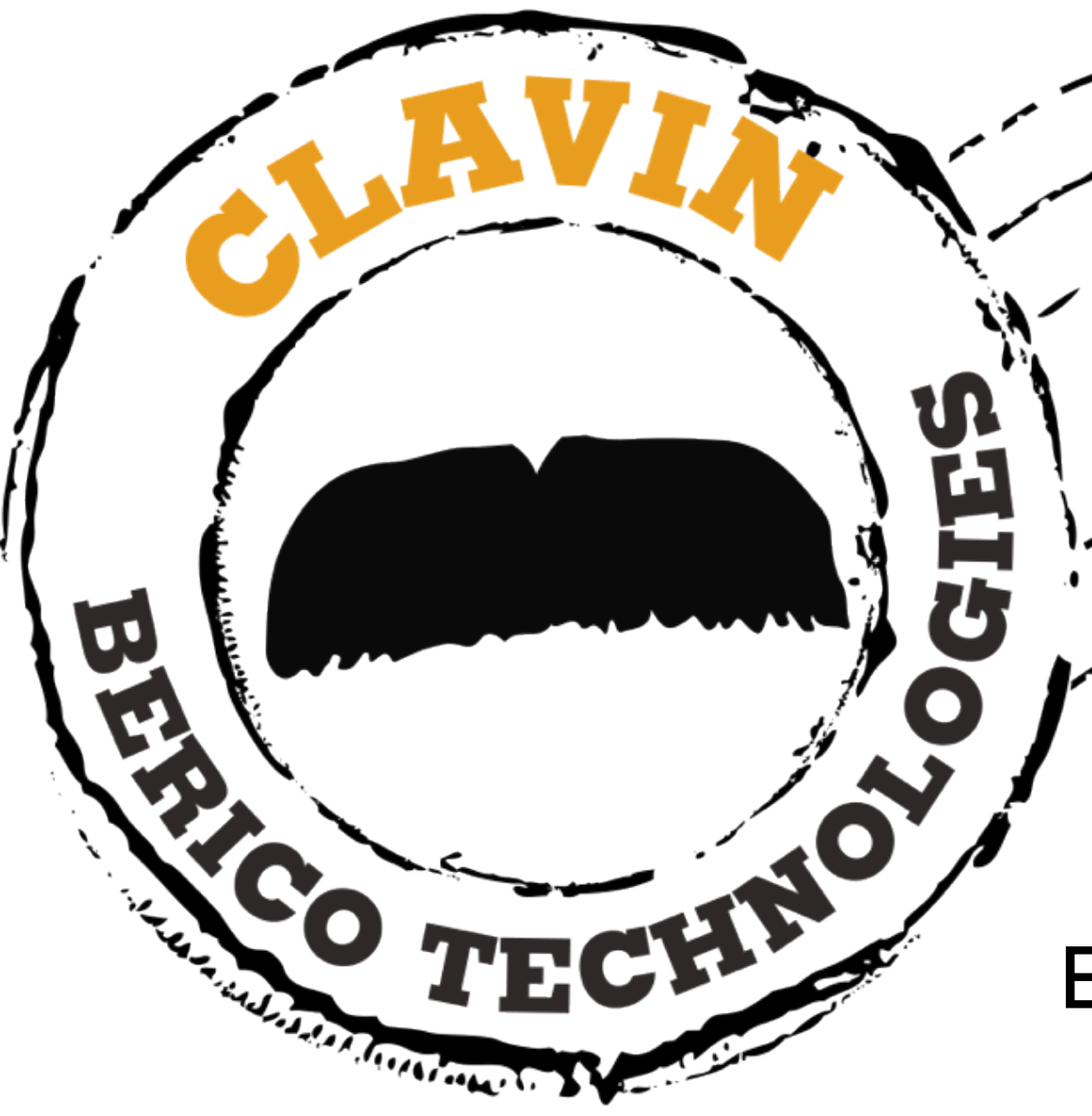


Enabling GEOINT and Cyber Security

Travis Pinney
Berico Technologies
@tlpinney

Mil-OSS LANT2 – 2013-Aug-08

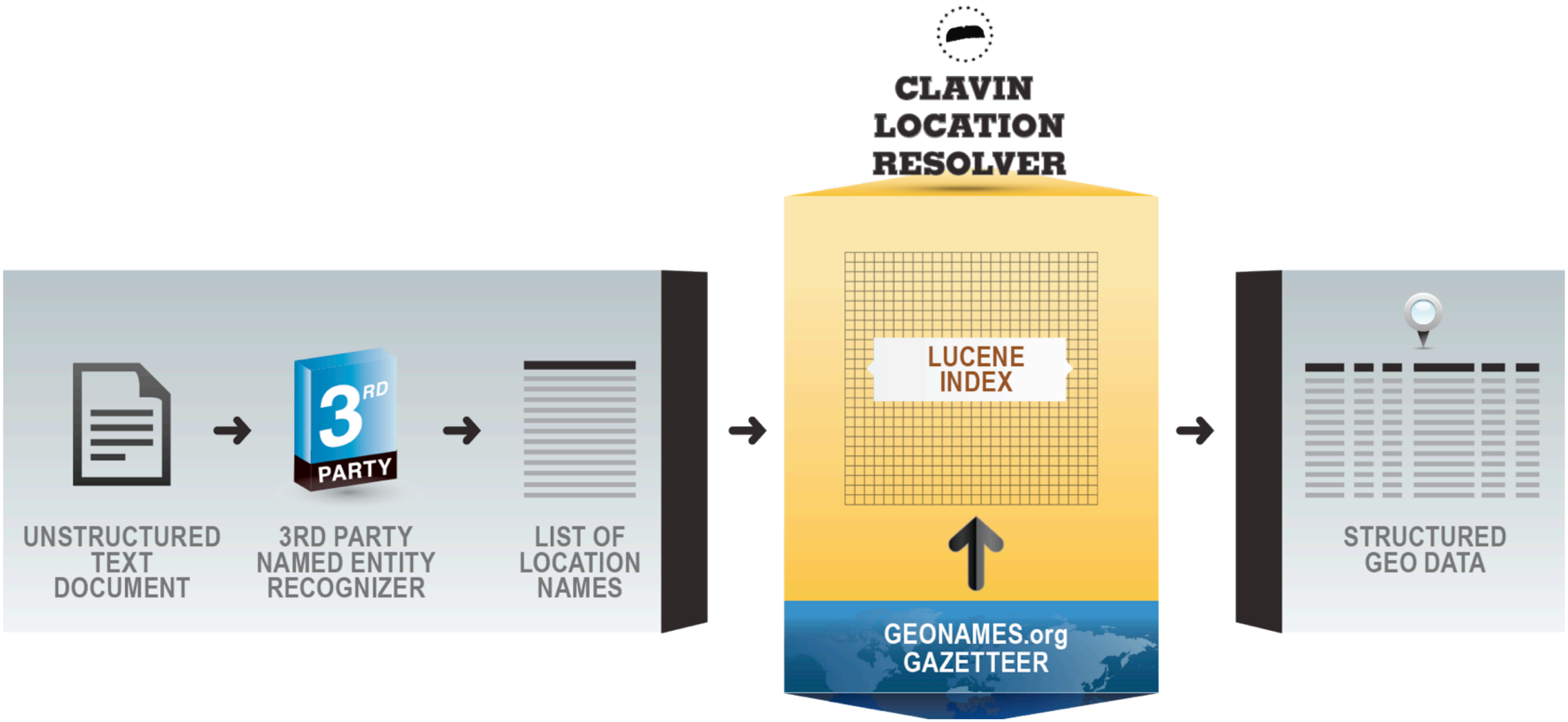
What is CLAVIN is NOT?

- A cyber security solution

What is CLAVIN?

- Cartographic Location And Vicinity INdexer
- Geoparser
- Extracts location names
- Resolves geospatial entities
- Open source
- Runs on Hadoop

Under the hood



CLAVIN handles

- Ambiguous references
- Typos & phonetic spellings
- "Ivory Coast" == "Côte d'Ivoire"

Demo Time

Only the top 20 locations are shown.

Locations Parsed and Resolved From Text

ID	Name	Lat, Lon	Country Code	#
6446345	US	49.1, 1.96667	FR	9
298795	Turkey	39.05901, 34.91155	TR	9
2215636	Libya	28, 17	LY	8
88319	Benghazi	32.11667, 20.06667	LY	7
2215636	Libyan	28, 17	LY	5
3199389	Herzegovina	43, 17.83333	BA	1
1149361	Afghanistan	33, 66	AF	1
529468	Southeast Europe	43.5333, 3.9833	FR	1
3319	Bengazi	32.11667, 20.06667	LY	1
0630	Cairo	30.06263, 31.24967	EG	1
5455014	American	33.52813, -105.74332	US	1
831053	Kosovo	42.58333, 21	XK	1
3277605	Bosnia	44.25, 17.83333	BA	1
685608	Balkans	44, 23	RO	1
6783140	Gadaffi	12.43269, 14.24894	CM	1
783754	Albania	41, 20	AL	1
718075	Macedonia			

