

**Big Data generated by
People:**

How is it being used?

After this video you will be able to..

- List the main advantages of Hadoop, Storm, Spark, and NoSQL to tackle the challenges of unstructured data
- State what a traditional warehouse does and why Big Data pushes for change
- Summarize examples of turning human generated data into value

Every challenge is an opportunity!

CHALLENGE →

OPPORTUNITY →



INTEGRATION



$$V = \frac{\Delta x}{\Delta t}$$







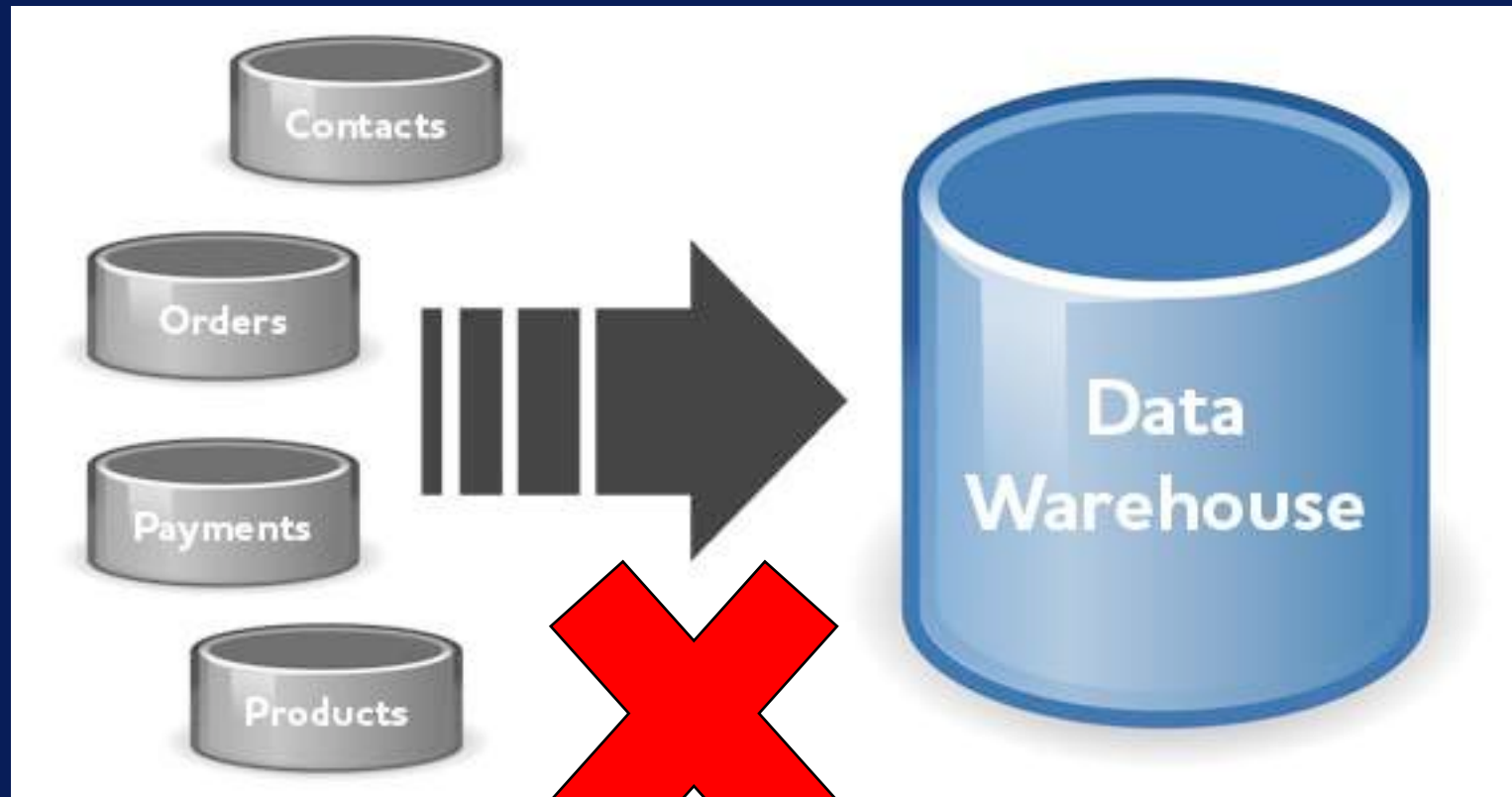


APACHE
STORM[™]
