The Twitter Project

INTRODUCTION

The Twitter Project is a social media analytics project that essentially amalgamates and presents data acquired through social media network(s). The objective was to gather data about pre-determined entities, store relevant information and display them through the most appropriate medium of visualization. This would allow user to scrutinize data and analyze them more specifically using Twitter User Names. The venture used Javascript as the programming language with the predominant use of Nodejs.

The project utilizes high-performance computing (HPC) resources including a small local cluster containing 2 nodes, 10 cores, 25 Gigabytes of memory and resources provided by the Janus Linux Cluster.10 Terabytes of local storage was set aside for project activities. In addition, the project will provide archiving utilizing Google Code as the hosting service to provide long-term archiving and future collaboration. Also, the service provides project wikis to allow for project management and maintenance.

Project and data management services include:

a permanent URL

secure replicated online storage location (multiple copies of data)

accurate metadata

a globally accessible repository

the option for contextual linking between data, and published research results

RethinkDB Access

The project will examine each research question through, experimental analysis of data streams. The social media sites Google+, Twitter, and Facebook all provide live data streams: In which actors on such sites public data can be programmatically extract using free/existing application programming interfaces (APIs) to provide a never-ending stream of data.

DATABASE

The first task was to establish a list of entities and twitter accounts upon which the experiment would be carried out. Therefore a list of 292 university research libraries was acquired. The list is produced below:

