



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

SOUPED UP SEARCH



Julie Kanzler, GISP
DC Office of the CTO





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

AN EXAMPLE – TYPICAL GEO SEARCH DIALOG

Property Search

Search by Address

Example: [441 th St NW](#)

Search by Square, suffix, and lot

<input type="text"/>	<input type="text"/>	<input type="text"/>
----------------------	----------------------	----------------------

Example: [0489 E 0800](#)

Or Do An Advanced Search

Basic Property Search Instructions

You can use this search to quickly find your property by address, or square suffix lot number.

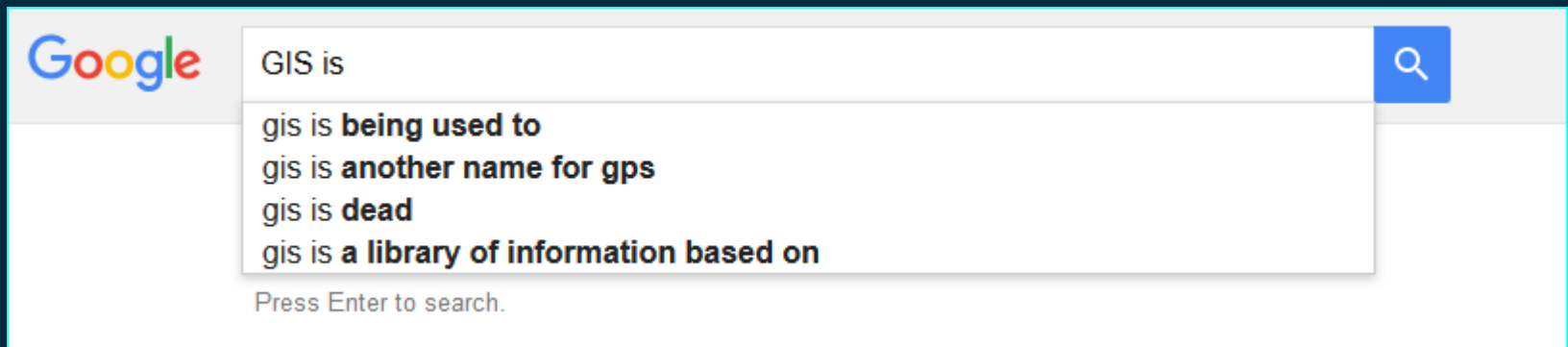
Search Tips

Remember that all addresses within the District of Columbia must end with a quadrant designation. (NW, NE, SW, or SE).

Property ID Numbers or Square Suffix Lot (SSL)

numbers, in most cases, are two sets of four numerals each. In some cases, there may be two sets of numerals, and one set of characters. (e.g. 0489 E 0800).

EXAMPLE - GOOGLE IS THE MODEL. OF COURSE.



No prior knowledge needed.

Start with a fragment of a question, and iterate.

Searching is a journey.

