

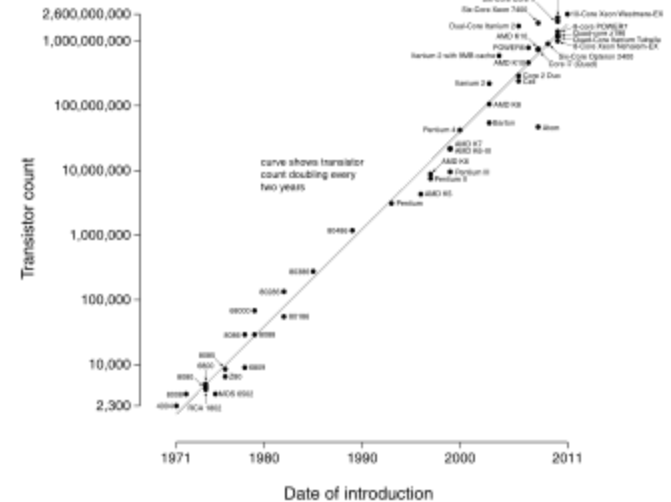
Big Data ?

- lots and lots of web pages ...
- a billion Facebook users
- billion+ Facebook pages
- hundreds of million Twitter accounts
- hundreds of million tweets per day
- Billions of Google queries per day
- Millions of servers, petabytes of data

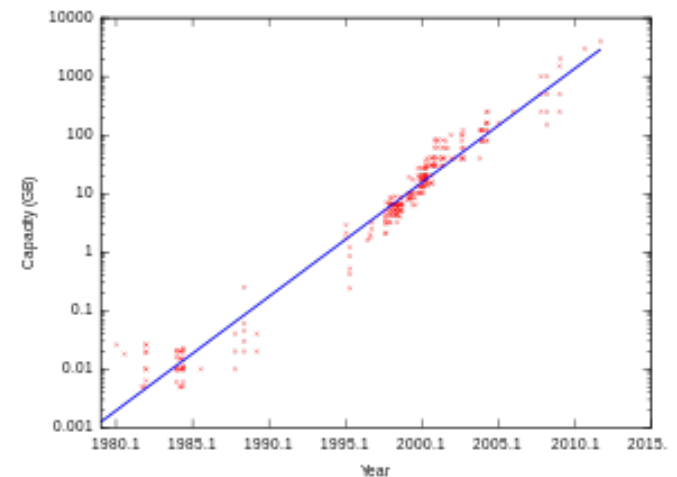
In contrast, typical large enterprise:

- ❑ 5000-50,000 servers,
- ❑ Terabytes of data, millions of Txn/day

Microprocessor Transistor Counts 1971-2011 & Moore's Law



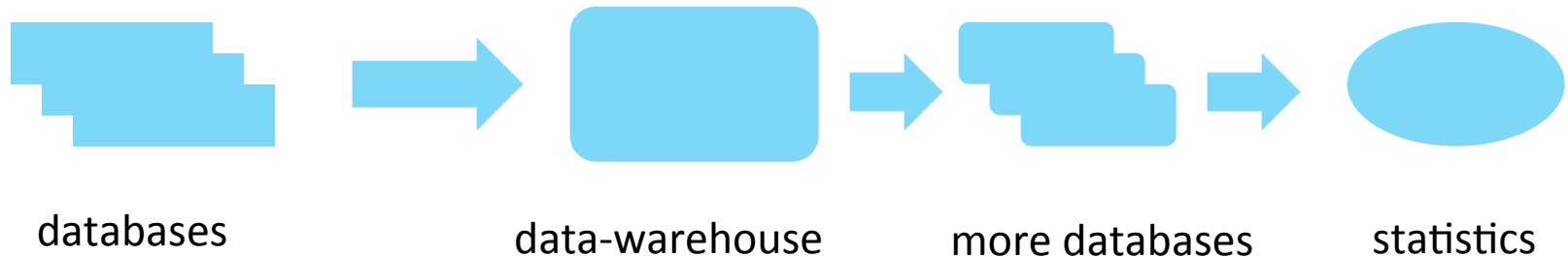
Moore's Law



Kryder's Law

Big-Data *technology*

traditional `business intelligence' using databases:

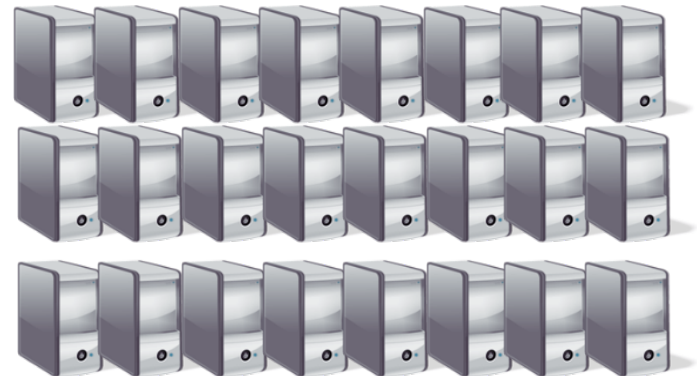


Google, Facebook, Linkedin, eBay, Amazon ...
did *not* use `traditional' databases for `big data'

why?

what?