CLAVIN stats

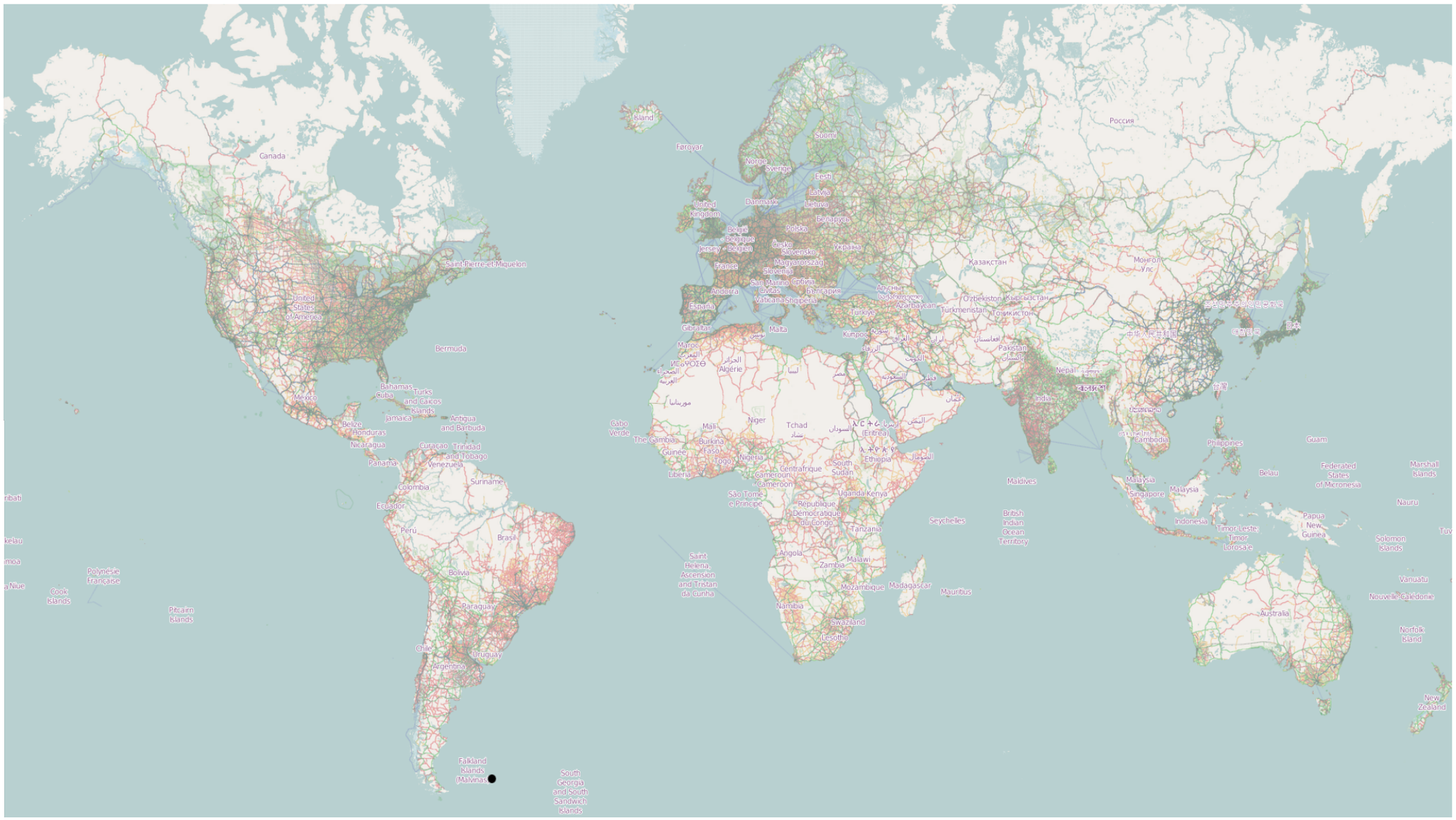
- **Accurate**: 0.75
- **Fast**: 100 locations per second per CPU
- **Scalable**: processes 1 million documents in under 1 hour on a 9-node Hadoop cluster
- **Open Source:** Apache License on GitHub

CLAVIN Extensions

- Custom Geo Gazetteers (OpenStreetMap)
- IP Address Gazetteer Integration
- GeoJson Support
- Log Extractors

OpenStreetMap

OpenStreetMap



OpenStreetMap

- Node
- Ways
- Relations

OpenStreetMap

- Node

  place=*

  name=*

IP Address Gazetteer

MaxMind GeoLite2 Database

Parsing Strategies

GeoHash

IpAddress

LatLon

Dms (Degrees Minutes Seconds)

