

MAKE SENSE OF YOUR (BIG) DATA!

David Pilato
Technical advocate

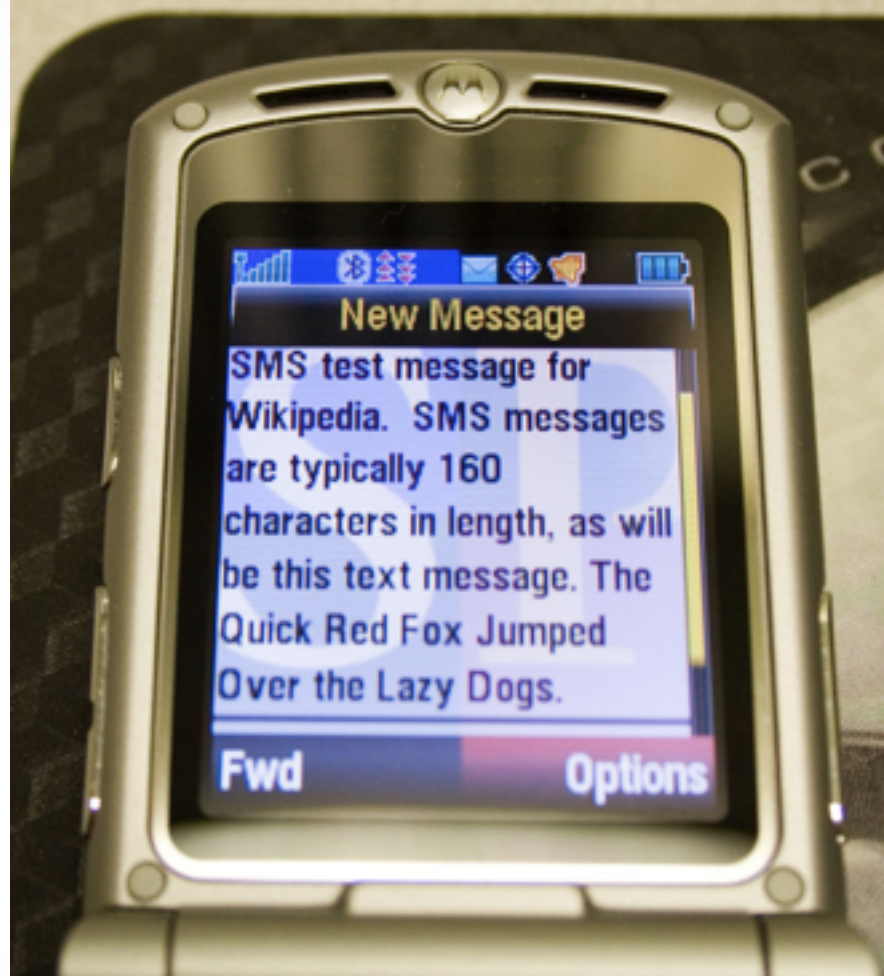


elasticsearch.
@dadoonet





data ?





elasticsearch

@elasticsearch

Elasticsearch 1.4.0.Beta1 is out, with improvements galore!

elasticsearch.org/blog/elasticse...

5:48 PM - 1 Oct 2014

84 RETWEETS 17 FAVORITES

```
{
  "created_at": "Wed Oct 01 15:48:54 +0000 2014",
  "id": 517340385371099140,
  "text": "Elasticsearch 1.4.0.Beta1 is out, with resiliency improvemen",
  "user": {
    "name": "elasticsearch",
    "screen_name": "elasticsearch",
    "description": "Open Source, Distributed, RESTful Search Engine",
    "followers_count": 10394,
    "created_at": "Fri Oct 23 05:25:26 +0000 2009",
    "statuses_count": 1911,
    "lang": "en",
  },
  "retweet_count": 84,
  "favorite_count": 17,
  "entities": {
    "urls": [
      {
        "url": "http://t.co/eMrv2XHJms",
        "expanded_url": "http://www.elasticsearch.org/blog/elasticsearch",
        "display_url": "elasticsearch.org/blog/elasticse_"
      }
    ]
  },
  "lang": "en"
}
```



