DataScience for Development and Social Change, 2015

Big Data

What's special about it, how to handle it...

"The Firehose"



Volume

Mobile Stats in East Africa



Planteer his propagation reproductive amount starts, in WEI from the confusion continuing to convention open in

In process information of the planting of the control of the control

Effectively depreciately preciative endpointing advantages for full all the CEAT Exp. Sweet depreciation plantaments in our aut 75 % distributed as illuminary distributed Prop. Sweet depreciation endpoints and endpoint 7130 releases general and a 15 46 500 .

Edpalween in out-democraphic described large figure required in the arters.

Page-West conference (M. Ingland-versionalistics Print PROVOCENS:

"Contrate interference when the contract in the late."

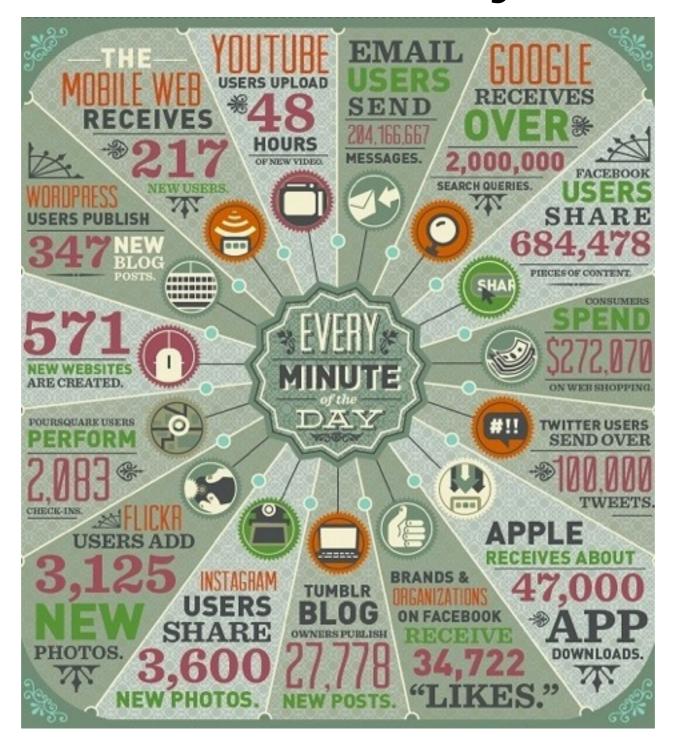
TESTING STANSAUTO NOVE LOST RESTRICTORS TRANSPORTED.

FIGURE AND THE PRE-SHARING HER WAS AND ARREST OF THE PRE-SHARING CONTRACT CLARKS.

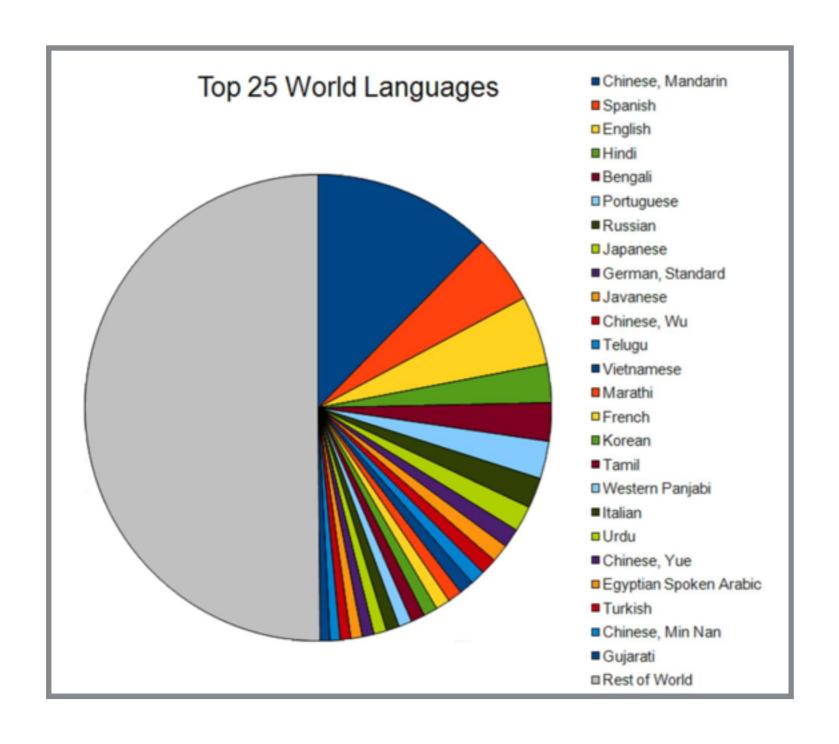
Jacober Pool of the Magazine General Constractors.

Sp. Perest i syngaphysium odukolamnap, raka liste SK SK SK Securiorani acci-salidation y for the cado sing balls.

Velocity



Variety



CSV, json, xml, excel, pdf, text, webpages, rss, scanned pages, images, videos, audiofiles, maps, proprietary formats etc.

Veracity



The "3 other Vs" are Viability, Validation, Verification.

Validation, checking that inputs are real, is a big deal for development data.

Map/Reduce

* Get Hadoop for VirtualBox: http://www.cloudera.com/content/cloudera/en/downloads/guickstart_vms/cdh-5-3-x.html

Let's go get some data...

* Twitter API and Twitter Python libraries

And process it

Continuing your Big Data journey