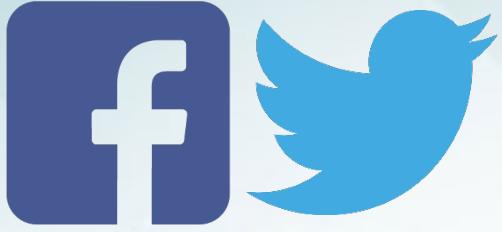


**BIG DATA**

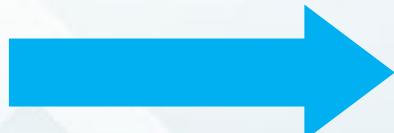
RESULTS  
SOCIAL YEARS  
DRIVE SAN NAS  
REPLICATION  
HANDLE SOURCES  
**LARGE** HAND CAPACITY  
**STORAGE**  
COMPLEX ANNOUNCED LIMITS  
PAST ONE  
MILLION SETS  
INCLUDE NETWORKS  
HYPOTHESIS SCIENCE  
END NEARLY MAPREDUCE  
FUTURE MOST PARALLEL STATE CRITIQUE USE  
CRITIQUES BILLION APPROACH NEW  
OUT DEVELOPMENT TECHNOLOGY APPROACH  
ONE  
GOVERN SYSTEMS PROCESS MAKING INTEGRATION  
VOLUME WORLD  
TECHNOLOGY FRAMEWORK BETWEEN BUSINESS EFFECTIVE  
SIZE STATISTICS VISUALIZATION  
MANAGEMENT DATABASES PROCESSING ANALYTICS  
IMPLEMENTATION SIMULATIONS  
BECOMES  
DATA FLOW NATIONAL INTELLIGENCE  
INITIATIVE ECONOMIC PROJECT ANNUAL  
TERABYTES MARKET VARIETY  
RELATED TIME SENSOR  
WOULD  
EXabytes DETERMINED SET DEPARTMENT  
HIGH USERS AMOUNT RATE  
COMPANIES ALGORITHMS REAL STRUCTURE WORK  
ANALYSIS RESEARCH LEARNING LESS HUGE  
PRIVATE EVERY WORLD VARIOUS BASED TIMES  
APPROACHES APPROACHES RECORDS COST PARADIGM FUNDING  
SEARCH INSIGHT TOOLS CENTER ARCHITECTURE  
INTERNET COMPATIBILITIES



