Big Data

SUKHJIT (BOB) SINGH

SINGHSUKHJIT@DEANZA.EDU
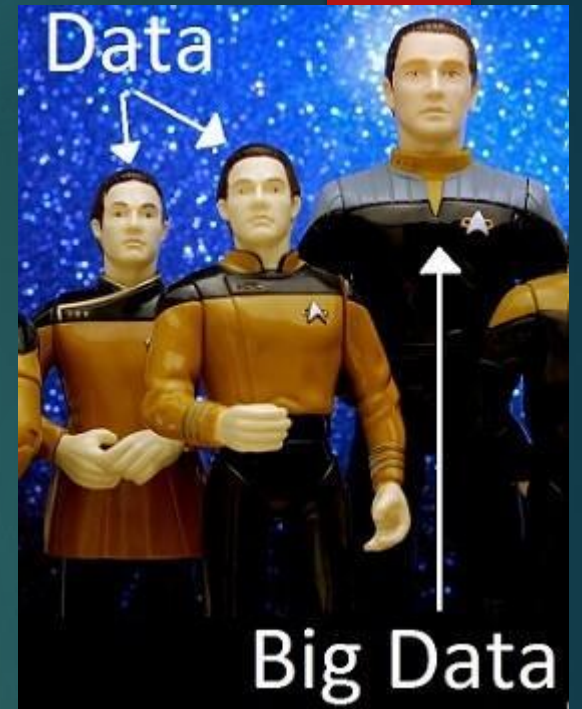
Post your questions here
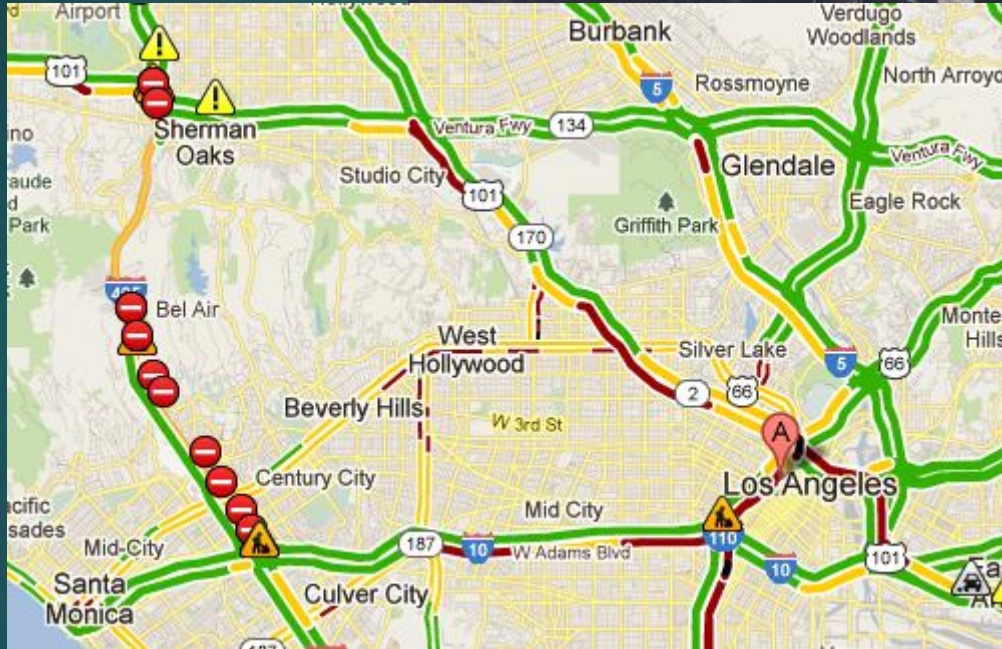
- https://todaysmeet.com/ACMClubBigData

Agenda

- What is it?

- Why we need it?

- How it works?

- What can you learn at De Anza?

- How can you use it in the world?

What is Big Data

"extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions."