Distributed • Resilient • Real-time

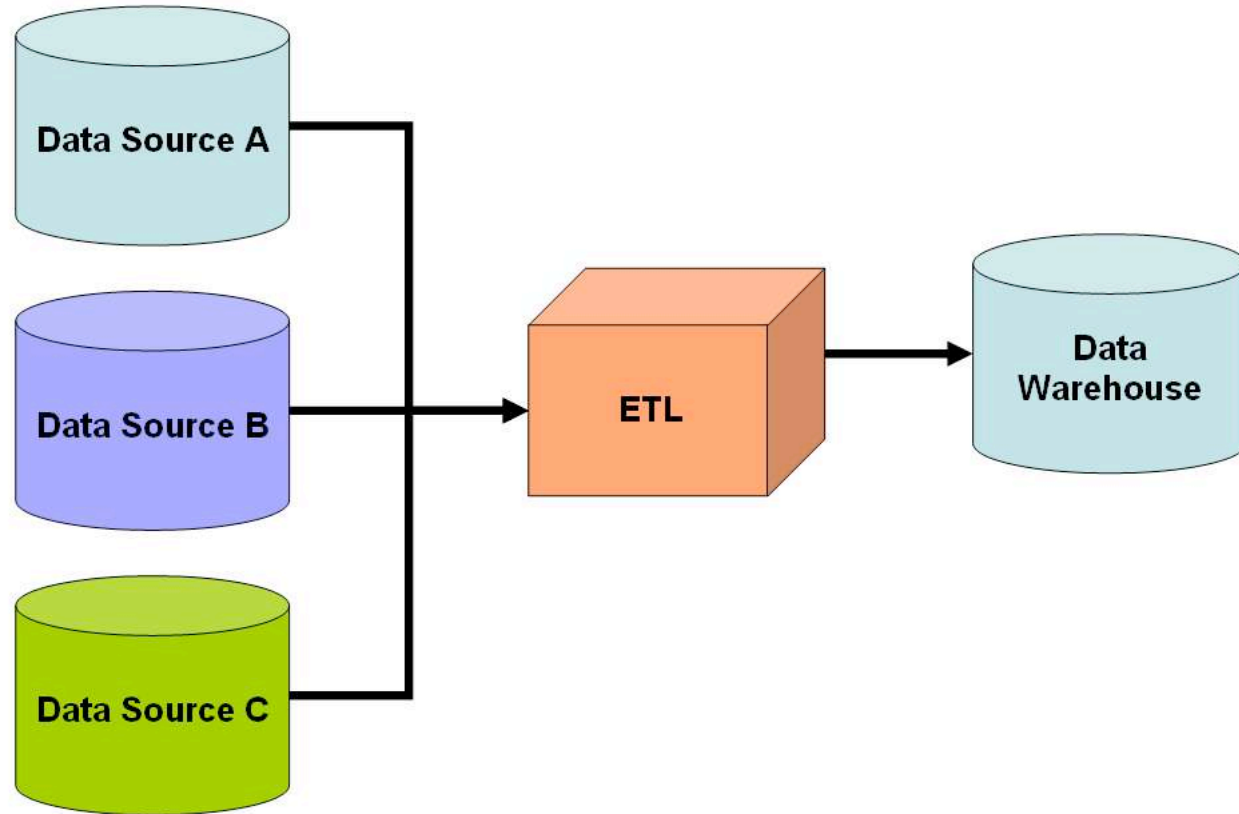


$$V = \frac{\Delta x}{\Delta t}$$

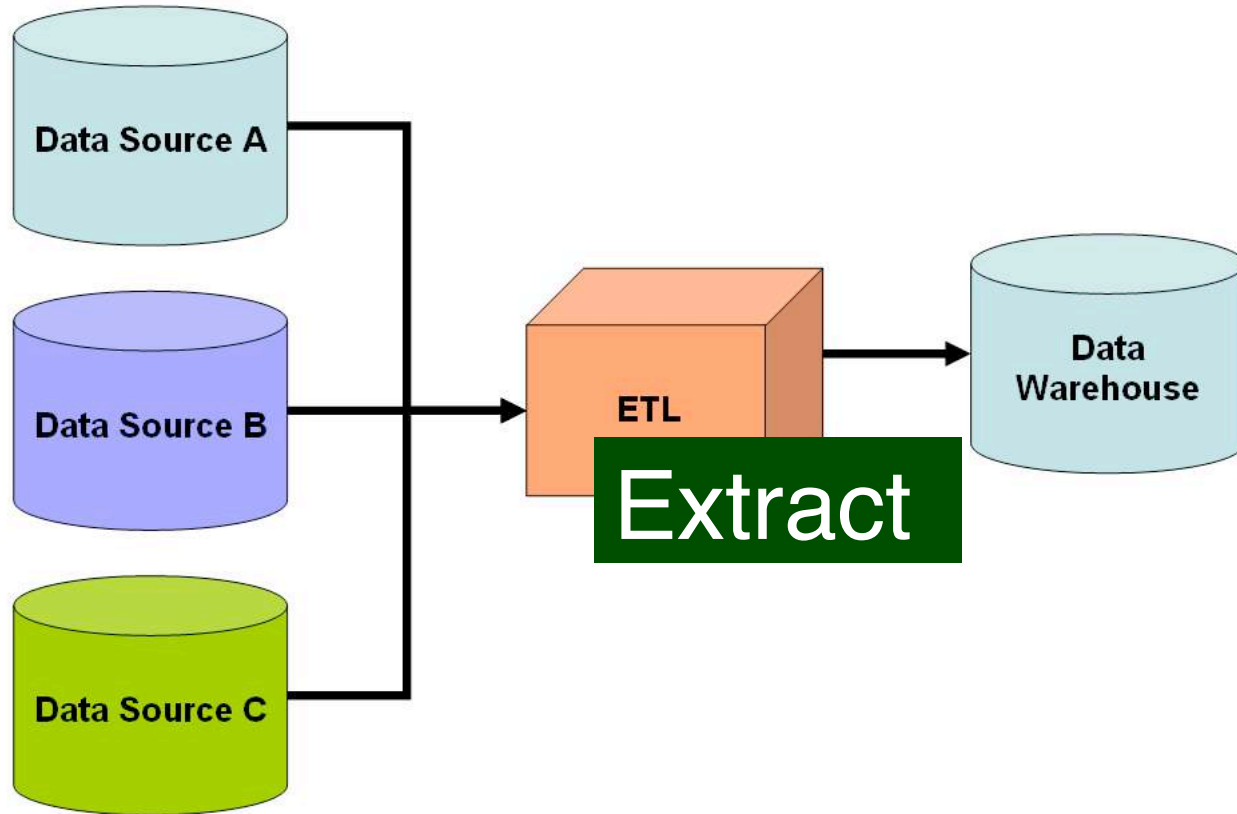




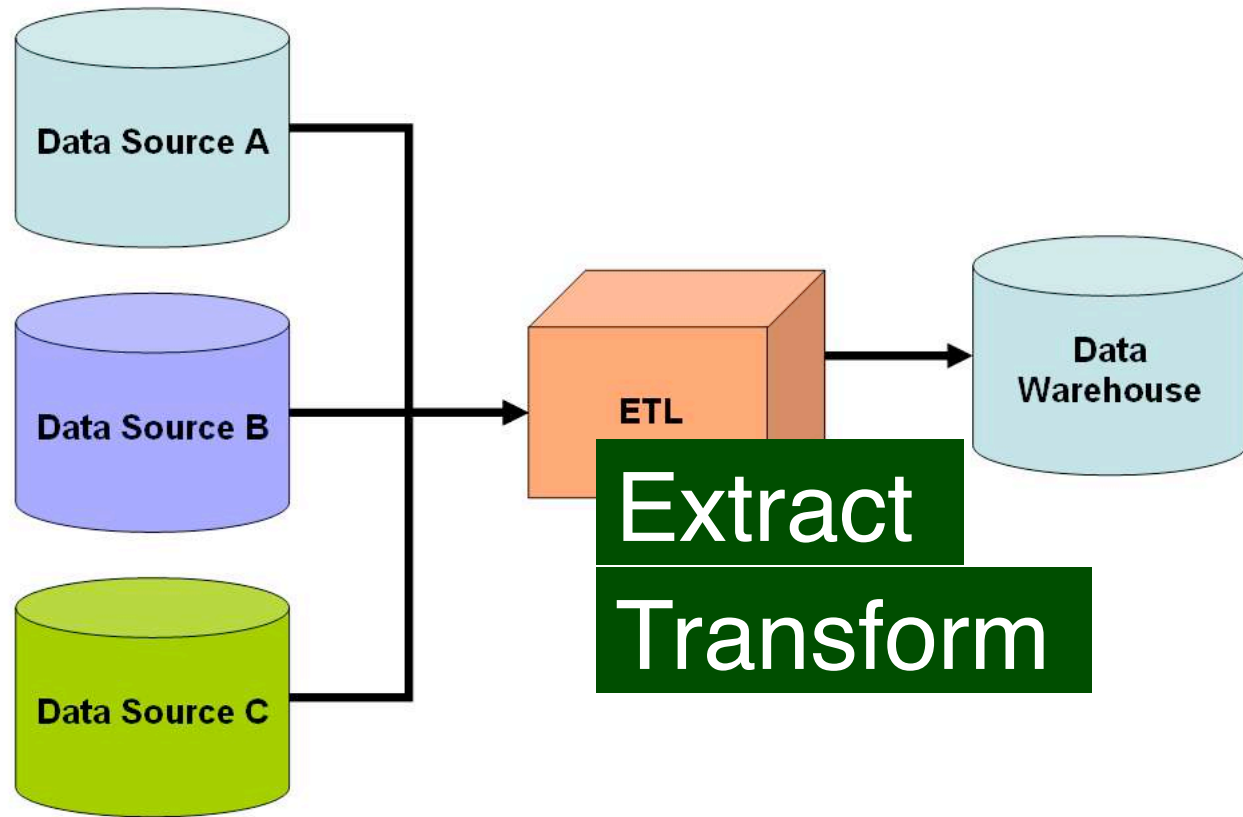