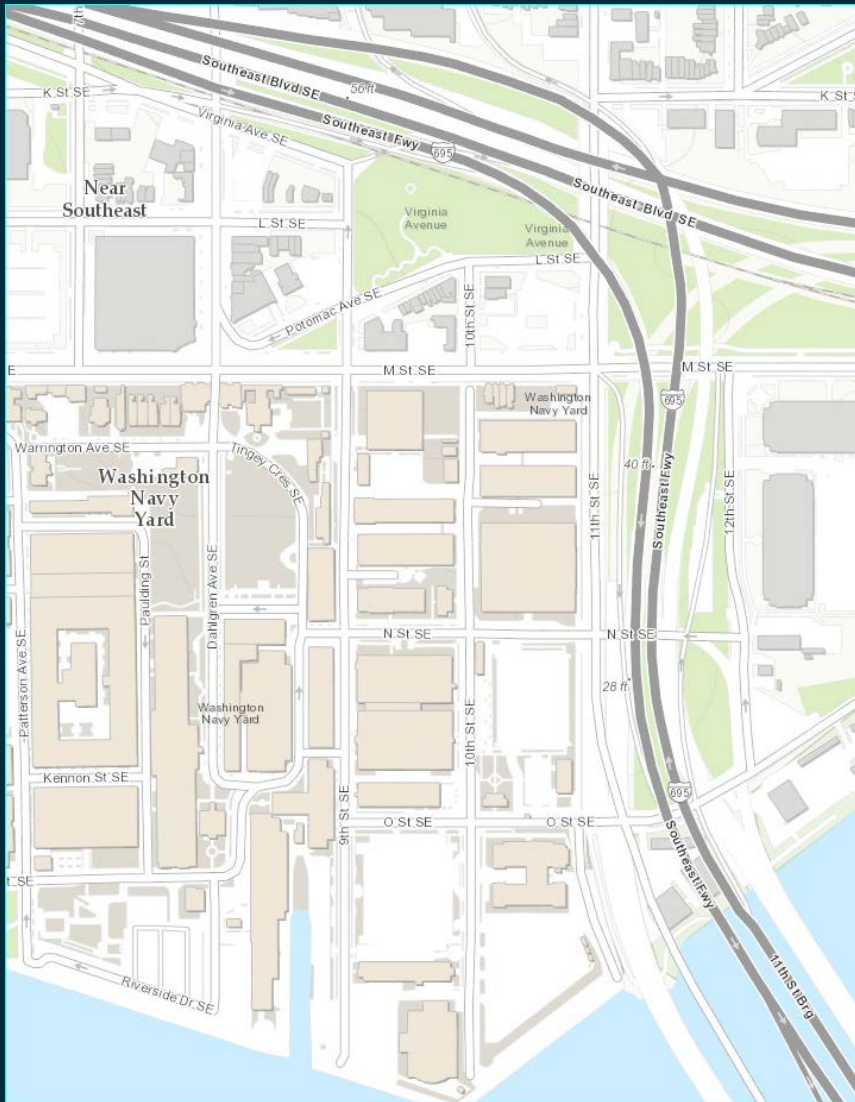


CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

HALLMARKS OF A SEARCH ENGINE

- “One box to rule them all”
- Fuzzy Searching
- Stemming
- Tokenizing
- Search-as-you-type Suggesters
- Relevance scoring
- Data Aggregation
- Faceted searching

If we want to change how we think about search, we need to change how we think about data



← this is
“Map-based
content”



-- Jared Spool
Founder of User Interface Engineering (UIE)

A world where everyone creates content
gets confusing pretty quickly



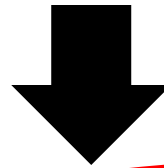
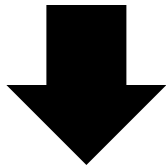
without a good search engine.

--Ethan Zuckerman, Director of the MIT Center for Civic Media

Search Engines for (geo) Data?