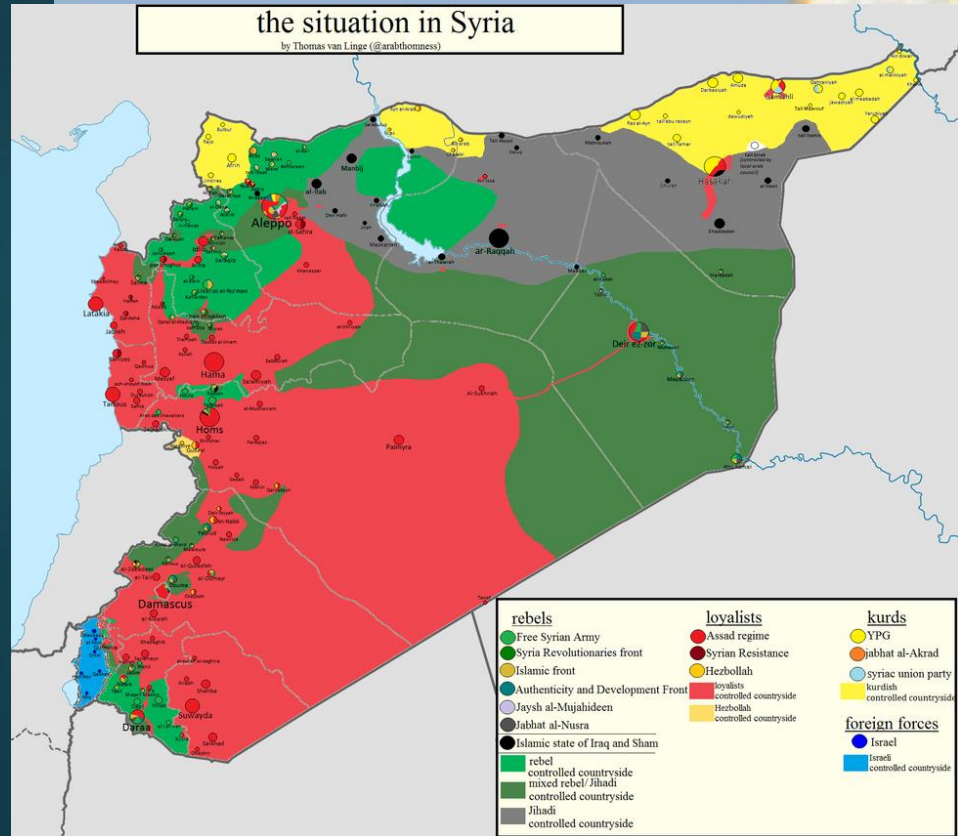# What is Big Data

What is Big Data

all geocoded tweets in October 2012
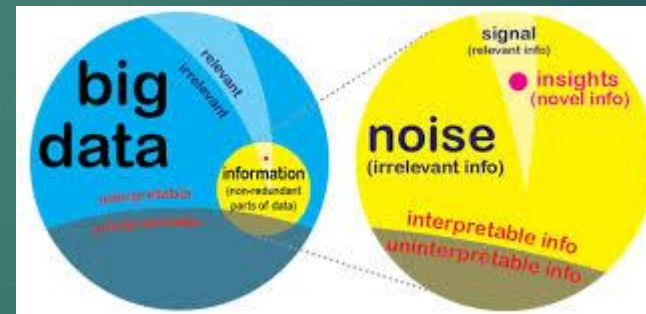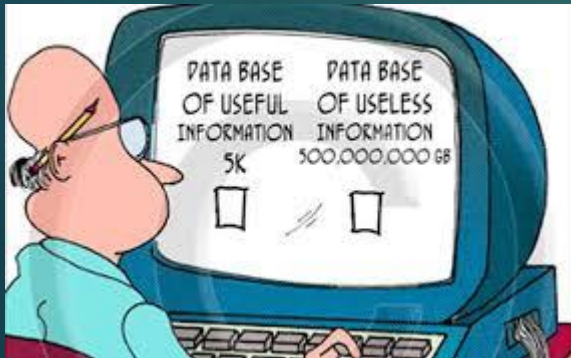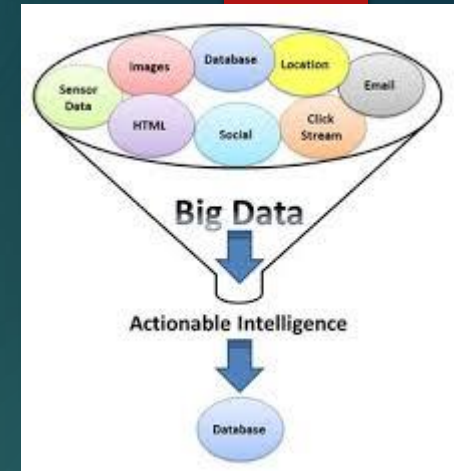
What Is Big Data?