Parsing Strategies

GeoHash

IpAddress

LatLon

Dms (Degrees Minutes Seconds)

Parsing Strategies

```
GeoParserFactory.DefaultCoordinateParsingStrategies
                                .add(new IpAddressParsingStrategy());

    // Get a parser instance.
    GeoParser parser =
GeoParserFactory.getDefault("IndexDirectory/");

// Extract and resolve IP Addresses and coordinates from the
text below.

    ResolutionContext results =
                parser.parse(“This is a test  41.222.183.44");
```

Parsing Strategies

41.222.183.44 (ipaddress) was extracted as -6.8, 39.2833.

41.222.183.44 (ipaddress) was resolved as Golden Tulip Dar Es Salaam, Tanzania.

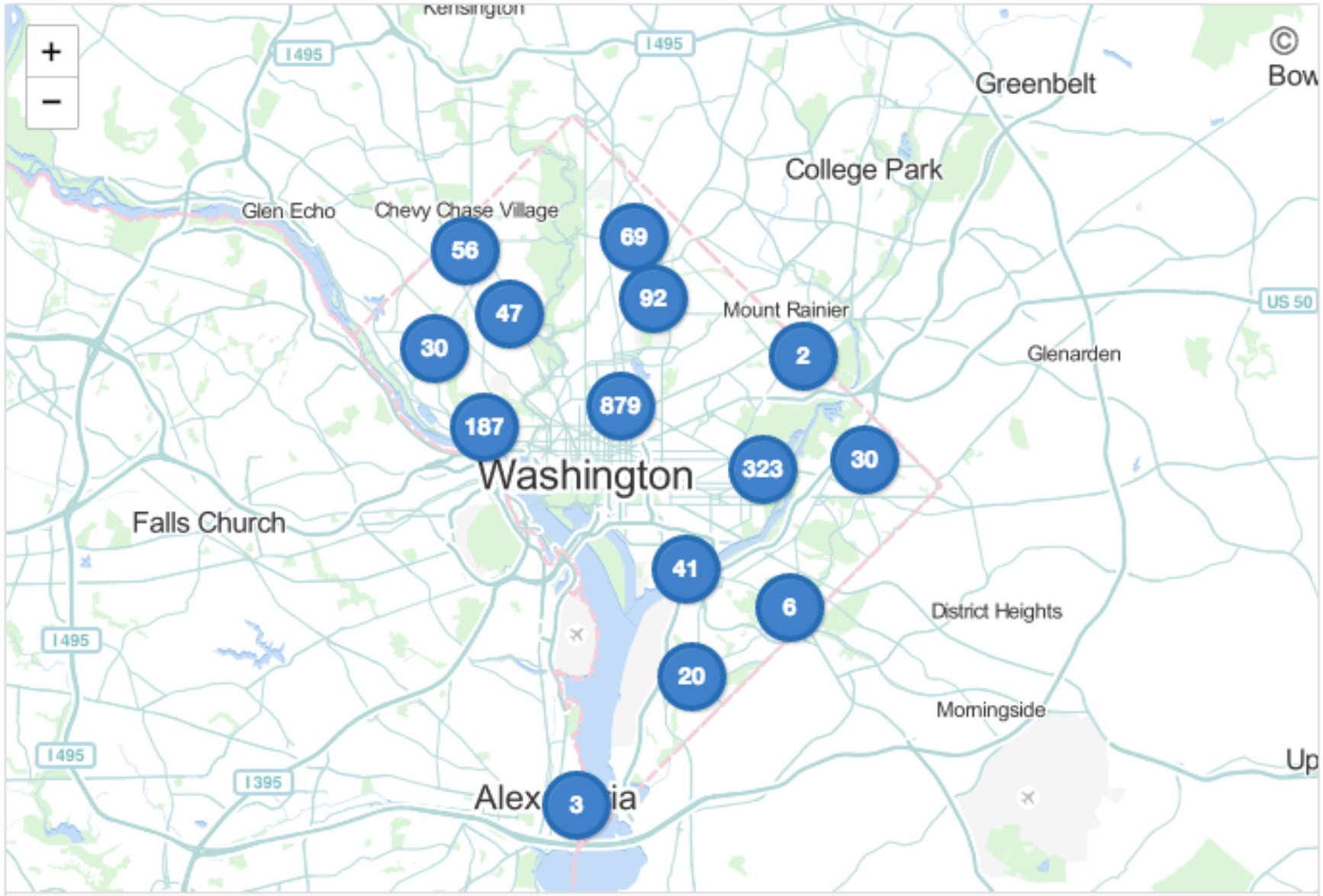
GeoJSON

geojson.org

1.0

https://github.com/Leaflet/Leaflet

Github


Kensington
I 495
I 495
Greenbelt
College Park
Glen Echo
Chevy Chase Village
56
69
47
92
30
Mount Rainier
2
US 50
Glenarden
879
187
Washington
323
30
Falls Church
41
6
District Heights
I 495
20
Morningside
I 495
I 395
3
Rendered with ❤ by GitHub