|  |
| --- |
| **University Research Library** |
|  |
| University of Alabama Libraries |
| (UoA) Bruno Business Library |
| (UoA) W.S. Hole Library (Special Collections) |
| (UoA) Gorgas Library |
| (UoA) Rodgers Library |
| (UoA) McLure Library |
| (UoA) Digitial Services Library |
| (UoA) Music Library |
| (UoA) Media Center |
| University at Albany; SUNY, Libraries |
| University of Alberta Libraries |
| (UoAL) Augustana Library |
| (UoAL) Cameron Library |
| (UoAL) Rutherford Library |
| University of Arizona Libraries |
| Arizona State University Libraries |
| (ASUL) Law Library |
| Auburn University Libraries |
| Boston College Libraries |
| (BCL) Law Library |
| Boston Public Library |
| Boston University Libraries |
| (BUL) Howard Gotieb Archival Research Center |
| Brigham Young University Library |
| (BYUL) Howard Hunter Law Library |
| University of British Columbia Library |
| Brown University Library |
| University at Buffalo, SUNY, Libraries |
| (UaB) Health Sciences Library |
| (UaB) Law Library |
| University of Calgary Libraries and Cultural Resources |
| (UoCLaCR) Business Library |
| (UoCLaCR) Downtown Campus Library |
| University of California, Berkeley, Library |
| (UoCBL) Bancroft Library |
| (UoCBL) ITS Library |
| University of California, Davis, Library |
| University of California, Irvine, Libraries |
| University of California, Los Angles, Library |
| (UoCLA) Arts Library |
| (UoCLA) Bio/Medical Library |
| (UoCAL) Powell Library |
| (UoCAL) Law Library |
| (UoCAL) Science and Engineering library |
| University of California, Riverside, Libraries |
| University of California, San Diego, Library |
| University of California, Santa Barbara, Libraries |
| Case Western Reserve University Libraries |
| Center for Research Libraries |
| **University of Chicago Library** |
| (UoCL) John Crerar Library |
| University of Cincinnati Library |
| University of Colorado Boulder Libraries |
| Colorado State University Libraries |
| Columbia University Libraries |
| (CUL) Avery Library |
| (CUL) Barnard Library |
| (CUL) University Libraries |
| (CUL) Science and Engineering |
| (CUL) Journalism |
| (CUL) Columbia Law Library |
| University of Connecticut Libraries |
| (UoCL) Stanford JRL Library |
| (UoCL) Magic Library |
| Cornell University Library |
| (CUL) EMPS Libraries |
| (CUL) Catherwood Library |
| (CUL) Weill Cornell Library |
| Dartmouth College Library |
|  |
| Duke University Libraries |
| (DUL) Law Library |
| Emory University Libraries |
| (EUL) Oxford Library |
| University of Florida Libraries |
| (UoFL) Library West |
| (UoFL) Education Library |
| (UoFL) College of Law Library |
| (UoFL) Marston Science Library |
| Florida State University Libraries |
| (FSUL) Goldstein Library |
| (FSUL) Law Library |
| George Washington University Library |
| Georgetown University Library |
| University of Georgia Libraries |
| (UoGL) College of Education |
| (UoGL) Zell Miller Learning Center |
| (UoGL) Russel Library |
| Georgia Tech Library and Information Center |
| University of Guelph Library |
| Harvard University Libraries |
| University of Hawai'i at Manoa Library |
| (UoH) Law School Library |
| (UoH) Manoa Library |
| (UoH) West Oahu Library |
| University of Houston Libraries |
| (UoHu) Digital Library |
| (UoHU) Law Library |
| Howard University Libraries |
| University of Illinois at Chicago Library |
| University of Illinois at Urbana-Champaign Library |
| Indiana University Libraries Bloomington |
| (IULB) Schurz Library |
| (IULB) University Library |
| (IULB) Medicine Library |
| University of Iowa Libraries |
| (UOIL) Enginerring Library |
| (UOIL) Hardin Library |
| (UOIL) Law Library |
| (UOIL) Science Library |
| Iowa State University Library |
| John Hopkins University Libraries |
| (JHUL) Welch Library |
| (JHUL) Montgomary County Campus |
| (JHUL) Milton Eisenhower Library |
| University of Kansas Libraries |
| Kent State University Libraries |
| (KSUL) Performing Arts Library |
| University of Kentucky Libraries |
| (UOK) Med Center Library |
| (UOK) W.T. Young Library |
| Bibliotheque de l' University Laval |
| Library of Congress |
| Louisiana State University Libraries |
| University of Louisville Libraries |
| McGill University Library |
| McMaster University Library |
| (McUL) Mills Library |
| (McUL) Business Library |
| (McUL) Thode Library |
| (McUL) Ron Joyce Center's Library |
| (McUL) Media Center |
| University of Manitoba Libraries |
| (UoM) Archives & Collections |
| (UoM) Health & Sciences Library |
| (UoM) Science & Technology Library |
| University of Maryland Libraries |
| (UoML) Special Collections Library (**not Working**) |
| (UoML) Performace Arts Library |
| (UoML) UMD Librarians |
| (UoML) Institute for Technology Library |
| (UoML) Engineering Library |
| (UoML) Digital/Media Library |
| University of Massachusetts Amherst Libraries |
| Massachusetts Institute of Technology Libraries |
| University of Miami Libraries |
| (UoML) Law Library |
| (UoML) Cuban Heritage Library |
| (UoML) Music Library |
| University of Michigan University |
| Michigan State University Libraries |
| (MSU) Business Library |
| University of Minnesota Libraries |
| (UOM) Business Library |
| (UOM) Wilson Library |
| (UOM) Walter Library |
| University of Missouri-Columbia Libraries |
| (UOMCL) Law Library |
| Bibliotheques del'Unversite de Montreal |
| National Agricultural Library |
| National Archives and Records Administration |