What is Big Data?

http://tweetping.net/

Why do we need it?

- Because it's here! Dude
- Data Driven (Communism) versus Data Informed (Democratic)
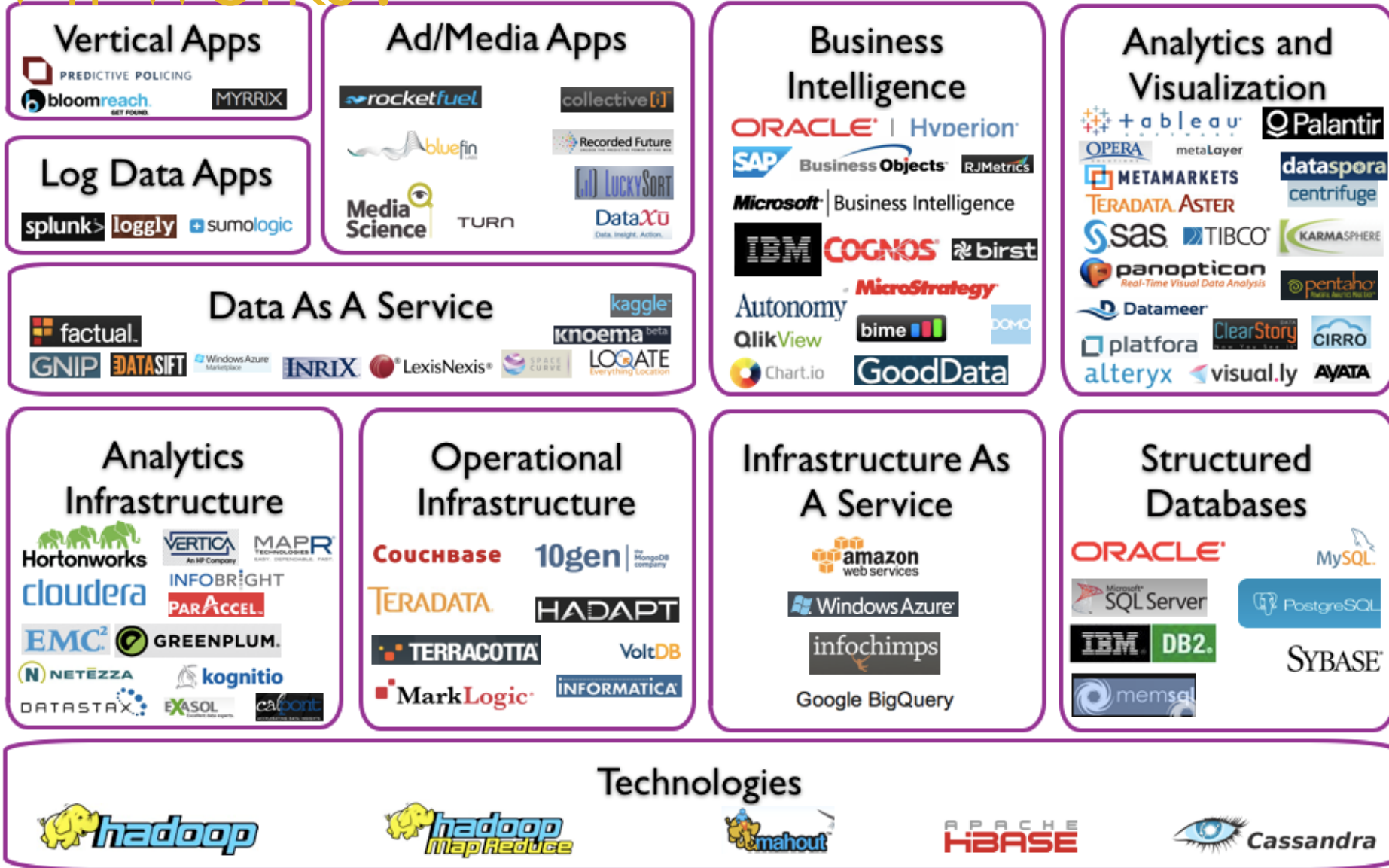- Usability
- ….Smarter Human Race
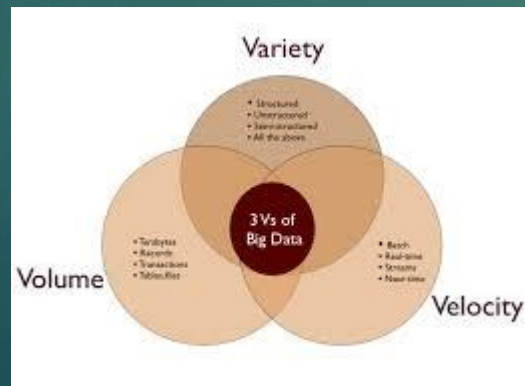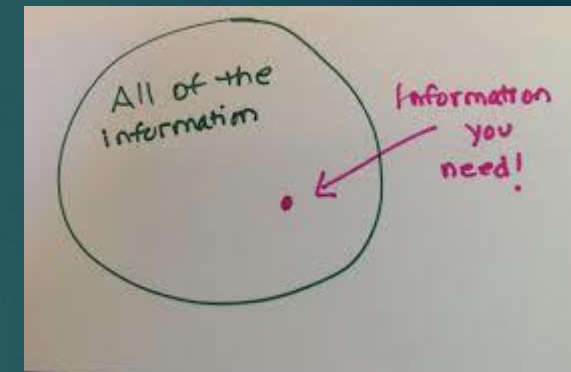
Why do we need it?