Traditional Data Warehouse



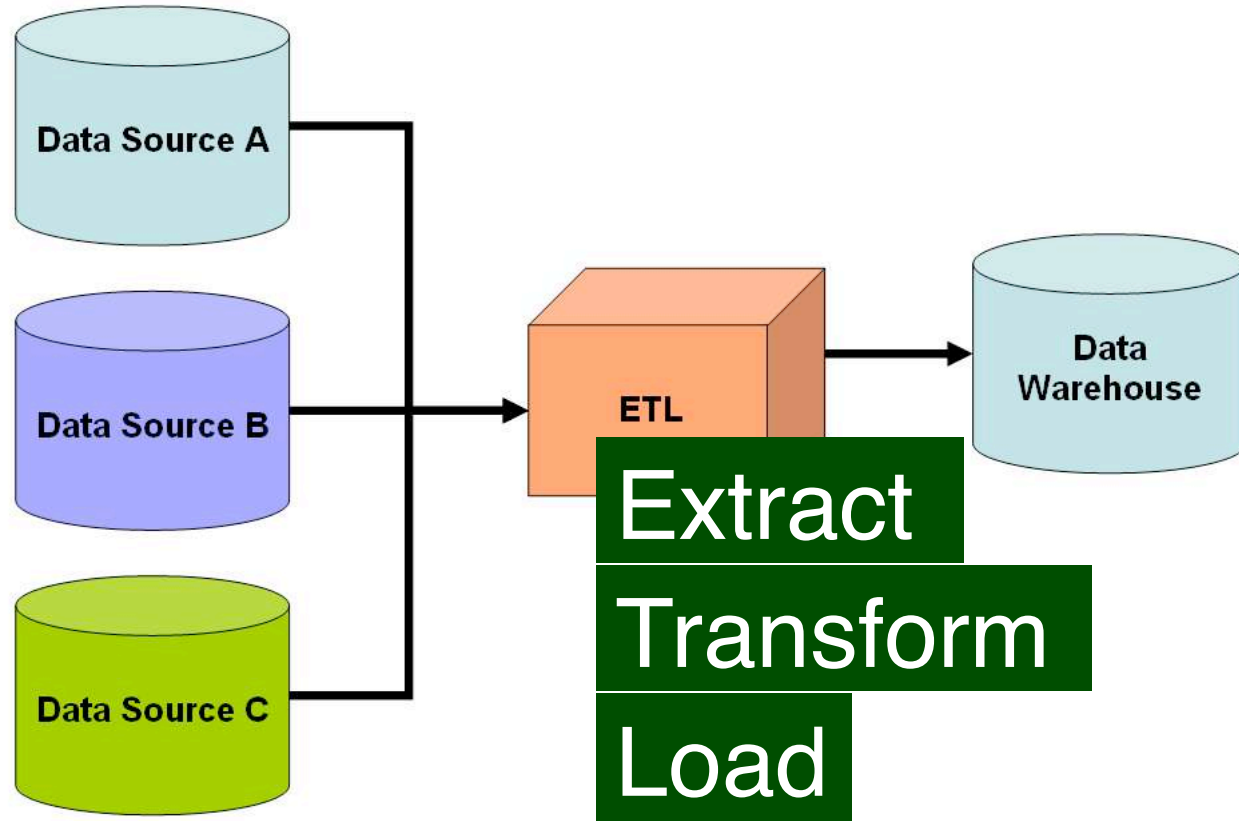
Traditional Data Warehouse



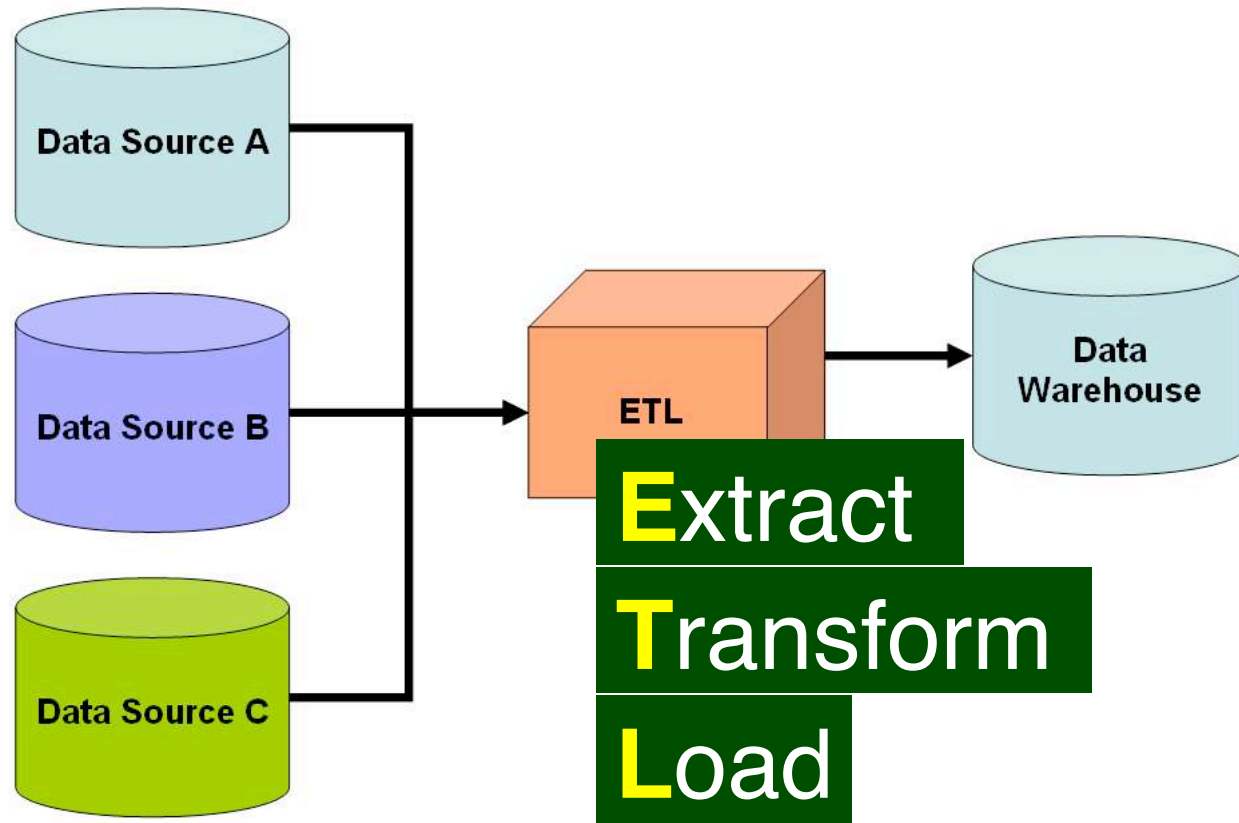
Traditional Data Warehouse



Traditional Data Warehouse



Traditional Data Warehouse



NoSQL Data Storage in the Cloud

Beyond relational databases!

Organize data to
suit the problem
and objectives!