- *massive parallelism*
- *Map-Reduce paradigm*

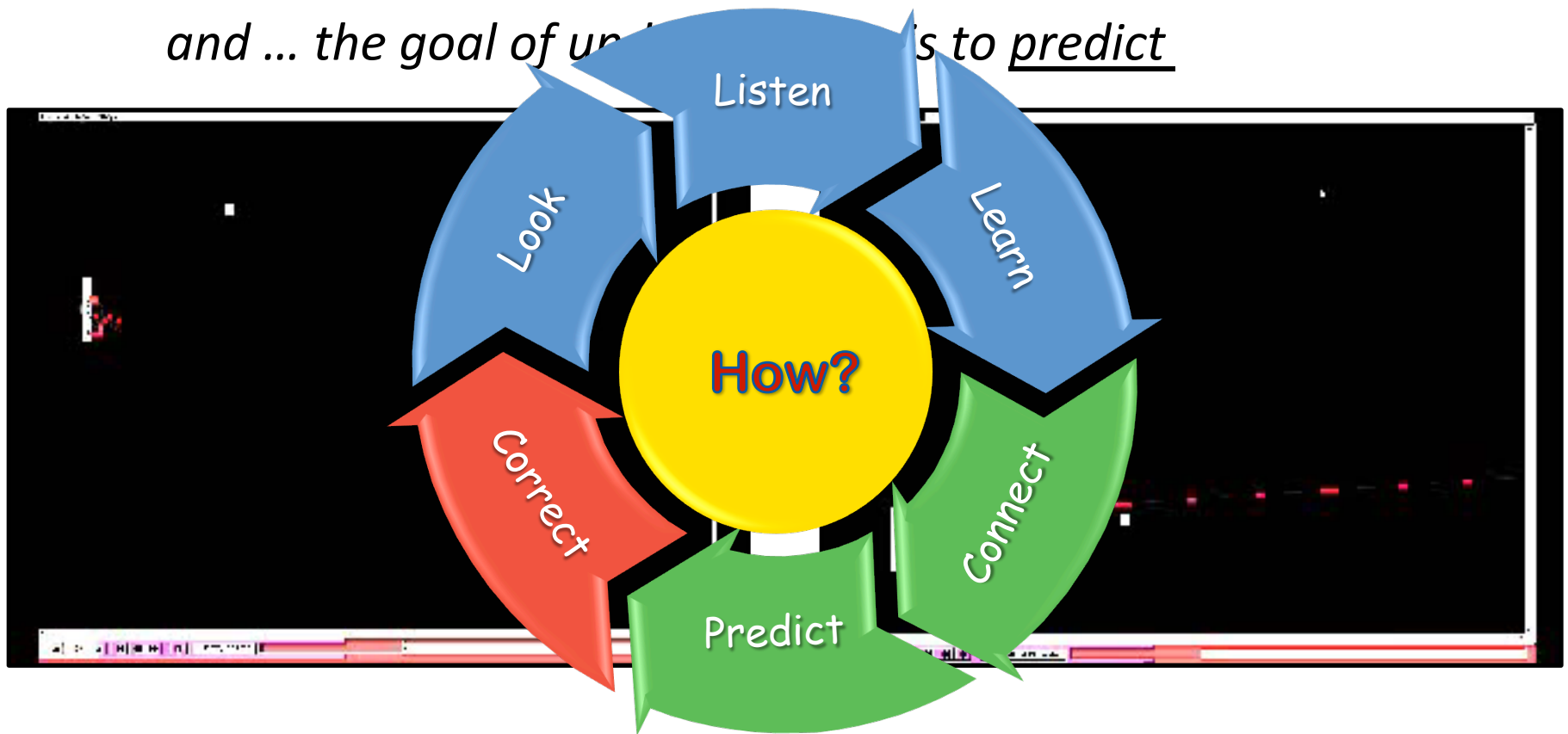


what does *data* have to do with *intelligence*?

“any fool can know ... the point is to understand.”