| National Library of Medicine |
| National Research Council Canada (NRCC), Knowledge Management |
| University of Nebraska-Lincoln Libraries |
| University of New Mexico Libraries |
| New York Public Library |
| New York State Library |
| New York University Libraries |
| (NYUL) Law Library |
| (NYUL) Bern Dibner Library |
| (NYUL) NYUAD Library |
| University of North Carolina at Chapel Hill Libraries |
| (UoNCaCHL) Davis Library |
| (UoNCaCHL) Health Sciences Library |
| (UoNCACHL) Undergraduate Library |
| (UoNCACHL) Scholarly Communications Library |
| North Carolina State University Libraries |
| Northwestern University Library |
| (NWUL) Archives |
| University of Notre Dame, Hesburgh Libraries |
| Ohio University Libraries |
| Ohio State University Libraries |
| (OSUL) Cartoon Library |
| (OSUL) Health Sciences Library |
| University of Oklahoma Libraries |
| Oklahoma State University Library |
| University of Oregon Library |
| University of Ottawa Library |
| University of Pennsylvania Libraries |
| (UoPL) Lippincott Library |
| (UoPL) Regan Kladstrup Special Collections |
| (UoPL) Van Pelt Library |
| (UoPL) Vet Library) |
| Pennsylvania State University Libraries |
| (PSUL) Science Library |
| (PSUL) Engineering Library |
| (PSUL) Life Sciences Library |
| (PSUL) PAMS Library |
| University of Pittsburgh Library |
| Princeton University Library |
| (PUL) Mudd Library Archives |
| Purdue University Libraries |
| (PUL) Archives Library |
| (PUL) Parrish Library |
| Queens University Library |
| Rice University Library |
| University of Rochester Libraries |
| Rutgers University Libraries |
| (RUL) Smith Library |
| University of Saskatchewan Library |
| Smithsonian Institution Libraries |
| University of South Carolina Libraries |
| University of Southern California Libraries |
| University Illinois University Carbondale Library |
| Stony Brook University, SUNY Libraries |
| Syracuse University Library |
| Temple University Libraries |
| (TUL) Health Science Library |
| (TUL) Law Library |
| (TUL) Science and Engineering Library |
| (TUL) Urban Archives |
| University of Tennessee, Knoxville, Libraries |
| University of Texas Libraries |
|  |
| (UoT) Human Rights Archive |
| Texas A&M University Libraries |
| (TA&M) Jack Williams Library |
| Texas Tech University Libraries |
| (TTUL) Law Library |
| (TTUL) Vietnam Center and Archive |
| (TTUL) Health Sciences Library |
| University of Toronto Libraries |
| (UoTL) Daniels Faculty Library |
| (UoTL) Business Information Center |
| (UoTL) First Nations House Library |
| (UoTL) Hart House Library |
| (UoTL) Law Library |
| (UoTL) Royal Ontario Museum) |
| Tulane University Library |
| (TUL) Rudolph Matas Health Services library |
| (Tul) Architecture Library |
| University of Utah Library |
| (UoUL) Mountain West Digital Library |
| Vanderbilt University Library |
| (VUL) Heard Library |
| (VUL) Bio Medical Library |
| University of Virginia Library |
| (UoVL) Fiske Fine Arts Library |
| (UoVL) Scholars Lab |
| Virginia Tech Libraries |
| University of Washington Library |
| (UoW) Law Library |
| (OoW) Odegaard Library |
| (Oow) Tacoma Library |
| Washington State University Libraries |
| Washington University in St. Louis Libraries |
| University of Waterloo |
| (UoW) Musagetes Architectural Library |
| (UoW) Writer Learning Resource Center |
| (UoW) St. Jerome Library |
| Wayne State University Libraries |
| (WSU) Shiffman Medical Library |
| (WSU) Arthur Neef Law Library |
| Western University Libraries |
| (WUL) Archives & Research Collection |
| (WUL) C.B. "Bud" Johnson Library |
| (WUL) Education Library |
| (WUL) Allyn and Betty Taylor Library |
| (WUL) D.B. Weldon Library |
| University of Wisconsin-Madison Libraries |
| (UoWM) Archives |
| (UoWM) Business |
| (UoWM) CCBC |
| (UoWM) Chemistry |
| (UoWM) College Library |
| (UoWM) Ebling Library |
| (UoWM) Journalism Reading Room |
| (UoWM) Little Magazines |
| (UoWM) Steenbock Library |
| (UoWM) Merit Library |
| (UoWM) Digitial Collections |
| (UoWM) Women's Studies Librarian |
| (UoWM) Wendt Commons |
| (UoWM) Wisconsin Water Library |
| Yale University Library |
| York University Libraries |
| (YUL) Bronfman Business Library |
| (YUL) Osgoode Hall Law Library  THE PROGRAM/ CODE  The next part of the project was to write a code which would facilitate the extraction of data from Twitter. Thus the following code was written in Javascript in order to accomplish that purpose.  var r = require('rethinkdb'),  connection;  var Twit = require('twit');  var util = require('util');  var T = new Twit({  consumer\_key:'O9ys1nNI2fWNZMWIwKBX3Q'  ,consumer\_secret:'n2bv9jstToSHNW1eBMkiXaUqwYwS2hxltFqABwhkjaw'  ,access\_token:'47155115-6bqwssOBeGPxDn9tH0OXukk0nWaxmcVgildvHPI16'  ,access\_token\_secret:'NsgCFmiQiBtgn5CEBulatuds1KrZJr8vXDZa0KXfq0'  })  r.connect( {host: 'localhost', port: 28016}, function(err, conn) {  if (err) throw err;  connection = conn;    RetrieveTwitter();  });  //Twitter Stuff  RetrieveTwitter = function() {  console.log("Tweet Stream Start");  //Tracks multiple screen names at one time.  var stream = T.stream('user', { track:'NAME' } && { track:'NAME' })  stream.on('tweet', function (tweet) {  //r.db('Twitter').table('Twitt').run(connection)  r.db('Twitter').table("Twitt").insert({  username : tweet.user.name,  userid: tweet.user.id,  fid: tweet.id,  data: tweet.created\_at,  text: tweet.text,  coordinates:tweet.coordinates  }).run(connection)  });  } |