Netstat

Active Connections

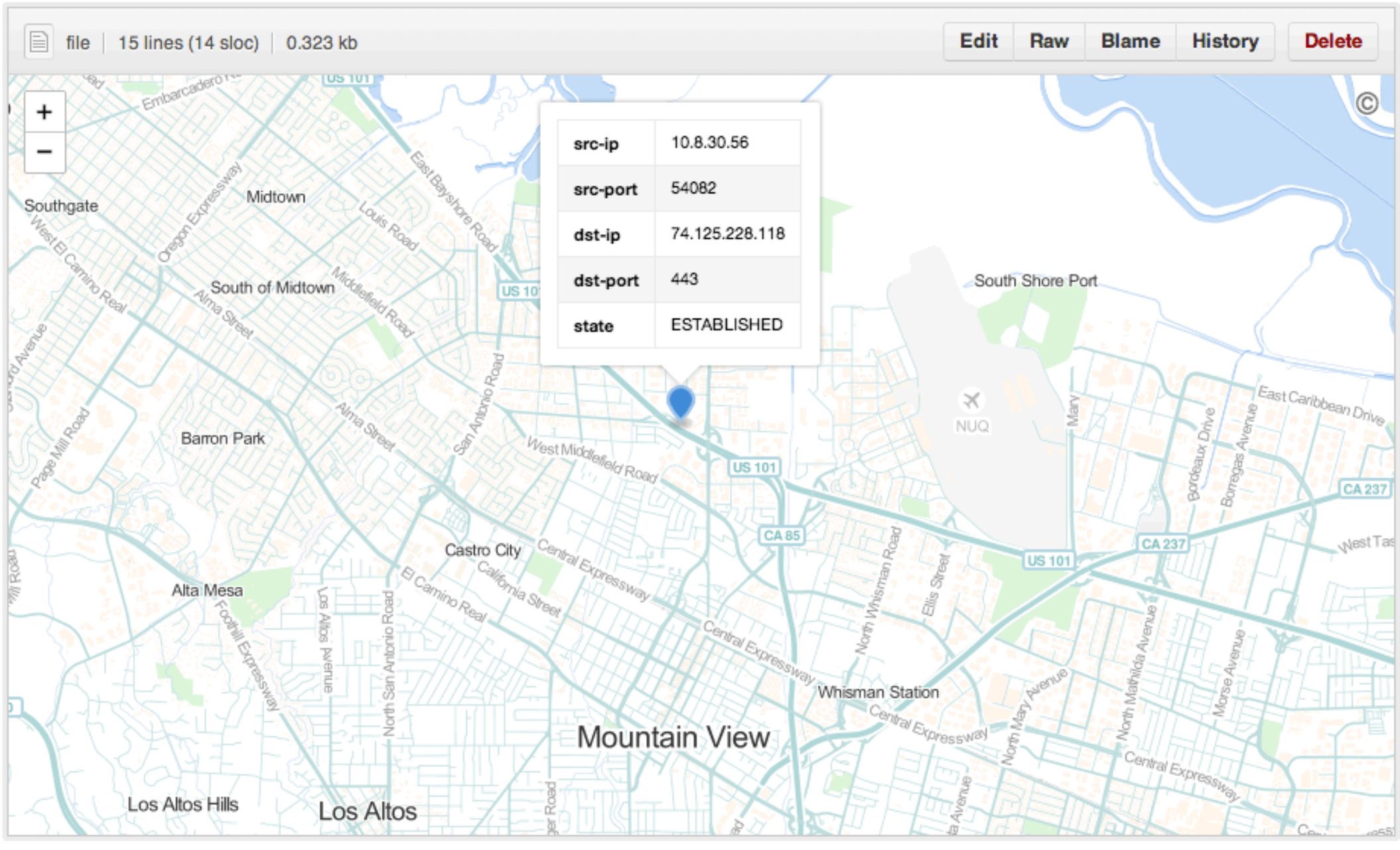
Proto	Local Address	Foreign Address	State
TCP	10.8.30.56:54082	74.125.228.118:443	ESTABLISHED

