statistics and facts on the U.S. Newspaper Industry." 2016. Statista from http://www.statista.com/topics/994/newspapers/. Retrieved April 26, 2016.

StatisticalBrainReserchInstitute (2015). "Total Number of Pages Indexed by Google." Statistical Brain Reserch Institute from http://www.statisticbrain.com/total-number-of-pages-indexed-by-google/. Retrieved April 26, 2016.

Straubhaar, J. (1999). "Media now: Communications media in the information age", Wadsworth Publishing, 2nd Edition.

Teitler, B. E., Lieberman, M. D., Panozzo, D., Sankaranarayanan, J., Samet, H., & Sperling, J. (2008, November). "NewsStand: A new view on news." In Proceedings Of The 16th ACM SIGSPATIAL International Conference On Advances In Geographic Information Systems (p. 18). ACM.

TextSummarization. "TextSummarization." from http://textsummarization.net/. Retrieved April 26, 2016.

The-National-Center-for-Atmospheric-Research. "What Is Climate and Climate Change?". from https://eo.ucar.edu/basics/cc_1.html. Retrieved April 26, 2016.

W3Techs (2016). "Usage of traffic analysis tools for websites." from http://w3techs.com/technologies/overview/traffic_analysis/all. Retrieved April 26, 2016.

Waisberg, D. and A. Kaushik (2009). "Web Analytics 2.0: empowering customer centricity." The Original Search Engine Marketing Journal 2(1): 5-11.

Wang, L., Wang, G., & Alexander, C. A. (2015). "Big Data and Visualization: Methods, Challenges and Technology Progress." Digital Technologies, 1(1), 33-38.

Whitehill, B. V. and B. A. McDonald (1993). "Improving Learning Persistence of Military Personnel by Enhancing Motivation in a Technical Trainiing Program." Simulation & Gaming 24(3): 294-313.

Whitlam, S. and H. Preston (1998). "Sports journalism as an information resource: a case study." New Library World 99(5): 194-204.

Woodward, D. F. (2005). "A Whole New Ballgame: How Fantasy Sports Has Evolved In The Mass Media." Communication. Arlington, University of Texas. Master of Arts.

APPENDIX 1

HEURISTIC EVALUATION QUESTIONNAIRE

Post Questionnaire

1. Describe your strategy for predicting the hottest location and the coldest location

   _____

   _____

2. Did your strategy change over time?

   - o No
   - o Yes (Explain below)

3. If yes to the previous question, please explain why and how

   _____

   _____

4. The news articles were important to your prediction making strategy (Please explain why in the next question below)

1	2	3	4	5
Not important	Slightly	Neutral	Moderately	Extremely

5. Please elaborate on your answer to the previous question

   _____

   _____

6. The news articles summaries had all the information I needed to make my prediction

1 2 3 4 5

Strongly Disagree Disagree Neutral Agree Strongly Agree

7. How often did you read the full article after viewing the summary?

   1 2 3 4 5

   Never Rarely Sometimes Often Always

8. The news presentation(interface) was intuitive and easy to navigate

   1 2 3 4 5

   Strongly Disagree Disagree Neutral Agree Strongly Agree

9. The news presentation(interface) was engaging and fun to use

   1 2 3 4 5

   Strongly Disagree Disagree Neutral Agree Strongly Agree

10. The location tags attached to each news article helped filter/find out articles you were looking for easily?

    1 2 3 4 5

    Strongly Disagree Disagree Neutral Agree Strongly Agree

11. I used the search bar to filter or find news articles

    1 2 3 4 5

    Never Almost Never Occasionally Almost every Time Frequently

12. What were your favorite features of the News tool?

    _____

    _____

13. What were your least favorite features of the News tool?

_____

_____

14. How could the news tool or its presentation be better integrated into the game?

_____

_____

15. Did you use any external (outside of the game) resource(s) to make your predictions? If (explain)

_____

_____

16. I played Fantasy Climate longer than I meant to

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

17. I really got in the game

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

18. Exploring the news articles kept you engrossed in the game

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

19. Exploring the news articles was part of what made the game fun

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

20. While playing Fantasy Climate, did you consume climate/weather related information more than usual? (explain)

_____

_____

21. What were your favorite features of Fantasy Climate?

_____

_____

22. What were your least favorite features of Fantasy Climate?

_____

_____

23. I would recommend Fantasy Climate to my friends and peers

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

24. I would like to be invited to play in future seasons of Fantasy Climate or other prediction games centered around science data

1	2	3	4	5
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

25. How can we improve Fantasy Climate?

_____

_____