SEARCH ENGINES FOR (GEO) DATA

Lucene



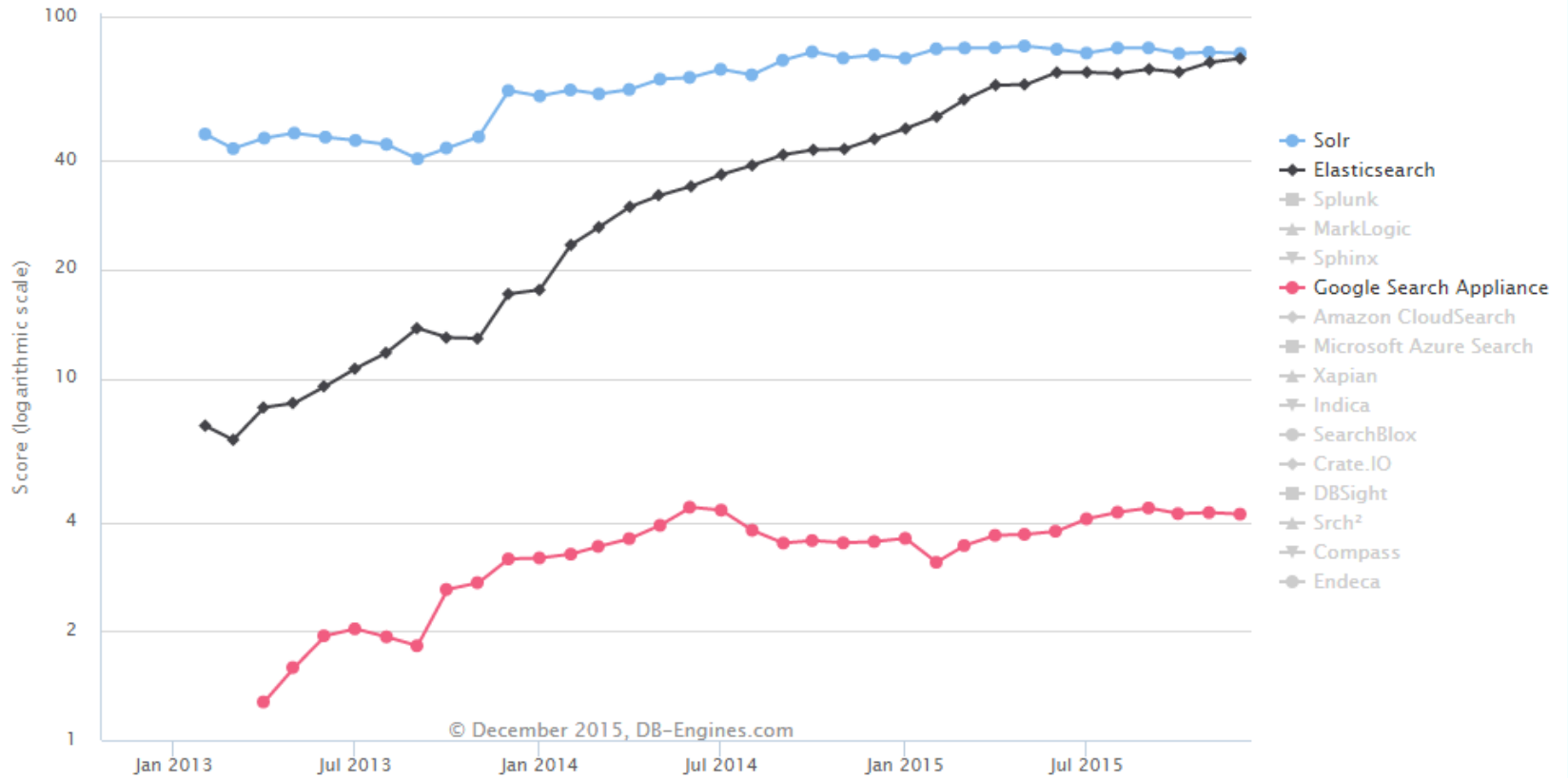
Apache

Solr



elastic

DB-Engines Ranking of Search Engines





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

HOW DID WE GET STARTED?

- Prior experience using elasticsearch to monitor logs
- Esri EAP engagement with Mansour Raad – Spring of 2015

WHAT IS ELASTICSEARCH?

- A search engine
- A series of RESTful web APIs that speaks (schema-less) JSON
- Open source, Mostly Free
- Scales from a single instance on a laptop to multi-node clusters
- Built-in load-balancing and failover