Connections between data



Connections between data



Key-Value pairs



MULTIPLE DATA SOURCES



VALUE

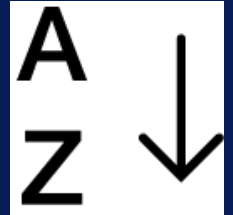
MULTIPLE DATA SOURCES



Retrieval and Storage



Pre-Processing



Analysis

VALUE

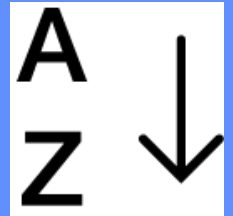
MULTIPLE DATA SOURCES



Retrieval and Storage



Pre-Processing



Analysis

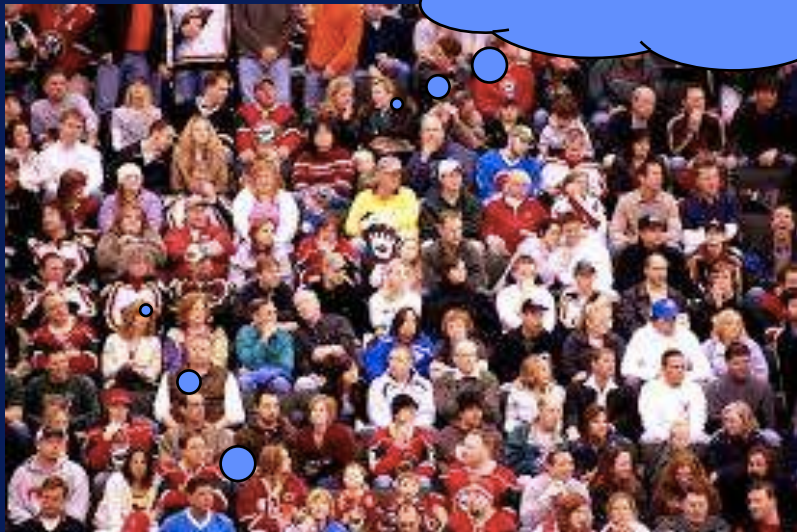
VALUE

Layers

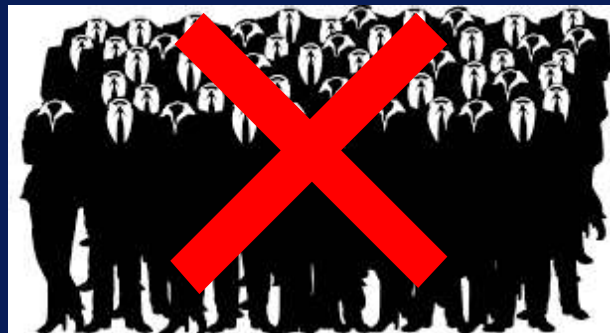
Value in Action

Sentiment Analysis

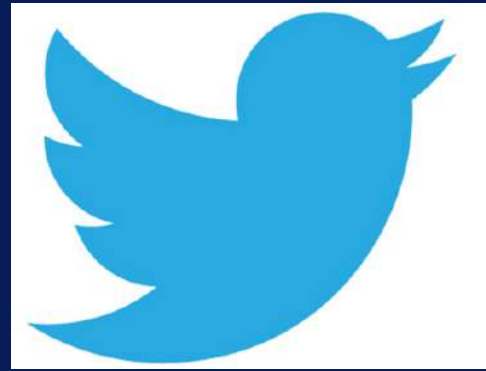
I like blue plates!



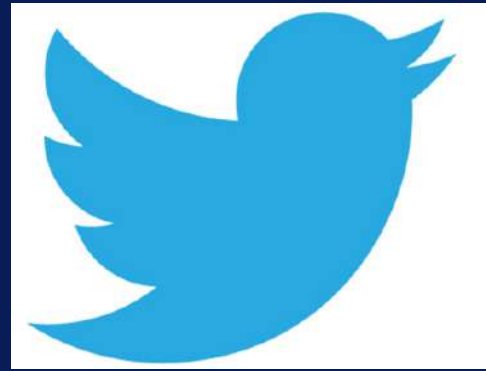
I like green plates!



Guess, how much Twitter data companies analyze everyday to measure “sentiment” around their products?



Guess, how much Twitter data companies analyze everyday to measure “sentiment” around their products?