This code in Javascript allows one to extract data about certain and relevant terms or user names in Twitter. Thus by replacing the word “NAME” in the code and replacing it with any other user name in Twitter will allow one to track all the tweets and retweets made about a certain user in Twitter.

RETHINKDB

The subsequent part of this project was to create a database to store the accumulated data. The Janus Server was used, as the base to store data, and the Rethinkdb was the designated platform. Tables were created to hold data captured about all the libraries involved in the experiment. The results produced are available on this website: <http://research-cs.org:8081/>

RUNNING THE PROGRAM WITH PM2

Running the libraries and the codes was the part that followed. The following intructions were used to run the programs that were created:

Start an app using all CPUs available + set a name :

pm2 start app.js -i max --name "api"

Restart the previous app launched, by name :

pm2 restart api

Stop the app :

pm2 stop api

Restart the app that is stopped :

pm2 restart api

Remove the app from the process list :

pm2 delete api

Kill daemon pm2 :

pm2 kill

Update pm2 :

npm install pm2@latest -g ; pm2 updatePM2

Deployment help:

pm2 deploy help

start [options] <file|json|stdin>

start and daemonize an app

deploy <file|environment>

deploy your json

startOrRestart <json>

start or restart JSON file

startOrReload <json>

start or gracefully reload JSON file

startOrGracefulReload <json>

start or gracefully reload JSON file

stop [options] <id|name|all|json|stdin>

stop a process (to start it again, do pm2 restart <app>)

restart [options] <id|name|all|json|stdin>

restart a process

reload <name|all>

reload processes (note that its for app using HTTP/HTTPS)

gracefulReload <name|all>

gracefully reload a process. Send a "shutdown" message to close all connections.

delete <name|id|script|all|json|stdin>

stop and delete a process from pm2 process list

sendSignal <signal> <pm2\_id|name>

send a system signal to the target process

ping

ping pm2 daemon - if not up it will launch it

updatePM2

update in-memory PM2 with local PM2

update

(alias) update in-memory PM2 with local PM2

interact [secret\_key] [public\_key] [machine\_name]

agent actions for keymetrics.io

killInteract

stop agent

infoInteract

get information about agent

web

launch an health API on port 9615

dump

dump all processes for resurrecting them later

save

(alias) dump all processes for resurrecting them later

resurrect

resurrect previously dumped processes

startup <platform>

auto resurrect process at startup. [platform] = ubuntu, centos, gentoo or systemd

generate

generate an ecosystem.json configuration file

ecosystem

generate an ecosystem.json configuration file

reset <name|id|all>

reset counters for process

describe <id>

describe all parameters of a process id

desc <id>

describe all parameters of a process id

list

list all processes

ls

(alias) list all processes

l

(alias) list all processes

status

(alias) list all processes

jlist

list all processes in JSON format

prettylist

print json in a prettified JSON

monit

launch termcaps monitoring

m

(alias) launch termcaps monitoring

flush

flush logs

reloadLogs

reload all logs

logs [id|name]

stream logs file. Default stream all logs

ilogs

advanced interface to display logs

kill

kill daemon

The command pm2 resurrect will resume the programs that may have crashed or stopped subsequently.

THE ANALYSIS AND VISUALIZATION

This final stage involves the representation, analysis and visualization of data. The Information Flow Networks and Friend-Follower networks were used to gain insight into how and why users interact. A highly used method for visualizing the network is the *force-directed layout* which will be used in the implementation of D3.

*GenerateDataTrend* defined in the class *ExtractDatasetTrend will be used to analyze the velocity of tweets and retweets. And finally TweetXplorer will be used to analyze the data gathered through twitter.*