



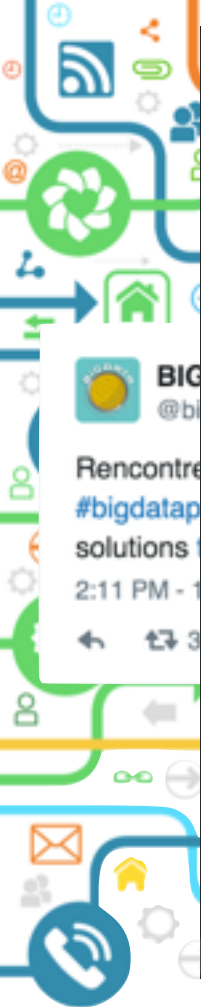
David Pilato
Developer | Evangelist
@dadoonet

make sense of your (BIG) data!



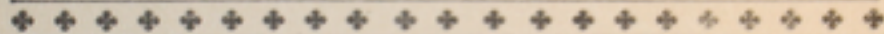






```
{
  "created_at": "Thu Dec 10 13:11:28 +0000 2015",
  "id": 674939490406920192,
  "text": "Rencontrez @dadoonet et @ElasticFR sur le congr\u00e8s #bigdataparis les 7 et 8 mars 2016 et d\u00e9couvrez les solutions de Big Data",
  "user": {
    "name": "BIG DATA PARIS",
    "screen_name": "bigdataparis",
    "location": "Paris",
    "description": "Le Congr\u00e8s BIG DATA PARIS. Faites de l'analyse de #donn\u00e9es le facteur cl\u00e9 de la performance",
    "followers_count": 3536,
    "statuses_count": 3674,
    "lang": "fr"
  },
  "is_quote_status": true,
  "retweet_count": 3,
  "favorite_count": 1,
  "entities": {
    "hashtags": [
      { "text": "bigdataparis" }
    ],
    "user_mentions": [
      { "screen_name": "dadoonet", "name": "David Pilato" },
      { "screen_name": "ElasticFR", "name": "Elastic FR" }
    ]
  }
}
```


[illegible]



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."

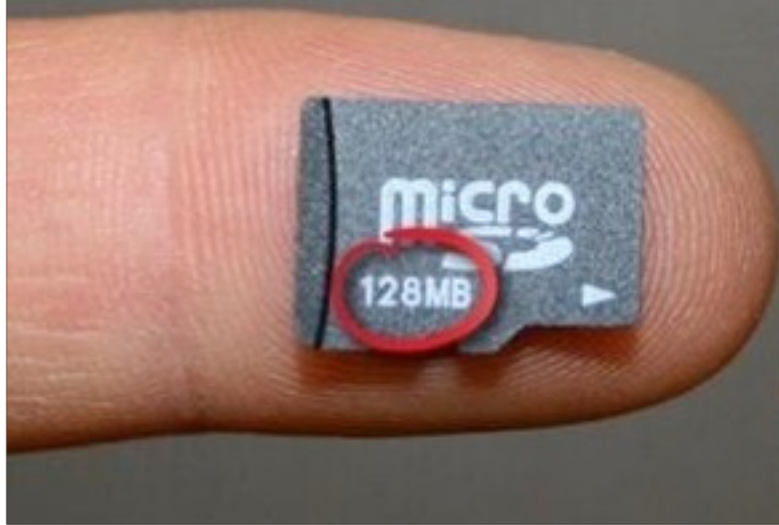
"It happens that gentlemen can't converse without being watched," continued Mr. Easthupp, pulling up his shirt-collar."

"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."

Big data?



2005



2014



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

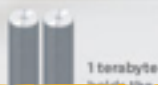
2020: MORE THAN 1/3 OF THE DATA PRODUCED WILL LIVE IN OR PASS THROUGH THE CLOUD.

2012: CUSTOMERS WILL START STORING 1 EB OF INFORMATION.



WHAT IS A ZETTABYTE?