- Albert Einstein

and ... the goal of understanding is to predict



Reactive Intelligence

Predictive Intelligence

web intelligence using big data

AI techniques at web-scale for `predictive intelligence`

- online advertising – predicting intent and interest
- gauging consumer sentiment and predicting behavior
- detecting adverse events and predicting their impact
- intelligent question answering such as in Watson
- categorizing and recognizing places, faces, people, ..
- personalized genomic medicine of the future
- building more intelligent public services: energy, water
- securing ourselves better

big data analytics

exploiting more efficient technology developed by web companies for their web-intelligence tasks

fusing social intelligence and business intelligence

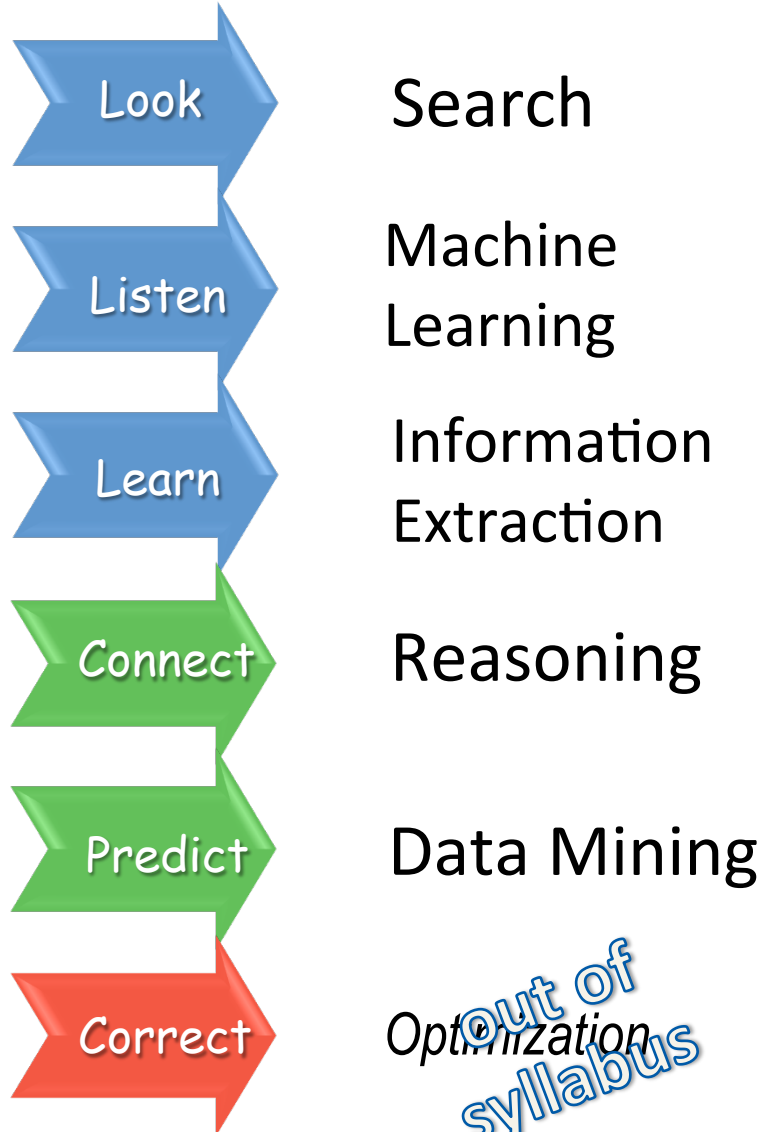
web-intelligence techniques on a mix of private and web data

- sales and marketing
- intelligent supply chains
- digital, mobile, *data-driven* business models and processes

“brick-and-mortar firms emulating web companies”

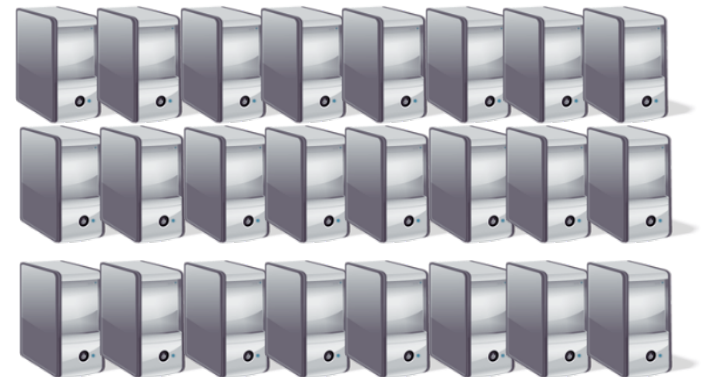
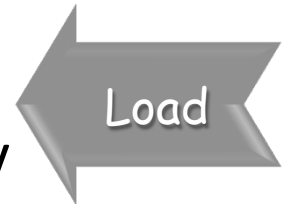
Web Intelligence and Big Data

“predict the future using AI and big data”



out of syllabus

Big Data Technology



parallel programming
using map-reduce