Resource



Data



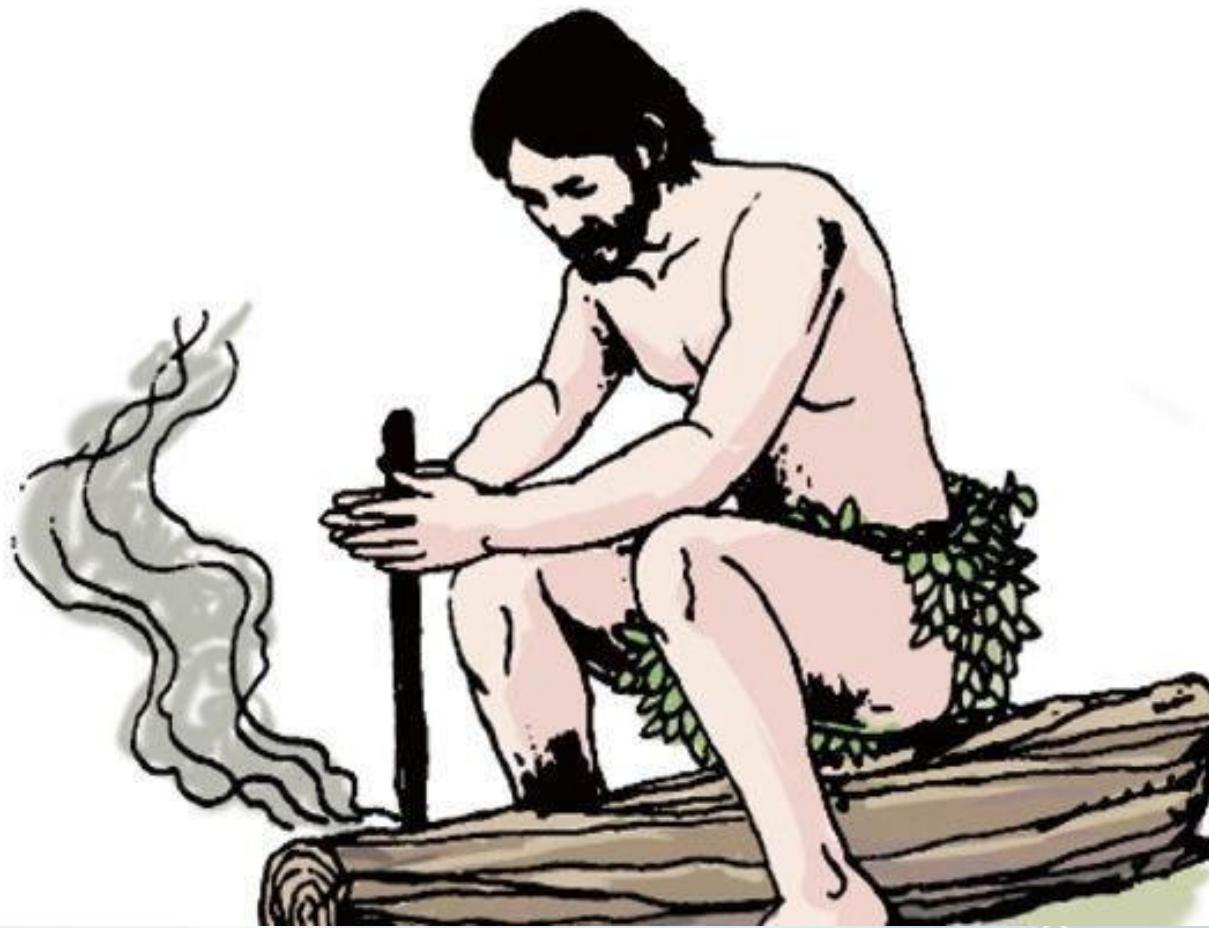
## Data Visualization



# What is the Big Data ?

Big data is data sets that are so voluminous and complex that traditional data-processing application software are inadequate to deal with them.

# What is the Big Data ?



# What is the Big Data ?

Big data (n.名詞)

無法用”傳統方法”處理的”資料們”

(DB2、Oracle、MSSQL、MySQL)

# What is the Big Data ?

Volume

大量

Velocity

快速

Variety

資料類型不定

Veracity

正確性不定

The “three V’s”, i.e the Volume, Variety and Velocity of the data coming in is what creates the challenge.

2

## VOLUME



Amount of  
Big Data  
stored  
across the  
world (in  
petabytes)

## VARIETY



### PEOPLE TO PEOPLE

NETIZENS, VIRTUAL COMMUNITIES, SOCIAL NETWORKS, WEB LOGS...



### PEOPLE TO MACHINE

ARCHIVES, MEDICAL DEVICES, DIGITAL TV, E-COMMERCE, SMART CARDS, BANK CARDS, COMPUTERS, MOBILES...



### MACHINE TO MACHINE

SENSORS, GPS DEVICES, BAR CODE SCANNERS, SURVEILLANCE CAMERAS, SCIENTIFIC RESEARCH...