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

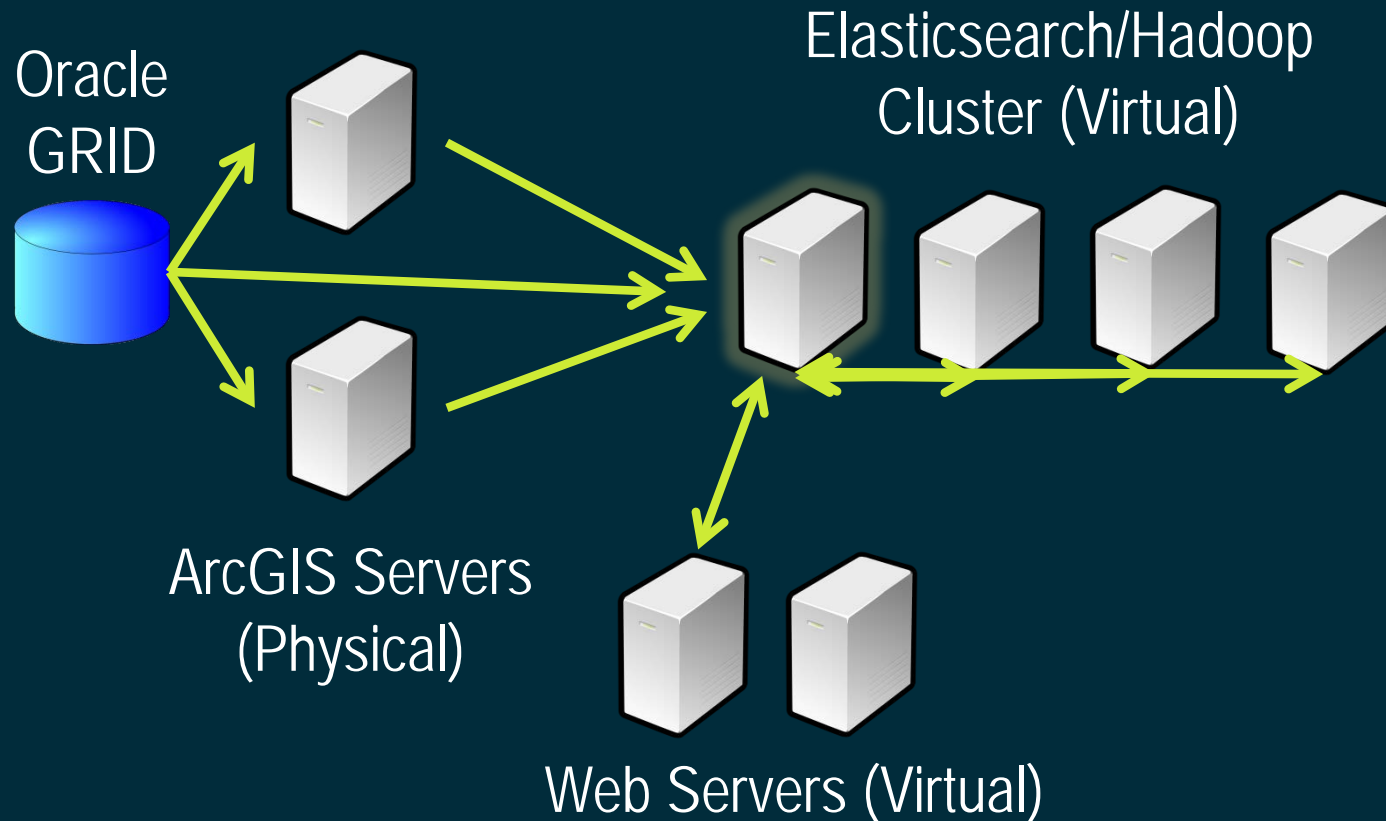
WHY ELASTIC IS ESPECIALLY HELPFUL FOR DATA

- It's an API in a box
- It's not a “black box” search engine – a high degree of control, can tune
- The node-based architecture is scalable and geared towards super-performant searching



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

HIGH LEVEL ARCHITECTURE





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

QUERY V. ELASTIC

Advanced Search Feature	Standard Query	Elasticsearch
Fuzzy Searching	No - Parameters must be exact	Yes
Data Aggregation	Yes – but requires significant finagling (and skill)	Yes – many options OOTB
Search-as-you-type Suggesters	Yes – but indexes are heavy and performance can be poor	Yes – across any data
Faceted searching	No – could do manually	Yes
Full text search – stemming, tokenizing, relevance scoring	Not really	Yes