DEEZER Découvrir Deezer ▾

⏮ ⏪ ⏩ ⏭ 🔊 00:00

À écouter

Explorer

Top Écoutes

Radios

APPLICATIONS

📱 Découvrez nos apps

📻 Shuffler.fm

BIBLIOTHÈQUE

👤 David Pilato

S'abonner

📖 Cours de cœur

Harder
Daft Punk | Disc

Partager

Ce titre est présent dans

daft punk

Discovery
par Daft Punk

```
{
  "id": "3135556",
  "readable": true,
  "title": "Harder Better Faster Stronger",
  "isrc": "GBDUM0000059",
  "link": "http://www.deezer.com/track/3135556",
  "duration": "225",
  "track_position": 4,
  "disk_number": 1,
  "rank": "783236",
  "explicit_lyrics": false,
  "preview": "http://cdn-preview-5.deezer.com/stream/51afcde9f56a132096c0496cc95eb24b-3.mp3",
  "bpm": 123.5,
  "gain": -6.48,
  "artist": {
    "id": "27",
    "name": "Daft Punk",
    "link": "http://www.deezer.com/artist/27",
    "picture": "https://api.deezer.com/artist/27/image",
    "radio": true,
    "type": "artist"
  },
  "album": {
    "id": "302127",
    "title": "Discovery",
    "link": "http://www.deezer.com/album/302127",
    "cover": "https://api.deezer.com/album/302127/image",
    "release_date": "2005-01-24",
    "type": "album"
  },
  "type": "track"
}
```




THE GREAT TRIANGULAR DUEL

By CAPTAIN FREDERICK MARRYAT

JACK walked up to the boatswain, and, taking off his hat, with the utmost politeness, said to him:
"If I mistake not, Mr. Biggs, your conversation refers to me."

"Very likely it does," replied the boatswain. "Listeners hear no good of themselves."

"It appears that gentlemen can't converse without being vatched," continued Mr. Easthupp, pulling up his shirt-collar."

"It is not the first time you have thought proper to make very offensive remarks, Mr. Biggs; and as you appear to consider yourself ill-treated in the affair of the trousers, for I tell you at once that it was I who brought them on board, I can only say," continued our hero, with a very polite bow, "that I shall be most happy to give you satisfaction."



DevOps Borat

@DEVOPS_BORAT



Follow

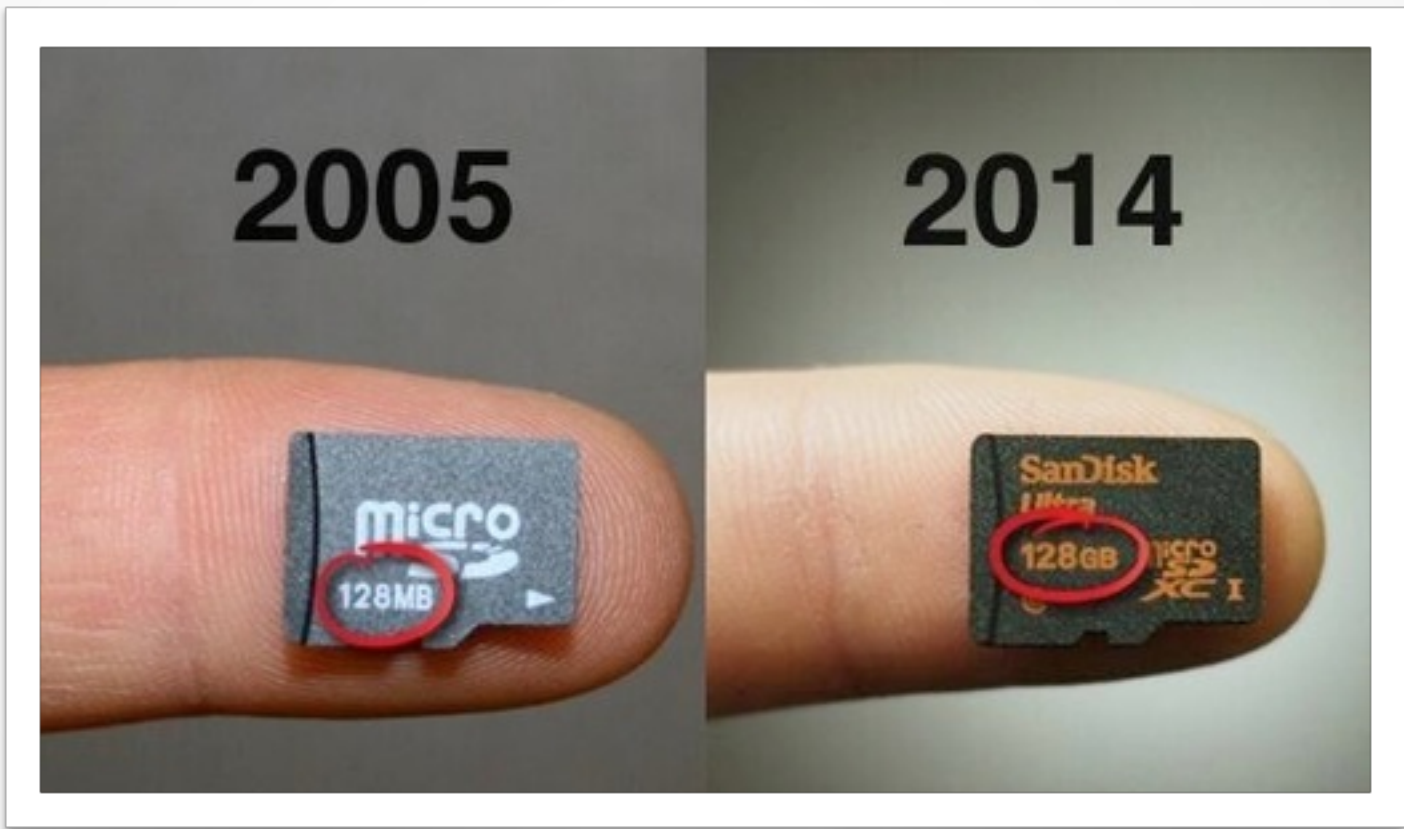
Big Data is any thing which is crash Excel.