1,000,000,000,000 gigabytes
1,000,000,000,000,000 bytes
1,000,000,000,000,000,000 bits
1,000,000,000,000,000,000,000 octets
1,000,000,000,000,000,000,000,000 dodecets
1,000,000,000,000,000,000,000,000,000 hectets
1,000,000,000,000,000,000,000,000,000,000 attets
1,000,000,000,000,000,000,000,000,000,000,000 zettets



1 terabyte

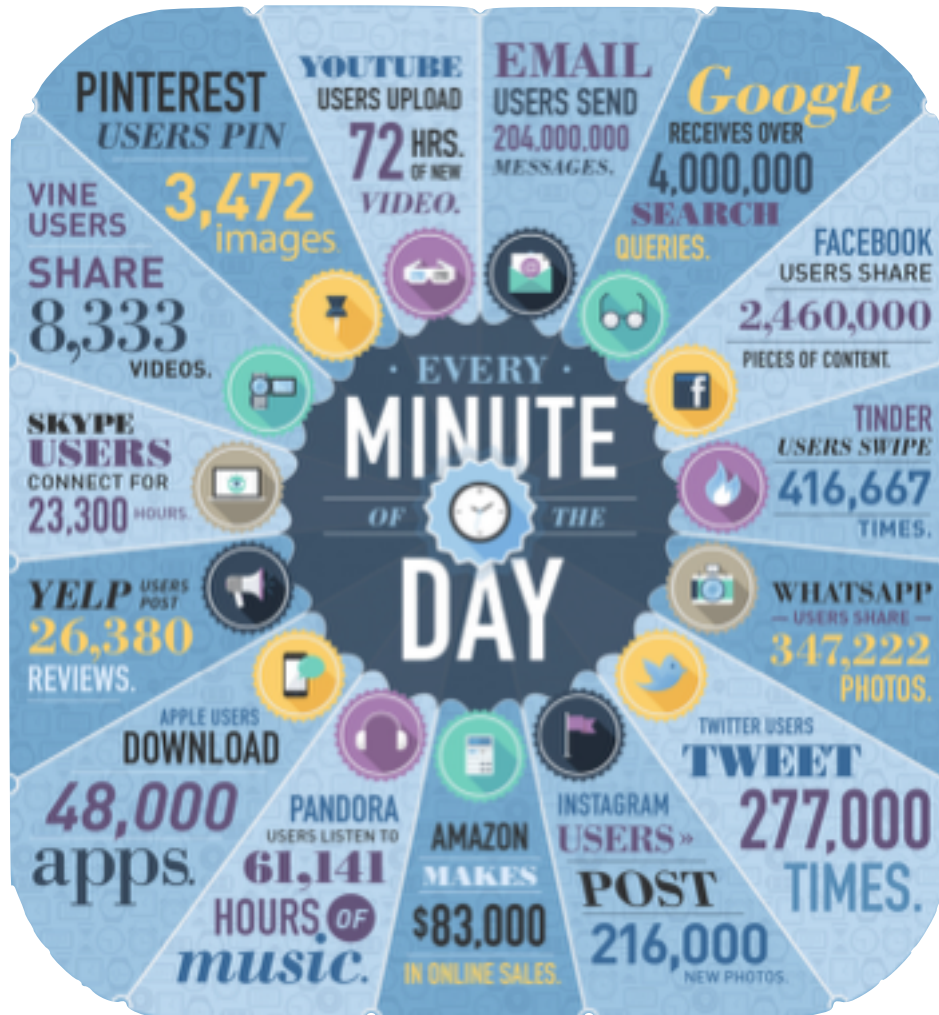
It took roughly 1 petabyte of local storage to render the



In 2007, the estimated

DATA PRODUCTION WILL BE 44 TIMES GREATER IN 2020 THAN IT WAS IN 2009

35.000.000.000.000.000.000 mb



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?

