## World Wide Geo Web

Steve McDonald SpacemanSteve@gmail.com

## **Tomorrow's Solution**

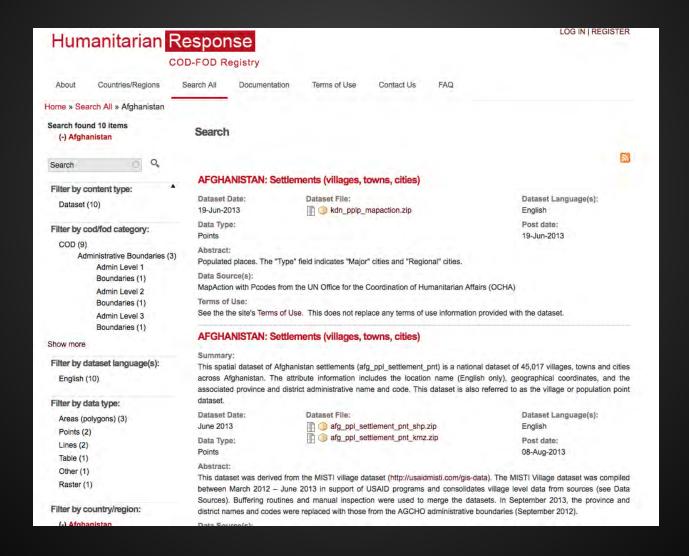
Crawl The Web For Data Files

**Preview Without Image Tiles** 

Download Over HTTP Get

No Need For OGC Protocols

## Typical Spatial Data Site



## Ingest Via Web Crawl

**User Provides URL For Site** 

Processes Links To Zip and Xml Files Parse Shape and Metadata

Uses crawl4j and jsoup

## cod.humanitarianresponse.info

#### 3.5 Hours On Home DSL

180 Layers With Title And Bounding Box

Provides Search, Not Preview And Download

## Rendering Via HTTP Get

Parse .shp, .shx, .dbf Files In Browser

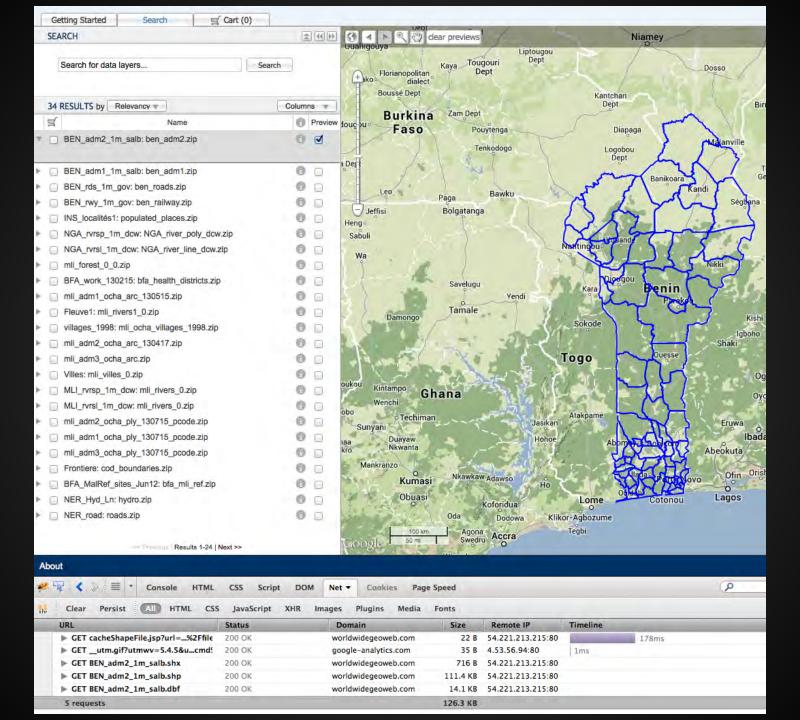
Render Elements In OpenLayers

Improved Interactivity For Small Files

Only Needs Zipped Shape Files

# Demo

http://WorldWideGeoWeb.com



### Vast Web Crawl

MapReduced 125GB Of HTML Pages

~One Million HTML Pages, 120 CPU Hours

Identified ~3500 Of KML/KMZ Links

Ingested 700 Files OGP Solr

**AWS Cluster Cost: \$4** 

# Spatial Data With HTTP

Steve McDonald SpacemanSteve@gmail.com