6:25 PM - 8 Jan 2013

2,110 RETWEETS **513** FAVORITES



BIG data ?



The production of data is expanding at an astonishing pace. Experts now point to a 4300% increase in annual data generation by 2020. Drivers include the switch from analog to digital technologies and the rapid increase in data generation by individuals and corporations alike.

Size of Total Data
Enterprise Created Data
Enterprise Managed Data

2012: CUSTOMERS WILL START STORING 1 EB OF INFORMATION.

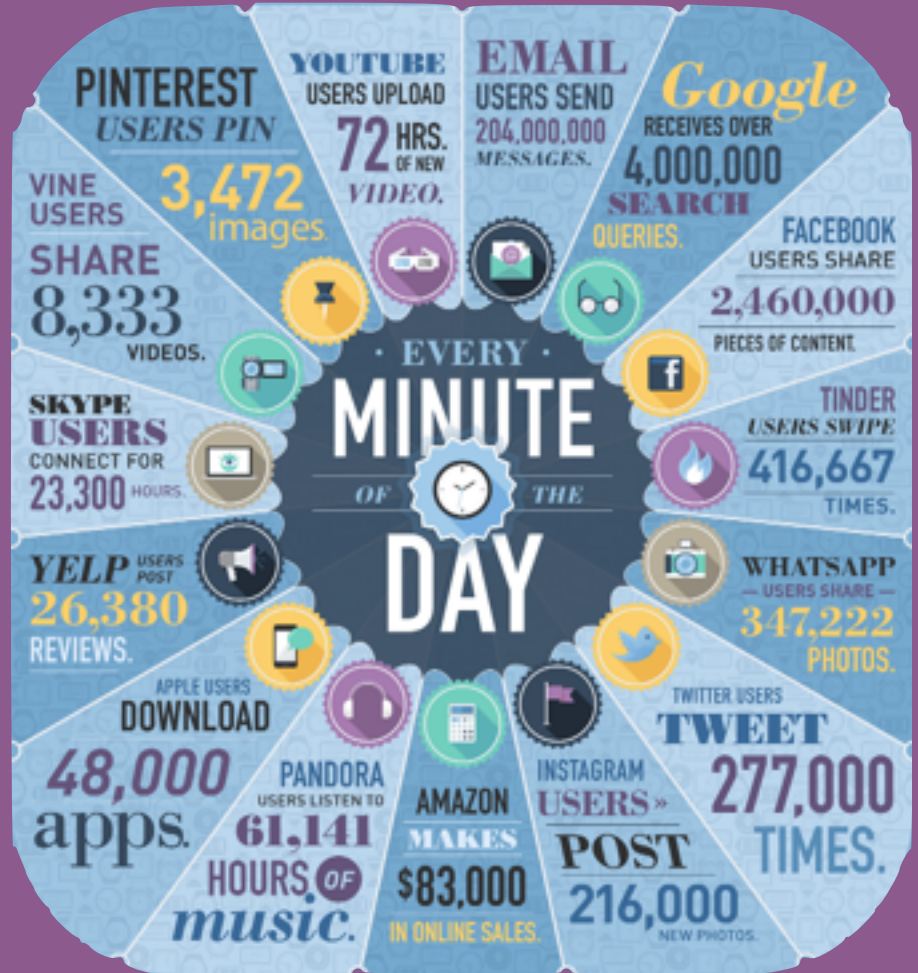


2020: MORE THAN 1/3 OF THE DATA PRODUCED WILL LIVE IN OR PASS THROUGH THE CLOUD.

35.000.000.000.000.000 mb

Source: http://www.csc.com/insights/flxwd/78931-big_data_just_beginning_to_explode





Source: <http://www.domo.com/learn/data-never-sleeps-2>





search = like % ?

```
SELECT
```

```
  doc.*, country.*
```

```
FROM
```

```
  doc, country
```

```
WHERE
```

```
  doc.country_code = country.code AND
```

```
  doc.date_doc > to_date('2011-12', 'yyyy-mm') AND
```

```
  doc.date_doc < to_date('2012-01', 'yyyy-mm') AND
```

```
  lower(country.name) = 'france' AND
```

```
  lower(doc.comment) LIKE '%product%' AND
```

```
  lower(doc.comment) LIKE '%david%';
```

Search engine ?

1964

Alphabetical **INDEX**

A	Accessory Groups.....	63
	Additional Accessories.....	62
	Adjustable Steering Wheel.....	54
	Air Cleaner.....	95
	Water-Heater System.....	52, 53
	Brake Control.....	71, 95
	Brake Lamp Timer.....	51
	Braking System.....	88, 89
		91, 101

B	Back-Up Lights.....	9
	Battery.....	70, 97
	Body Construction.....	82A
	Body Insulation.....	87
	Body Sealing.....	87
