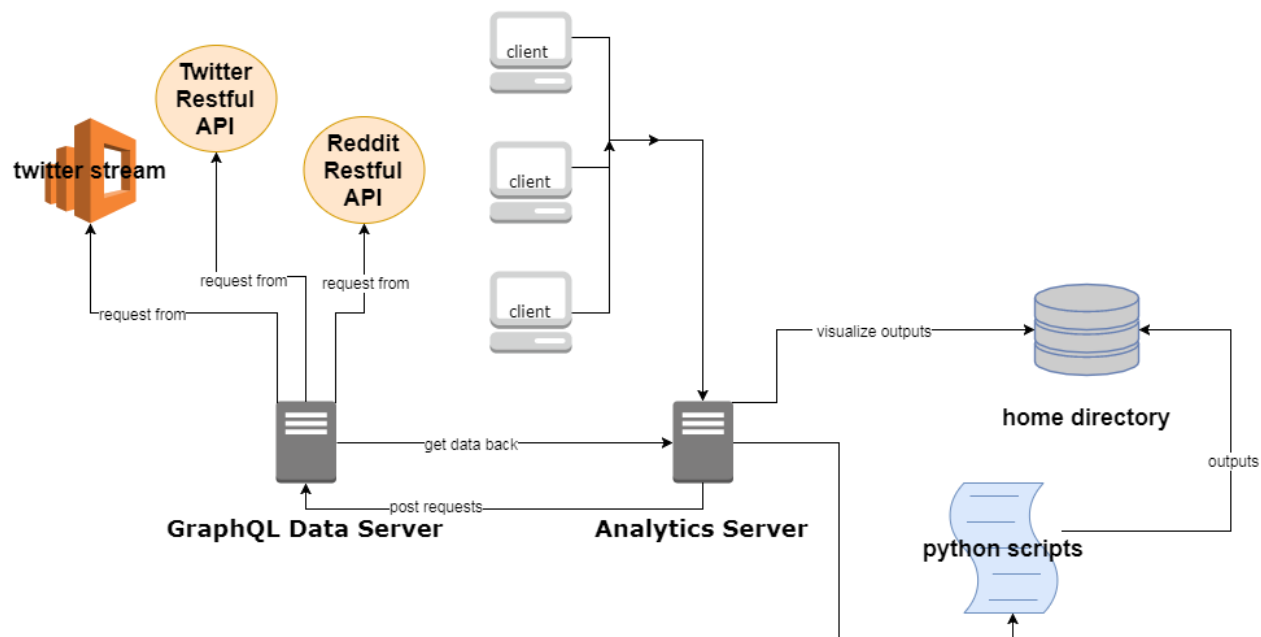


Memo V1.0: Dependencies and Libraries for Analytic Tools

Author: Chen Wang (Tech Services)

Mailto: cwang138@illinois.edu

Architecture



Code Structures

./package.json → npm start will invoke this package.json in which concurrently boost up the data server as well as the analytics server

./app.js → main of the analytics app

./public → bootstrap css, customized javascript code on client side, logos and images

./scripts → python scripts that do the computing and analysis of the social media data collected

./routes → analytics app's routes

./views → analytics app's webpage templates

./graphql → this directory hosts the GraphQL data server code

./graphql/app.js → main for the data server

./graphql/API → how graphql will actually interact with social media platforms and get data

./graphql/data/schema → graphql data schema

./graphql/bin/www → invoke the graphql server

Bash script to install all the dependency libraries

- ✓ Make sure you have a **python3**, and it's your default python.

Test it by type `$ python -V` anywhere in your command-line tool, see what version it is.

Linux comes with a default python 2.7, to overcome this, do: `$ alias python=python3` in your command-line tool.

- ✓ Make sure you have a working version (>v6) of NodeJS and its path is in your environment.

Test it by type `$ node --version`

- ✓ Run the bash script in the root directory by typing `$./install_libraries.sh`

This bash script calls another python script "ntlk_download.py" to install the nltk corpus

Manually install python and its dependency library

- ✓ Install python on Ubuntu:

`$ sudo apt-get update`

`$ sudo apt-get install python3.6`

- ✓ Check if pip is installed:

`$ command -v pip3`

If not installed, do:

`$ sudo apt-get install python3-setuptools`

`$ sudo easy_install3 pip`

- ~~✓ Add libraries via site-packages~~

~~Site-packages.zip contains all the dependency libraries (including a few minor custom change)~~

~~Unzip site-package.zip, and merge it to python's Lib/site-package directory. Usually this path depends on where you installed your python. Ask me if you cannot find it (cwang138@illinois.edu)~~

- ~~✓ Alternatively, install libraries via pip:~~

~~Install pip if it's not installed:~~

~~<https://pip.pypa.io/en/latest/installing/>~~

~~`curl https://bootstrap.pypa.io/get-pip.py -o get-pip.py`~~

~~`python get-pip.py`~~

- ✓ Use Pip to install dependency libraries:

pip3 install sklearn

pip3 install pyLDAvis

pip3 install ...