## VELOCITY



**2.9**  
MILLION

EMAILS  
SENT EVERY  
SECOND



**20**  
HOURS

OF VIDEO  
UPLOADED  
EVERY MIN

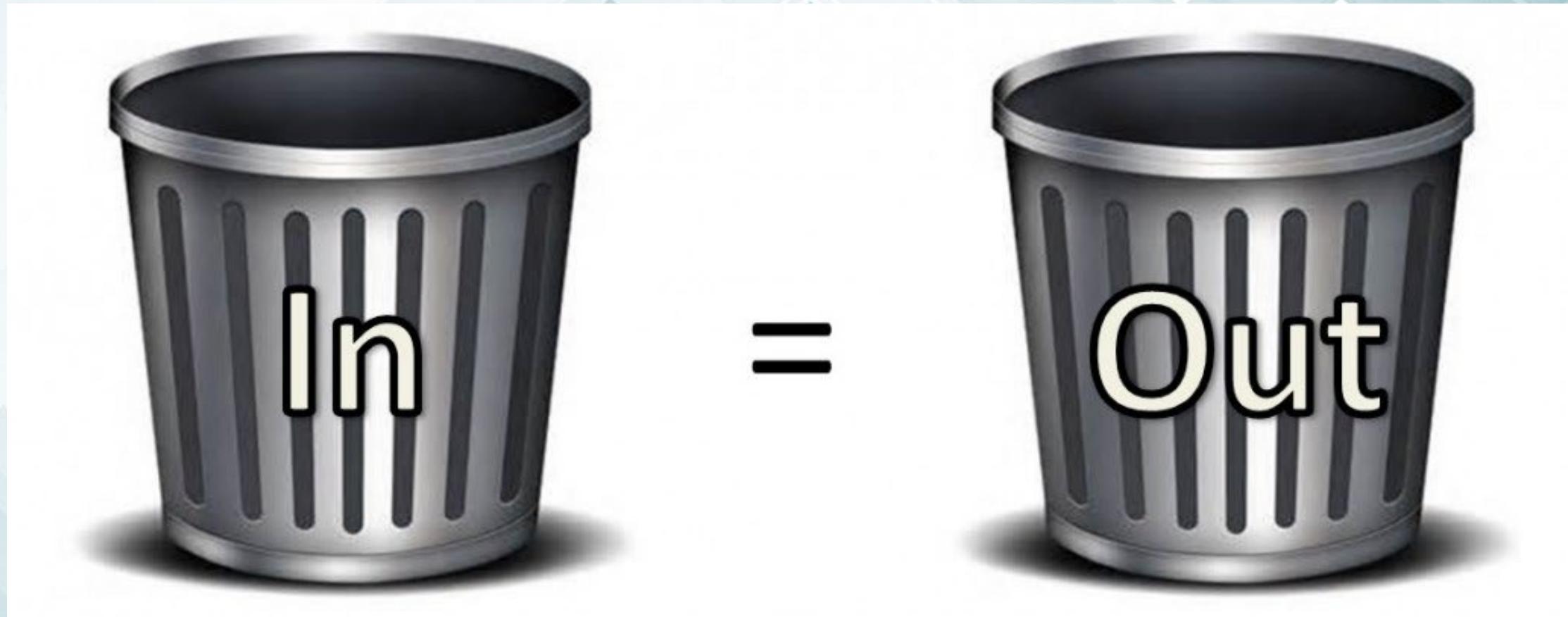


**50**  
MILLION

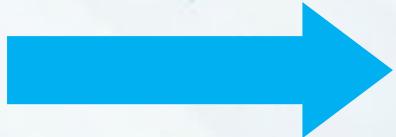
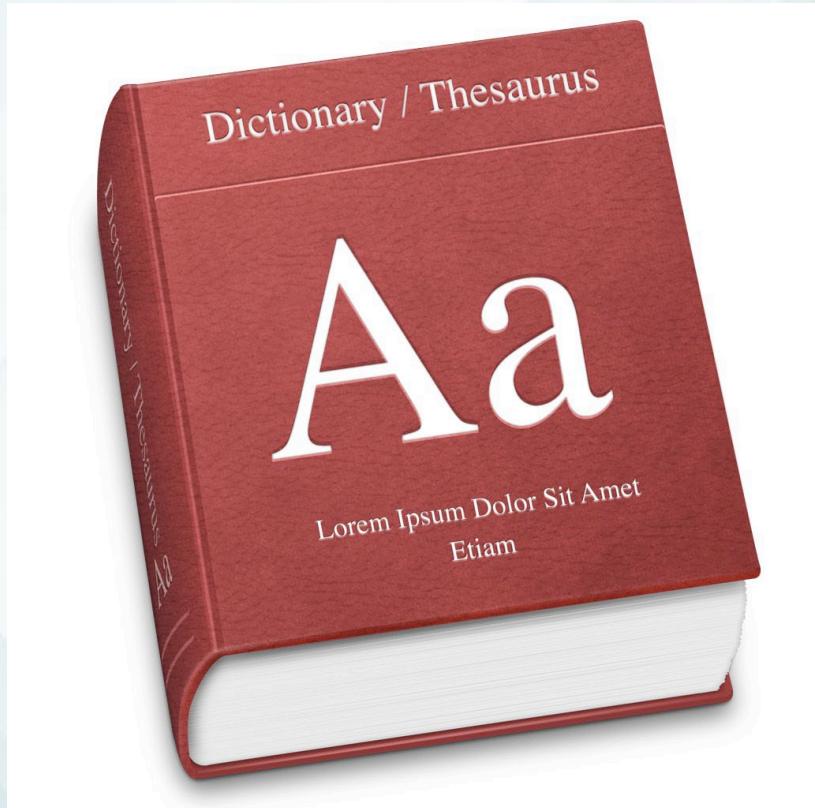
TWEETS  
PER DAY