	Body Styles.....	26-49
	Brake Drums.....	76, 77, 102
	Brake Lining.....	77, 102
	Brakes.....	76, 77, 102
	Brakes, Parking.....	76, 77, 102
	Brakes, Power.....	76, 77, 102
	Bucket Seat Console.....	25
	Bucket Seats.....	35-39, 45
	Bumpers.....	9

		69, 94
		71, 95
		72-81
	View.....	72A
	Department.....	48, 49
		71, 95
	Chamber.....	68
	Ratio.....	52, 53
		91, 92
	Connecting Rods.....	93
	Console, Bucket Seat.....	25
	Constant Velocity (CV) Joints.....	74, 75, 100
	Controlled Differential.....	58
	Cooling.....	95, 96



plug & play

scalable

elasticsearch

Apache 2 license

REST/JSON

Lucene

Start...

```
$ wget https://download.elasticsearch.org/elasticsearch/  
elasticsearch/elasticsearch-1.3.4.tar.gz  
$ tar -xf elasticsearch-1.3.4.tar.gz  
$ ./elasticsearch-1.3.4/bin/elasticsearch  
[INFO ][node                ][Ghost Maker] {1.3.4}[5645]: initializing
```

... and play!

```
$ curl -XPUT localhost:9200/sessions/session/1 -d '{
  "title"      : "Elasticsearch",
  "subtitle"   : "Make sense of your (BIG) data !",
  "date"       : "2014-10-17T14:30:00",
  "tags"       : [ "elasticsearch", "bdxio", "bigdata" ],
  "speakers"  : [{
    "first_name" : "David",
    "last_name"  : "Pilato"
  }]
}'
```

Search!

```
$ curl http://localhost:9200/sessions/session/_search -d'
{
  "query": {
    "multi_match": {
      "query": "elasticsearch bdxio david",
      "fields": [ "title^3", "tags^2", "speakers.first_name" ]
    }
  },
  "post_filter": {
    "range": {
      "date": {
        "from": "2014-10-01",
        "to": "2014-11-01"
      }
    }
  }
}
```

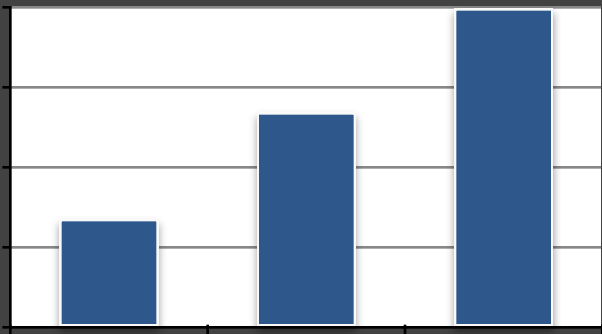

Compute?