Netstat Feature

```
var feature = {
    "type": "Feature",
    "geometry": {
        "type": "Point",
        "coordinates": [-122.0813, 37.4139]
    }
    "properties": {
        "src-ip” : "10.8.30.56",
        "src-port" : "54082”,
        "dst-ip” : "74.125.228.118",
        "dst-port" : "443",
        ”state” : ”ESTABLISHED"
     },

};
```

Netstat Feature



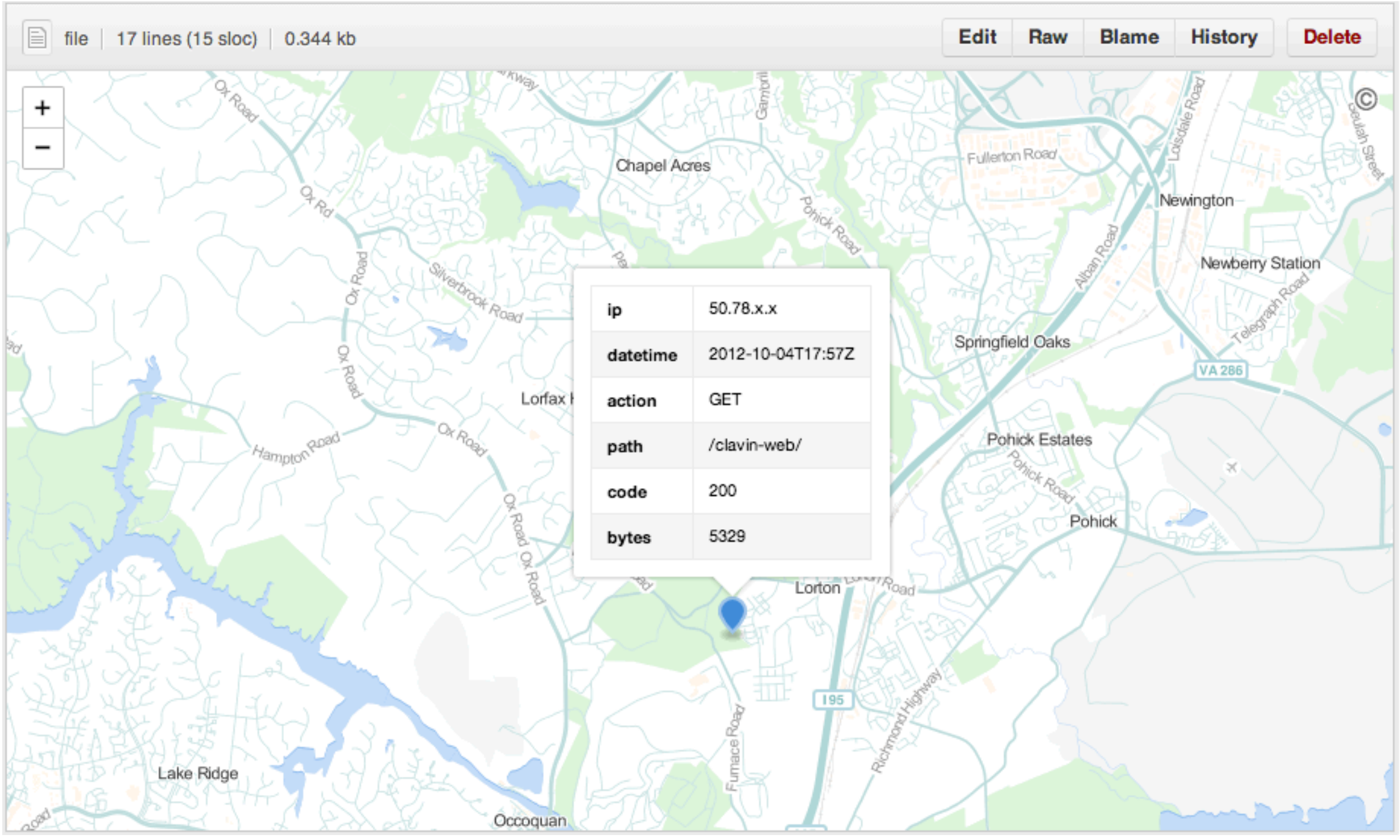
Tomcat Logs

50.78.x.x - - [04/Oct/2012:17:57:17 +0000] "GET /clavin-web/ HTTP/1.1" 200 5329

Tomcat Log Feature

```
var feature = {
    "type": "Feature",
"geometry": {
        "type": "Point",
        "coordinates": [122.0828, 37.3861]
    },
"properties": {
        "ip" : "50.78.X.X",
        "datetime" : ”2012-10-04T17:57Z"
        "action" : ”GET",
        "path" : "/clavin-web/",
        "code" : ”200”,
        ”bytes” : “5329”
    }
};
```