How it works?

Big Data Landscape

Copyright © 2012 Dave Feinleib   dave@vcdave.com   blogs.forbes.com/davefeinleib
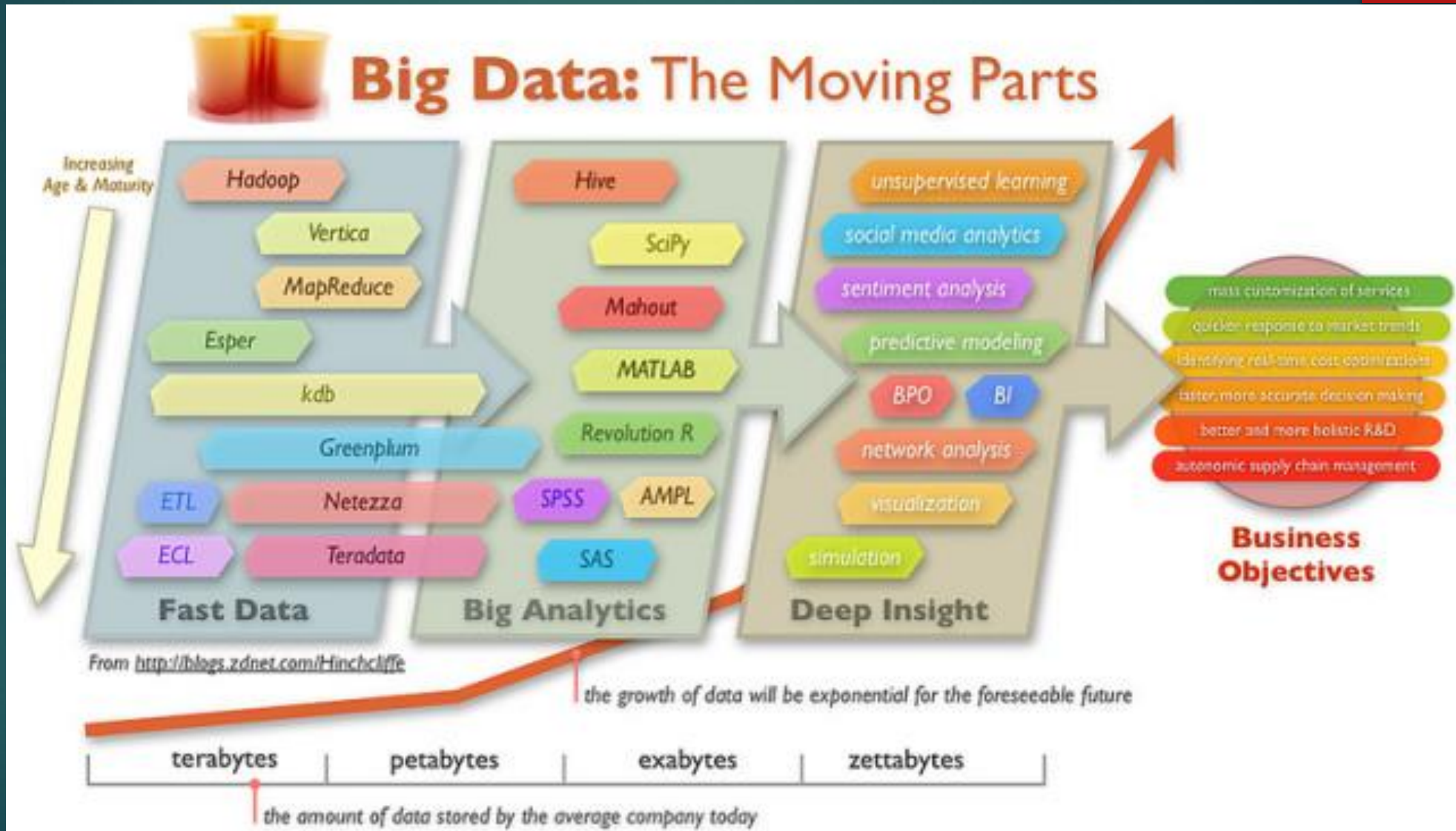
How it works?