<i>B</i>	Back-Up Lights.....	9
	Battery.....	70, 97
	Body Construction.....	82A
	Body Insulation.....	87
	Body Sealing.....	87
	Body Styles.....	26-49
	Brake Drums.....	76, 77, 102
	Brake Lining.....	77, 102
	Brakes.....	76, 77, 102
	Brakes, Parking.....	76, 77, 102
	Brakes, Power.....	76, 77, 102
	Bucket Seat Console.....	25
	Bucket Seats.....	35-39, 45
	Bumpers.....	9

Compute!

```
$ curl http://localhost:9200/sessions/session/_search -d'
```

```
{  
  "query": { ... },  
  "aggs": {  
    "by_date": {  
      "date_histogram": {  
        "field": "date",  
        "interval": "day",  
        "format": "dd/MM/yyyy"  
      }  
    }  
  }  
}'
```



```
"by_date": [  
  { "key_as_string": "03/04/2014", "doc_count": 1 },  
  { "key_as_string": "12/04/2014", "doc_count": 2 },  
  { "key_as_string": "16/04/2014", "doc_count": 3 }  
]
```



Let's make sense of ...

- logs
- twitter
- github
- marketing data
- ...
- your data
- your big data

Let's make sense of ...

- logs
- twitter
- github
- marketing data
- ...
- your data
- your big data

```
{  
  "name": "Pilato David",  
  "dateOfBirth": "1971-12-26",  
  "gender": "male",  
  "children": 3,  
  "marketing": {  
    "fashion": 334,  
    "music": 3363,  
    "hifi": 2351  
  },  
  "address": {  
    "country": "France",  
    "city": "Paris",  
    "location": [2.332395, 48.861871]  
  }  
}
```


démo

MAKE SENSE OF YOUR (BIG) DATA!

let's inject some marketing documents...

ELASTICSEARCH





elasticsearch.

elasticsearch



kibana



logstash



Marvel





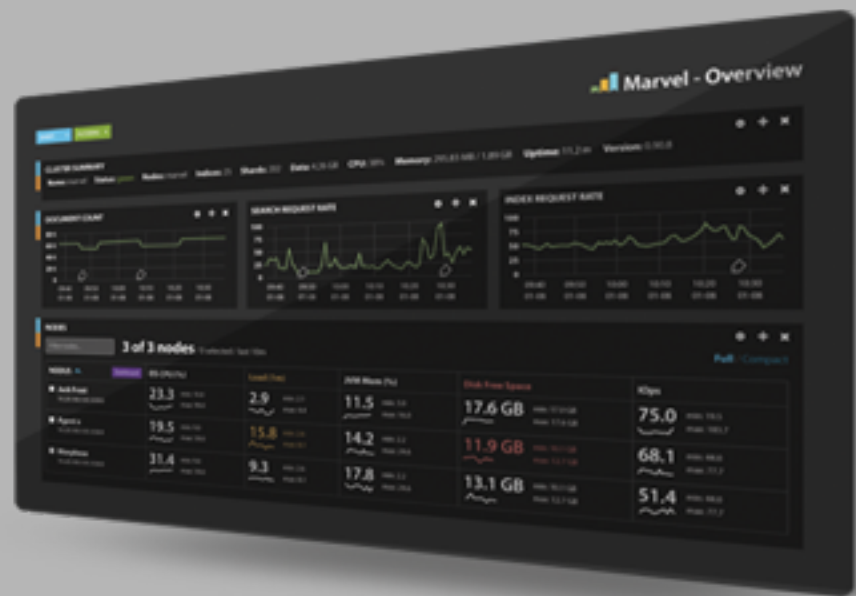
elasticsearch.

Training (public and on-site)

Development support

Production support

Marvel







thanks

@dadoonet

we are hiring!