Tomcat Log Feature



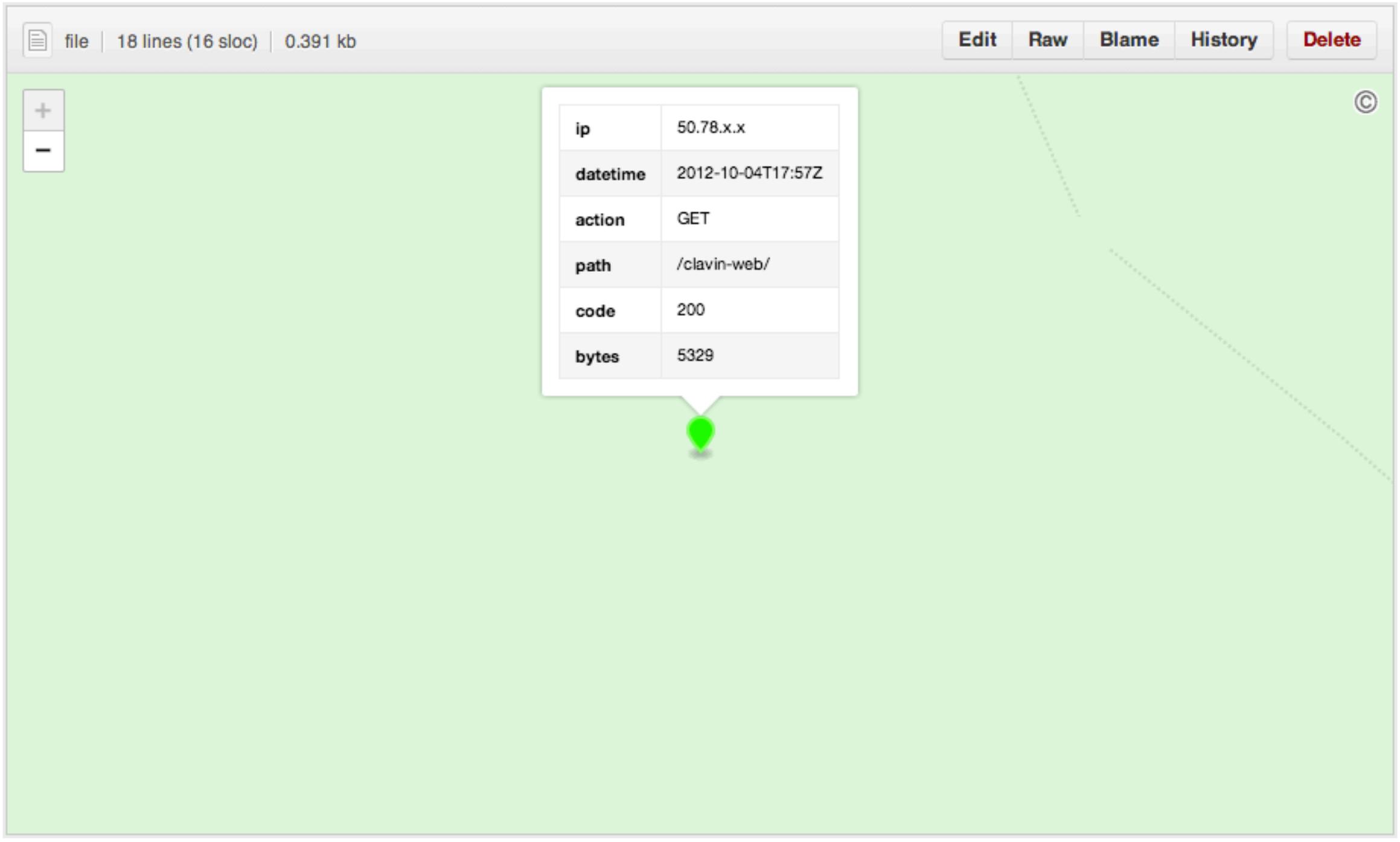
Styling

- https://help.github.com/articles/mapping-geojson-files-on-github#styling-markers

Feature Color

```
var feature = {
    "type": "Feature",
    "properties": {
        "ip" : "50.78.x.x",
        "datetime" : ”2012-10-04T17:57Z"
        "action" : ”GET",
        "path" : "/clavin-web/",
        ”state” : ”ESTABLISHED”,
       “marker-color”: “#00FF00”
     },
    "geometry": {
        "type": "Point",
        "coordinates": [122.0828, 37.3861]
    }
};
```