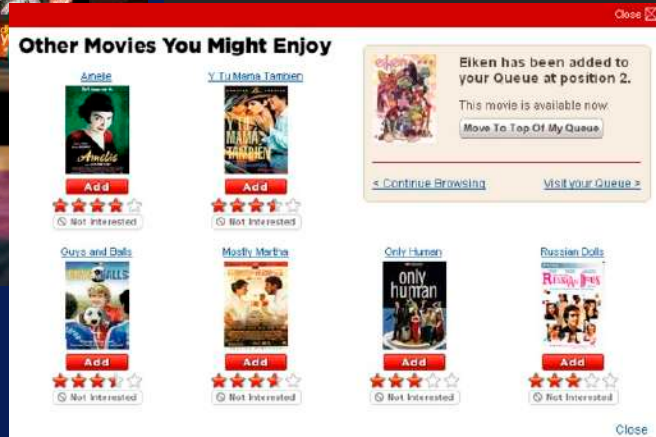


Answer: 12 Terabytes/Day

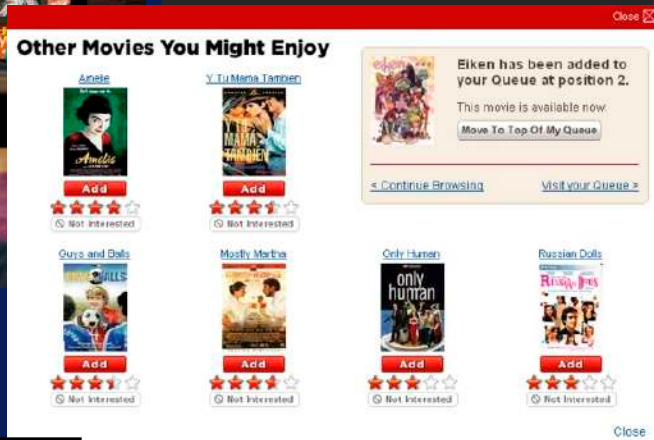
You would need to listen continuously for ~ 2 years to finish 1 TB of music !

Predict Customer Behavior

Personalized Recommendations



Predict Customer Behavior



Personalized
Recommendations

Happy Customer

More Profit



Collective Disaster Response



Nepal Earthquake

Find best
escape routes

Predict and
Evacuate

Super-storm
Sandy



Wildfire





crisismappers.net

Aerial &
Satellite
Imagery

Participatory
Maps

Crowdsourced
Event Data

Live
Twitter
Updates



Mobile and Web
Applications

Geospatial
Platforms

Statistical
Models

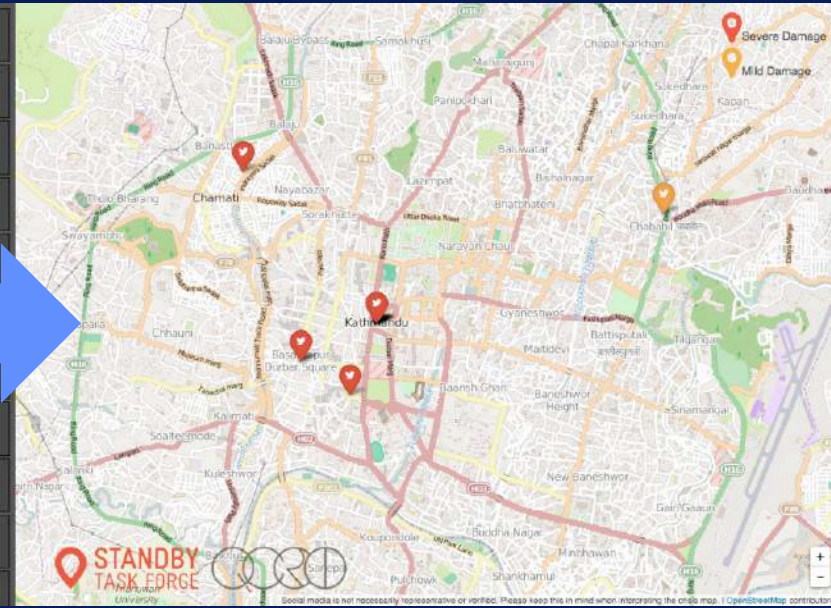
Live
Simulation

Advanced
Visualization

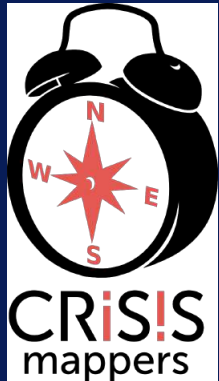


SEVERE	2015-04-28 01:40:05
SEVERE	2015-04-28 01:40:04
MILD	2015-04-28 01:40:03
MILD	2015-04-28 01:40:03
SEVERE	2015-04-28 01:39:43

MILD	2015-04-28 01:22:37
MILD	2015-04-28 01:33:06
SEVERE	2015-04-28 01:33:05
MILD	2015-04-28 01:33:06
SEVERE	2015-04-28 01:33:06



Nepal Earthquake 2015



Can you list 3 things you can do from Big Data analysis



Unstructured People
Generated Data



Emerging Big Data
Tools and Approaches



Better decisions,
Increased productivity, profits
and societal impact