3

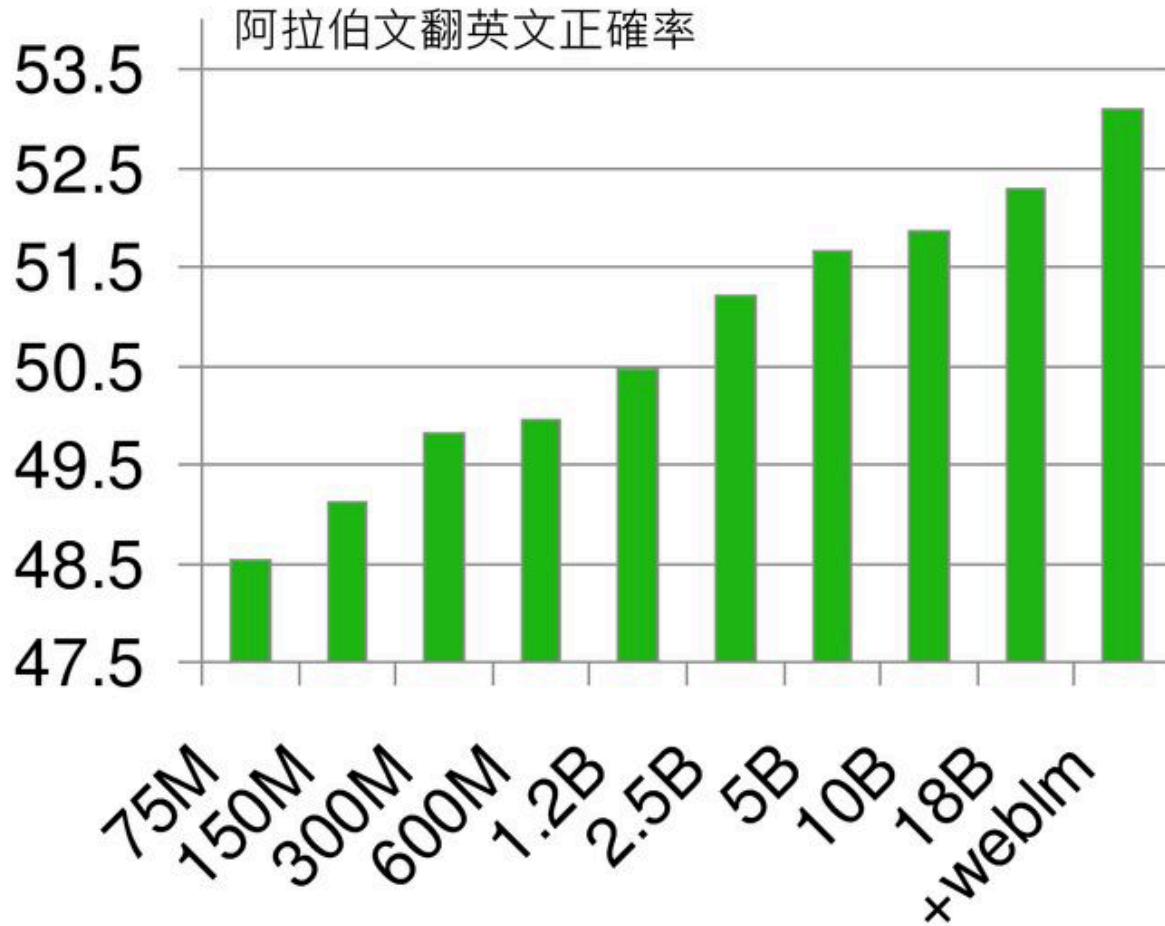
# Gaebage IN Garbage OUT?