How it works?

Count the number of words in the following paragraph.
Make sure you do not count "Stop Words" (the, a, is)

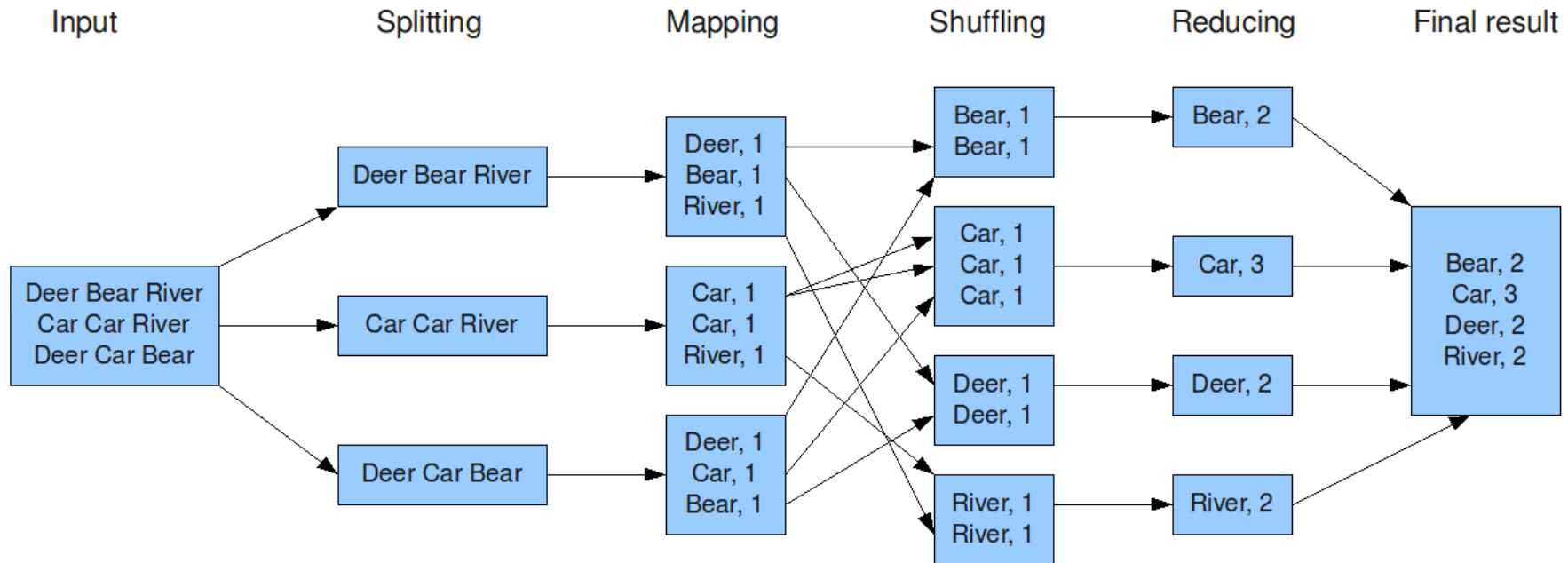**The quick brown fox jumps over the lazy dog**
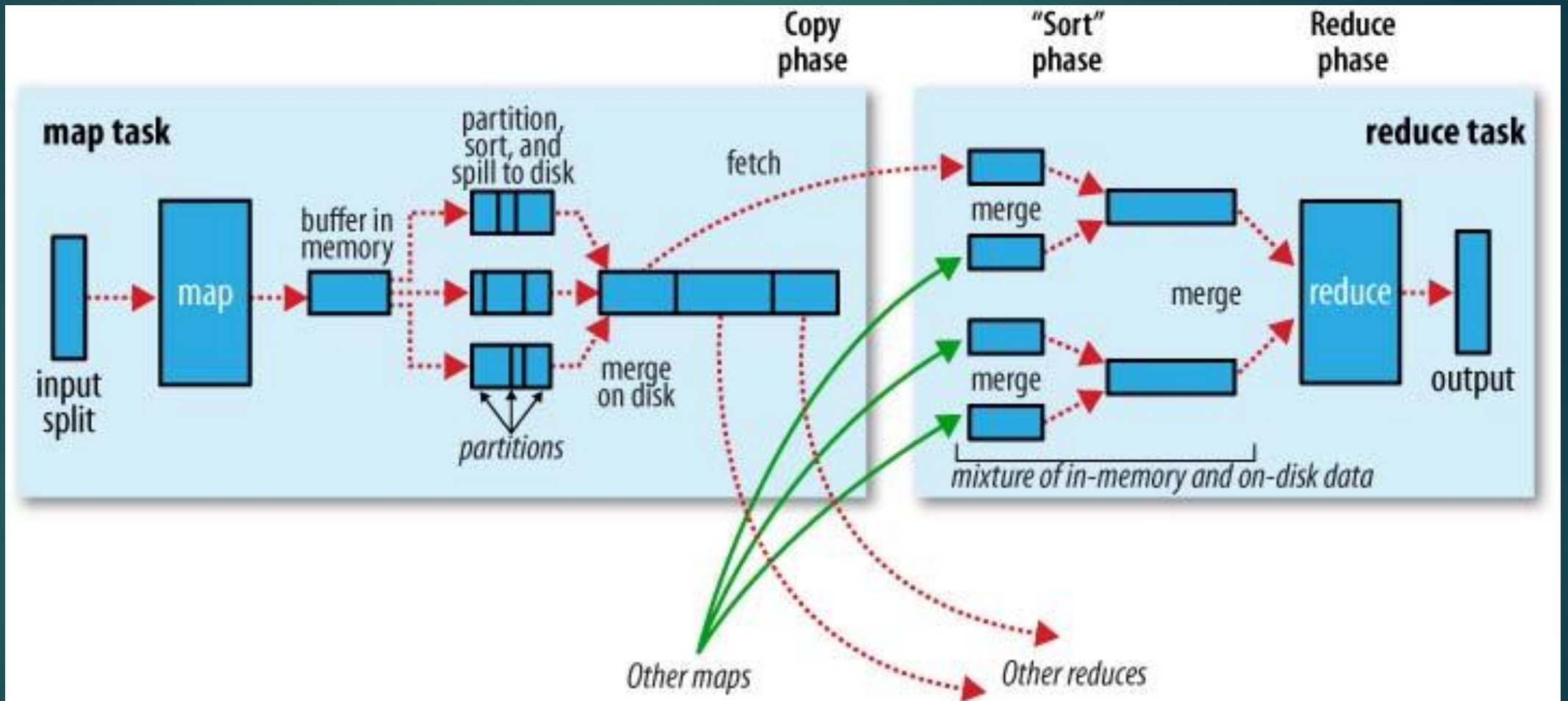**A dog is a man's best friend**

quick	=> 1
brown	=> 1
fox	=> 1
jumps	=> 1
over	=> 1
lazy	=> 1
dog	=> 2
mans	=> 1
best	=> 1
friend	=> 1

How it works?

The overall MapReduce word count process

How it works?

What can you learn at De Anza?

Google - De Anza Database Design

http://www.deanza.edu/counseling/pdf/degrees/cis2-database.pdf

How can we use it @ De Anza?

Questions

THANK YOU