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

USE MARVEL SENSE TO INTERACT & LEARN

The screenshot shows the Marvel Sense web interface in a browser. The address bar shows the URL `localhost:9200/_plugin/marvel/sense/`. The interface includes a server selection dropdown set to `localhost:9200`, a `Development Trial` badge, and a `Dashboards` menu. The main area is split into three panes: a REST client on the left, a request body in the middle, and a JSON response on the right.

```
1 GET _search
2 {
3   "query": {
4     "match_all": {}
5   }
6 }
7
8 GET /performance/kpi
9
```

```
793   "_score": 1,
794   "_source": {
795     "_source_node": {
796       "id": "CHAbwkBNQyOh_YS8FyHD9Q",
797       "name": "Typhon",
798       "transport_address": "inet[OCTO-JQYR12.ojs.dcgov.priv/192.168.1.102:9300]",
799       "ip": "192.168.1.102",
800       "host": "OCTO-JQYR12.ojs.dcgov.priv",
801       "ip_port": "192.168.1.102:9300",
802       "master_node": true,
803       "data_node": true,
804       "master": true
805     },
806     "@timestamp": "2015-09-01T01:11:06.570Z",
807     "cluster_name": "elasticsearch",
808     "message": "blocks changed, indices created, routing change",
809     "event": "cluster_state",
810     "event_source": "create-index [testindex], cause [api]",
811     "status": "red",
812     "version": 12,
813     "master_node": "CHAbwkBNQyOh_YS8FyHD9Q",
814     "blocks": {},
815     "nodes": {
816       "CHAbwkBNQyOh_YS8FyHD9Q": {
817         "name": "Typhon",
818         "transport_address": "inet[OCTO-JQYR12.ojs.dcgov.priv/192.168.1.102:9300]",
819         "attributes": {}
820       }
821     },
822     "routing_table": {
823       "indices": {
824         "testindex": {
825           "shards": {
826             "0": [
827               {
828                 "state": "INITIALIZING",
829                 "primary": true,
830                 "node": "CHAbwkBNQyOh_YS8FyHD9Q",
831                 "relocating_node": null,
832                 "shard": 0,
833                 "index": "testindex",
834                 "unassigned_info": {
835                   "reason": "INDEX_CREATED",
836                   "at": "2015-09-01T01:11:06.391Z"
837                 }
838               }
839             ]
840           }
841         }
842       }
843     }
844   }
845 }
```




CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

API'S, CONNECTORS, AND INTEGRATIONS

- Python API: `elasticsearch-py`
- Javascript API: `elasticsearch.js`
- .NET API
- PHP API
- Java API
- Perl API
- Ruby API
- Groovy API
- Elasticsearch for Apache Hadoop
- Community Contributions (e.g. Drupal Connector)
- **Probably your favorite client**



CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

THERE'S A LOT MORE TO IT THAN JUST SEARCH

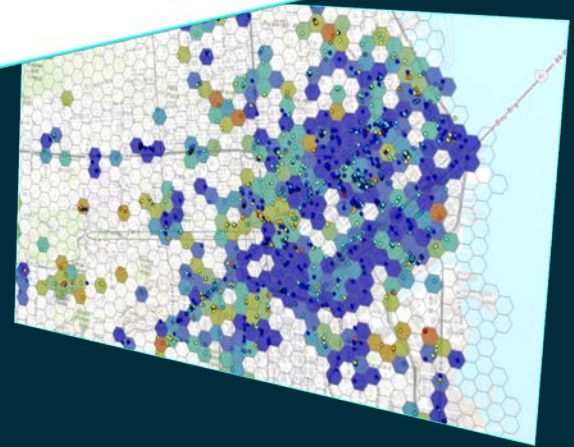
- Logstash
- Kibana
- Marvel
- Shield
- Others...

(some free, some not)



NEXT STEPS ...

- Integrate geospatial/data search with web search (pilot underway)
- Deploy as back-end for
 - transparency web app with performance metrics
 - Crime web app
- Integrate with geocoder





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

WANT TO LEARN MORE?

- The Definitive Guide

<https://www.elastic.co/guide/en/elasticsearch/guide/current/index.html>

- Exploring Elasticsearch

<http://exploringelasticsearch.com/>

- Using Spatial Data in Elastic Search

<https://mapbutcher.gitbooks.io/using-spatial-data-in-elasticsearch/content/>





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.



What I need is a search engine that,
no matter what I type in, comes
back with GO BACK TO WORK.

— *Dave Barry* —

AZ QUOTES





CUSTOMER SERVICE. ACCOUNTABILITY. EFFICIENCY. SECURITY.

THANK YOU

JULIE KANZLER, GISP
DC OFFICE OF THE CTO

 @JULIEKDANG