- ✓ List of dependency libraries used in the scripts

Command	Description	Custom change
pip3 install numpy scipy matplotlib pandas	Scipy stack for data science	no
pip3 install pyldavis (disabled temporarily)	Visualize LDA topic modeling	yes
pip3 install scikit-learn	Machine learning packages	no
pip3 install plotly	Create visualization in HTML format	no
pip3 install python-dotenv	Read environment variables from .env	no
pip3 install beautifulsoup4	Beautifulsoup is for parsing html	no
pip3 install nltk python -m nltk.downloader all.	Natural language processing stacks	no
(disabled temporarily) pip3 install git+https://github.com/erikavaris/tokenizer.git	Reddit tokenizer	yes
Pip3 install twython	Twitter specified	no
Pip3 install networkx	Network visualization	no

Manually install NodeJs and Analytics app libraries (NodeJS server 1)

- ✓ Install nodejs and npm (package manager)

\$ sudo apt-get install nodejs

\$ sudo apt-get install npm

\$ node -v

\$ npm -v

- ✓ package.json in the root directory is the dependencies for analytics app

under root directory, \$ npm install

✓ list of dependencies and its details

Libraries	Description
adm-zip	ADM-ZIP is a pure JavaScript implementation for zip data compression for NodeJS
body-parser	Parse incoming request bodies in a middleware before your handlers, available under the req.body property.
cookie-parser	Parse Cookie header and populate req.cookies with an object keyed by the cookie names.
csv-string	Parse and Stringify for CSV strings in the same fashion as JSON parser
debug	A tiny node.js debugging utility modelled after node core's debugging technique.
dotenv	Dotenv is a zero-dependency module that loads environment variables from a .env file into process.env.
express	Fast, unopinionated, minimalist web framework for node.
express-session	Create a session middleware in express framework.
fs	Nodejs's filesystem
jsonexport	This module makes easy to convert JSON to CSV and its very customizable.
morgan	HTTP request logger middleware for node.js
node-fetch	A light-weight module that brings window.fetch to Node.js
nodemailer	Send e-mails from Node.js
nodemon	For use during development of a node.js based application.
oauth	A simple oauth API for node.js .
path	Nodejs's path system
pug	Pug is a high performance template engine heavily influenced by Haml and implemented with JavaScript for Node.js and browsers.
python-shell	A simple way to run Python scripts from Node.js with basic but efficient inter-process communication and better error handling.

serve-favicon	Node.js middleware for serving a favicon.
---------------	---

Data Server (NodeJS server 2)

- ✓ Install nodejs and npm (package manager)

Same as above

- ✓ package.json in the “ ./graphql” directory is the dependencies for the data server

cd into “graphql” directory, \$ npm install

- ✓ list of dependencies and its details

Libraries	Description
aws-es	A Node.js binding for Amazon's Elasticsearch service which requires Amazon's Signature Version 4.
aws-sdk	AWS SDK for JavaScript
bluebird	Full featured Promises/A+ implementation with exceptionally good performance
body-parser	Parse incoming request bodies in a middleware before your handlers, available under the req.body property.
cookie-parser	Parse Cookie header and populate req.cookies with an object keyed by the cookie names.
cors	Node.js CORS middleware
dataloader	A data loading utility to reduce requests to a backend via batching and caching.
debug	A tiny node.js debugging utility modelled after node core's debugging technique.
dotenv	Dotenv is a zero-dependency module that loads environment variables from a .env file into process.env.
elasticsearch	The official low-level Elasticsearch client for Node.js and the browser.
express	Fast, unopinionated, minimalist web framework for node.
express-graphql	Production ready GraphQL HTTP middleware.
graphql	A Query Language and Runtime which can target any service.

http-aws-es	Use the elasticsearch-js client with Amazon ES
morgan	HTTP request logger middleware for node.js
node-fetch	A light-weight module that brings window.fetch to Node.js
nodemailer	Send e-mails from Node.js
nodemon	For use during development of a node.js based application.
oauth	A simple oauth API for node.js .
path	Nodejs's path system
pug	Pug is a high performance template engine heavily influenced by Haml and implemented with JavaScript for Node.js and browsers.
request	A simple way to run Python scripts from Node.js with basic but efficient inter-process communication and better error handling.
serve-favicon	Node.js middleware for serving a favicon.
snoowrap	A JavaScript wrapper for the reddit API
twitter	Twitter API client library for node.js
when	A lightweight Promises/A+ and when() implementation, plus other async goodies.