Feature Color



Tools

GEOJSONLINT.COM

Network

Firewall Logs

HTTP Logs

Pcap

Netflow

Social Media

Twitter

Facebook

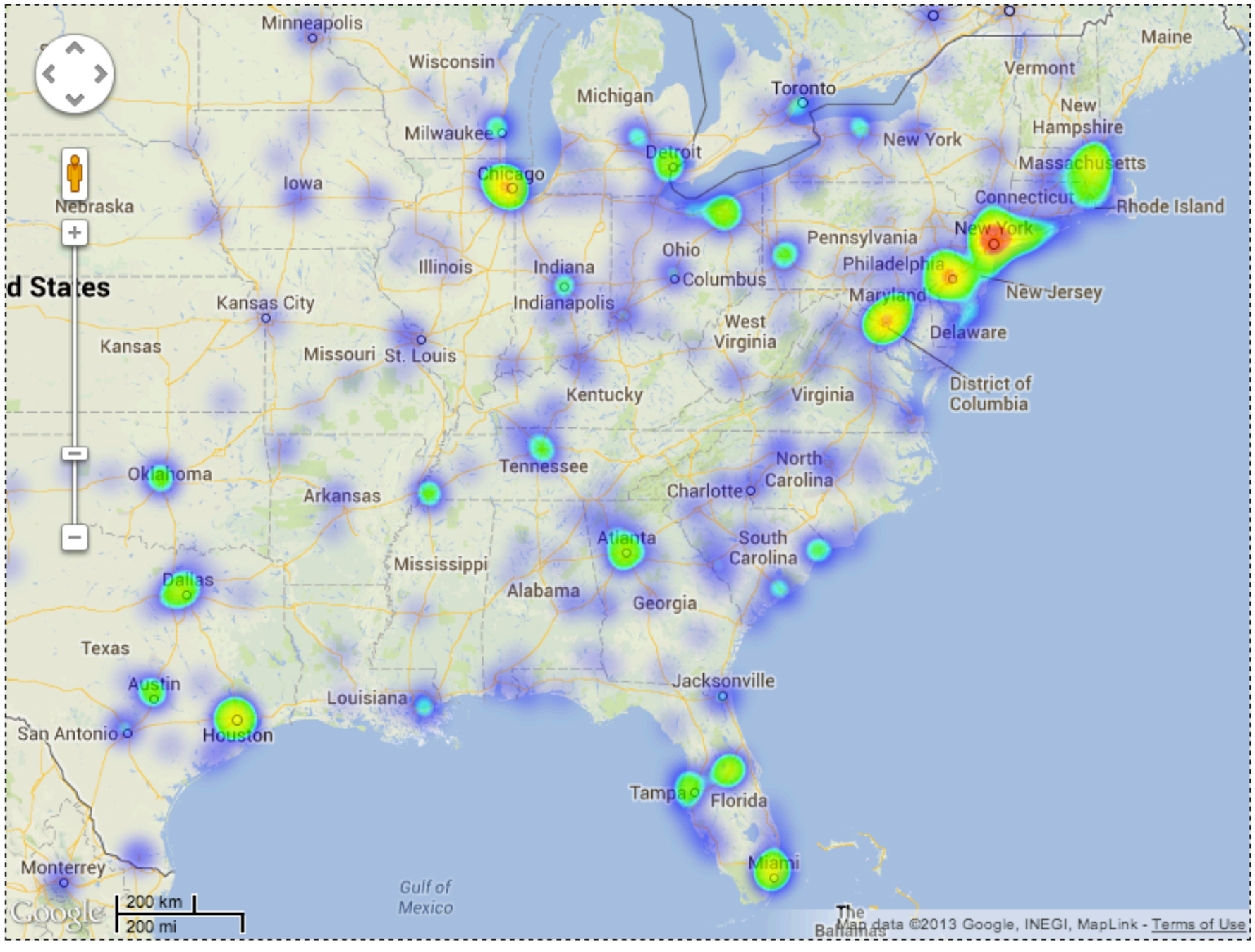
Panoramio

FourSquare

Social Media

heatmap.js

Social Media



Conclusion

- One small but important part of a network security analytic workflow

Links

https://github.com/Berico-Technologies/CLAVIN

https://github.com/Berico-Technologies/CLAVIN-contrib

clavin.bericotechnologies.com
@CLAVIN__ (two underscores)
@greenbacker
@tlpinney
@BericoTech (we're hiring!)