# Example 1 : translation



# Example 1 : translation



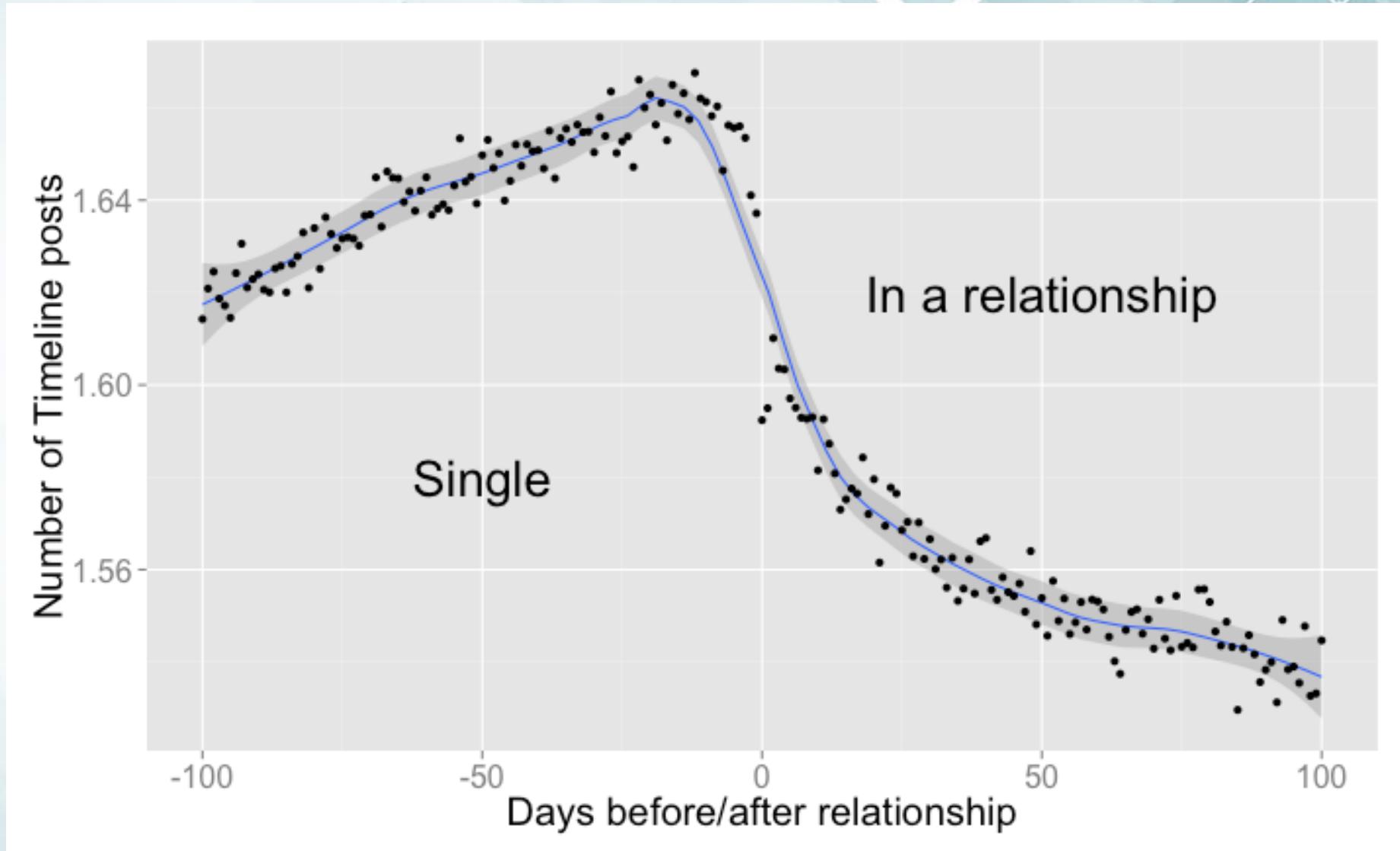
**+webLM =**  
219B words of  
web data!