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

1964

Alphabetical INDEX

A	Accessory Groups.....	63
	Additional Accessories.....	62
	Adjustable Steering Wheel.....	54
	Air Cleaner.....	95
	Air Conditioner-Heater System.....	52, 53
	Ash Receivers.....	23
	Automatic Choke Control.....	71, 95
	Automatic Headlamp Timer.....	51
	Automatic Heating System.....	88, 89
	Axle Ratios.....	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

C	Camshaft.....	69, 94
	Carburetor.....	71, 95
	Chassis Features.....	72-81
	Chassis Over-All View.....	72A
	Chauffeur Compartment.....	48, 49
	Choke Control.....	71, 95
	Combustion Chamber.....	68
	Comfort Control.....	52, 53
	Compression Ratio.....	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/release/  
org/elasticsearch/distribution/tar/elasticsearch/2.2.0/  
elasticsearch-2.2.0.tar.gz  
$ tar -xf elasticsearch-2.2.0.tar.gz  
$ ./elasticsearch-2.2.0/bin/elasticsearch  
[INFO ][node      ][Ghost Maker] version[2.2.0], pid[72965], ...  
[INFO ][transport][Ghost Maker] publish_address {[/127.0.0.1:9300]}  
[INFO ][http      ][Ghost Maker] publish_address {[/127.0.0.1:9200]}  
[INFO ][node      ][Ghost Maker] started
```


... and play!

```
$ curl -XPUT localhost:9200/sessions/session/1 -d '{
  "title"      : "Elasticsearch",
  "subtitle"   : "Make sense of your (BIG) data !",
  "date"       : "2016-03-07T16:30:00",
  "tags"       : [ "realtime", "bigdata" ],
  "speakers"   : [{
    "first_name" : "David",
    "last_name"  : "Pilato"
  }]
}'
```

search!

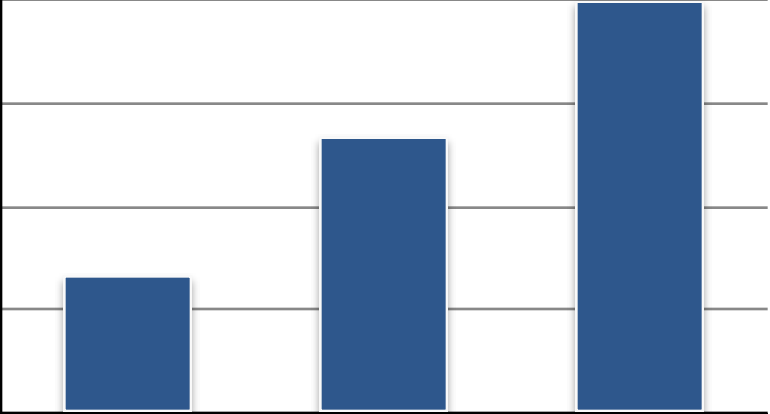
```
$ curl http://localhost:9200/sessions/session/_search -d'
{
  "query": {
    "multi_match": {
      "query": "elasticsearch bigdata david",
      "fields": [ "title^3", "tags^2", "speakers.first_name" ]
    }
  },
  "post_filter": {
    "range": {
      "date": {
        "from": "2016-03-01",
        "to": "2016-04-01"
      }
    }
  }
}'
```

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"
      }
    }
  }
}'
```



key_as_string	doc_count
03/03/2016	1
07/03/2016	2
10/03/2016	3

```
  "by_date": [
    { "key_as_string": "03/03/2016", "doc_count": 1 },
    { "key_as_string": "07/03/2016", "doc_count": 2 },
    { "key_as_string": "10/03/2016", "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]  
  }  
}
```

Demo

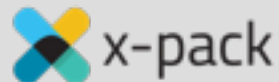
Let's inject **1 000 000**
marketing documents



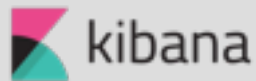
Demo

The Elastic Stack

Plugins



User Interface



Store, Index & Analyze



Ingest



Hosted Service







David Pilato
Developer | Evangelist
@dadoonet

Thank you!

<https://www.elastic.co/subscriptions>