# What is the Big Data ?

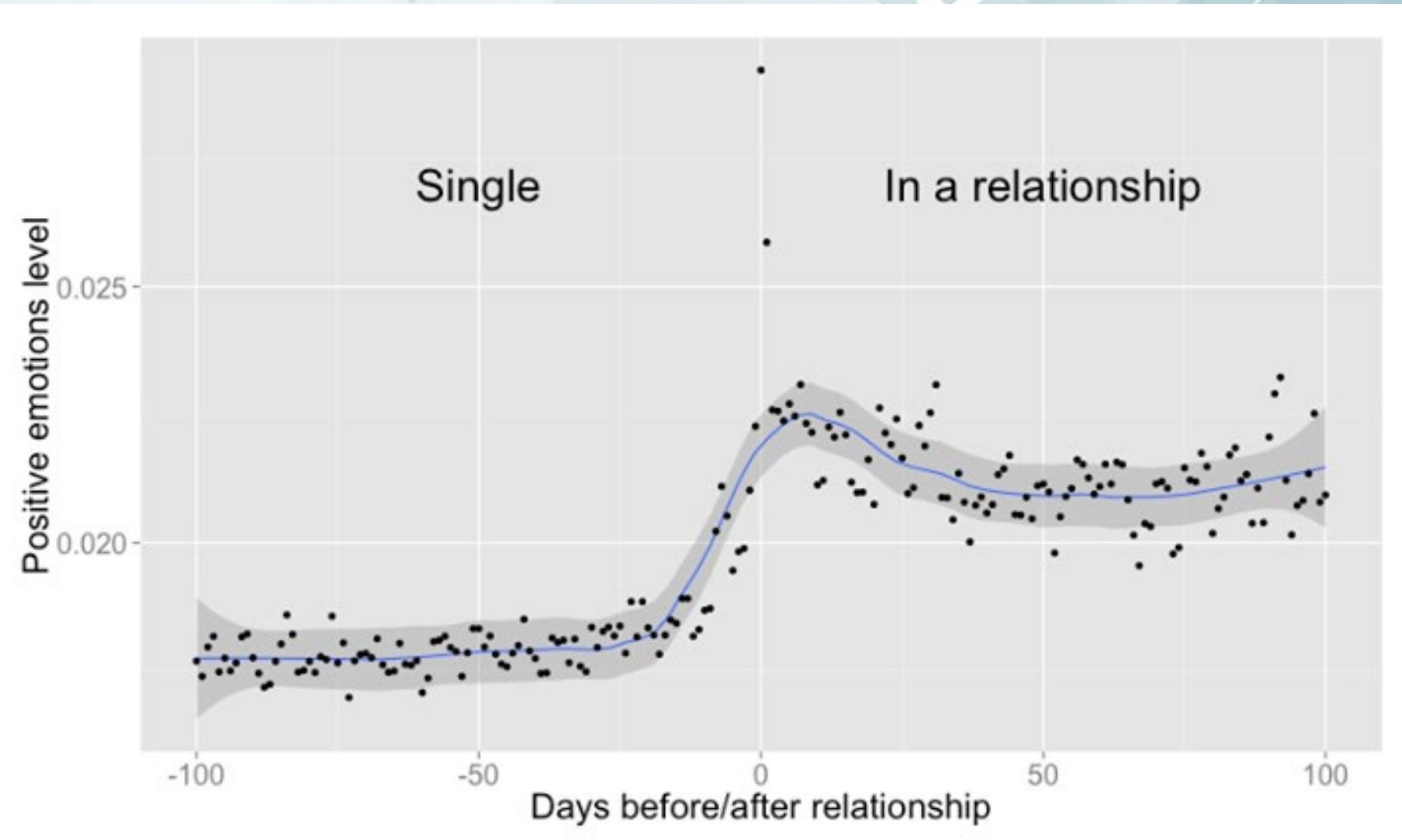
Big data (n.名詞.延伸)

了解如何挖掘、設計分析，  
從資料中提取出價值的技術。

# Example 2 : relationship



# Example 2 : relationship



# Level of the Big Data

What you can **save** is storage.

What you can **see** is Data.

What you can **understand** is information.

What you can **use** is intelligence.

# **Who can do the Big Data ?**

**擁有Big Data的人**

**會寫程式分析Big Data的人**

**知道Big Data價值何在的人**

**And, Who earn the most ?**



**So, What can we do ?**