UCLouvain

FIAL
Un éventail de formations, une multitude d'horizons

Big Data

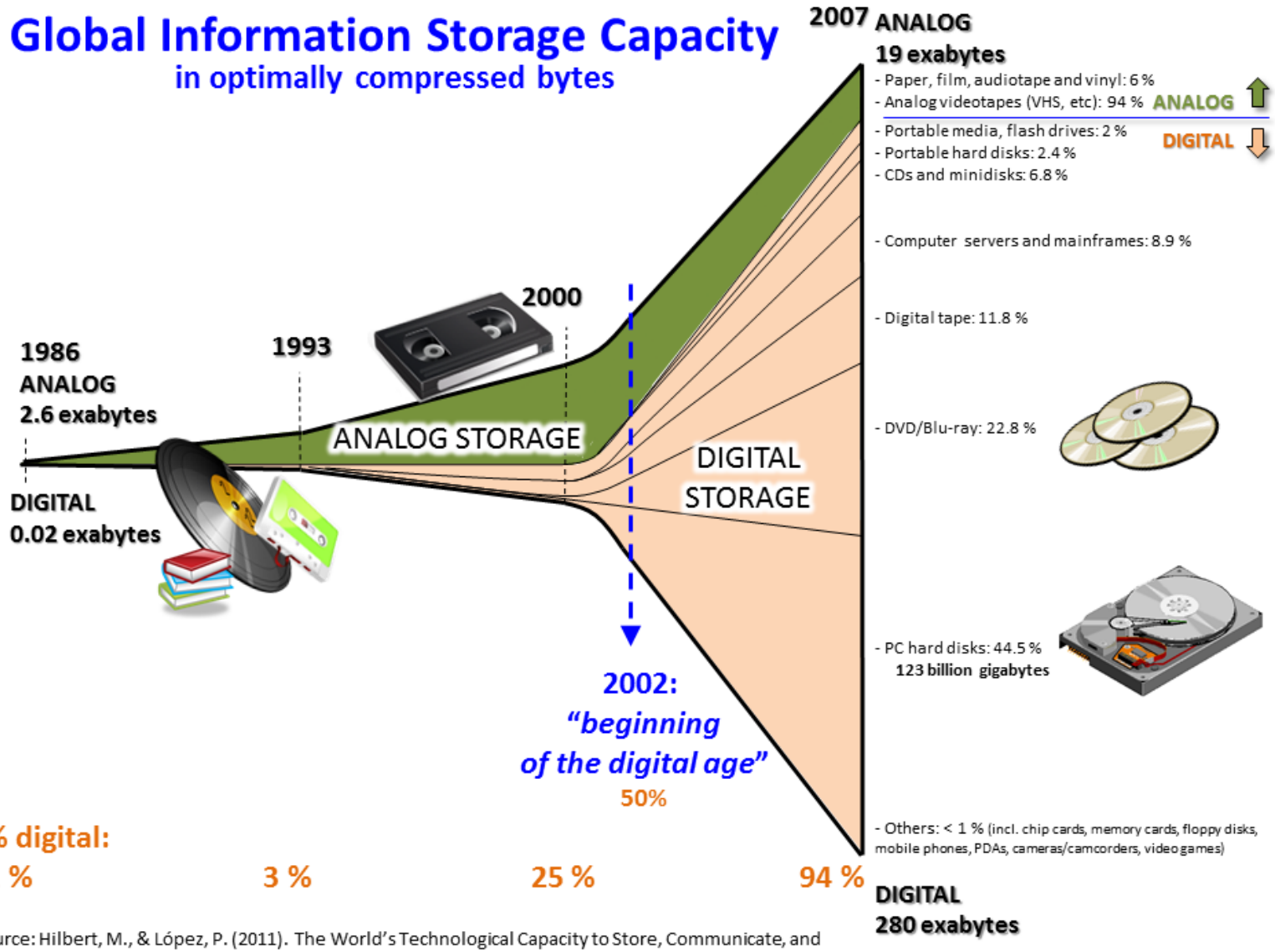
LFILO 2970
Class 4

Drivers of change

Technological drivers of change:

- Big data everywhere, even accessible while mobile
- Behavioral targeting and analytics
- Encryption, anonymity

Global Storage Capacity

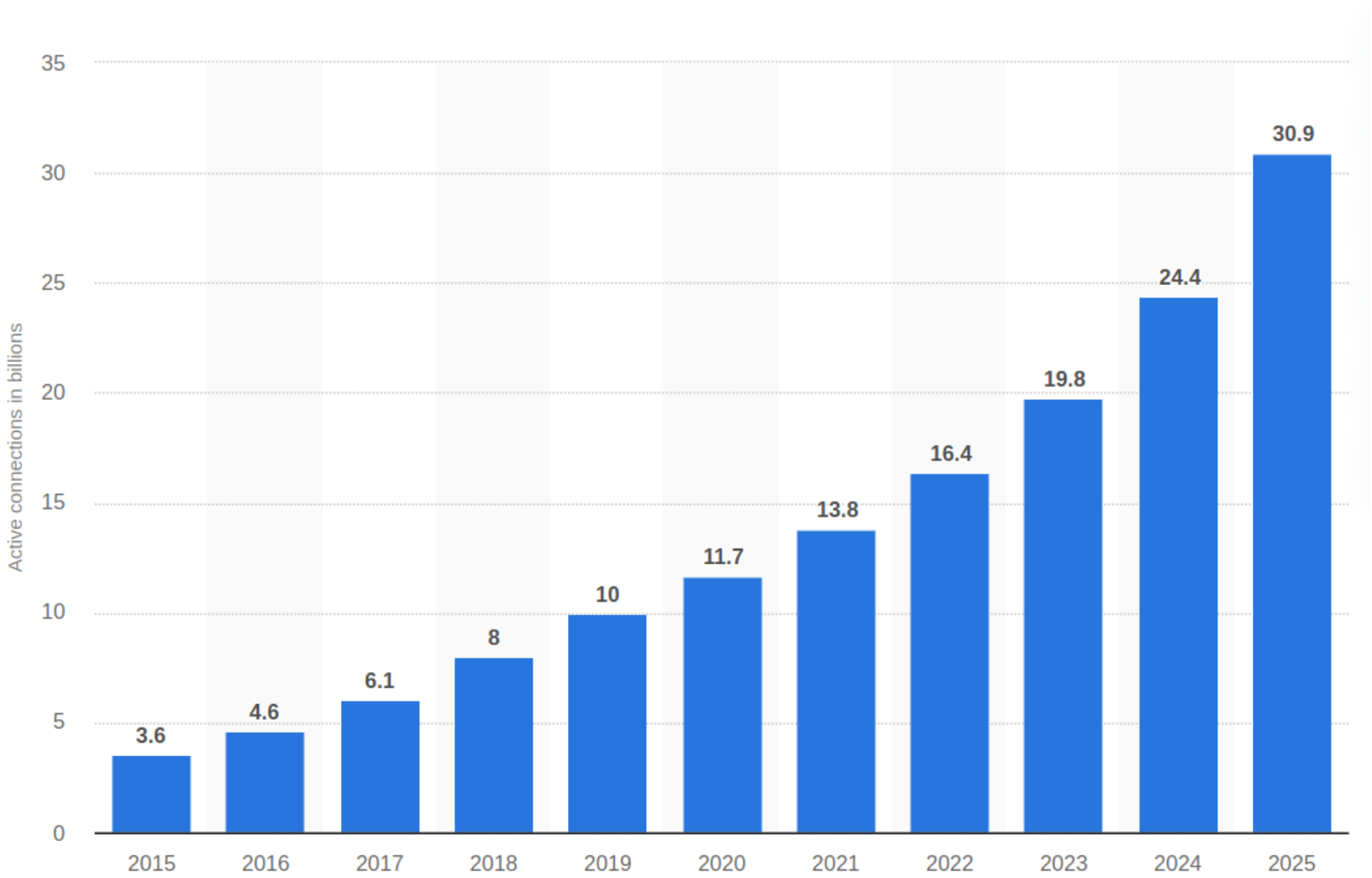


Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60 –65. http://www.martinhilbert.net/WorldInfoCapacity.html

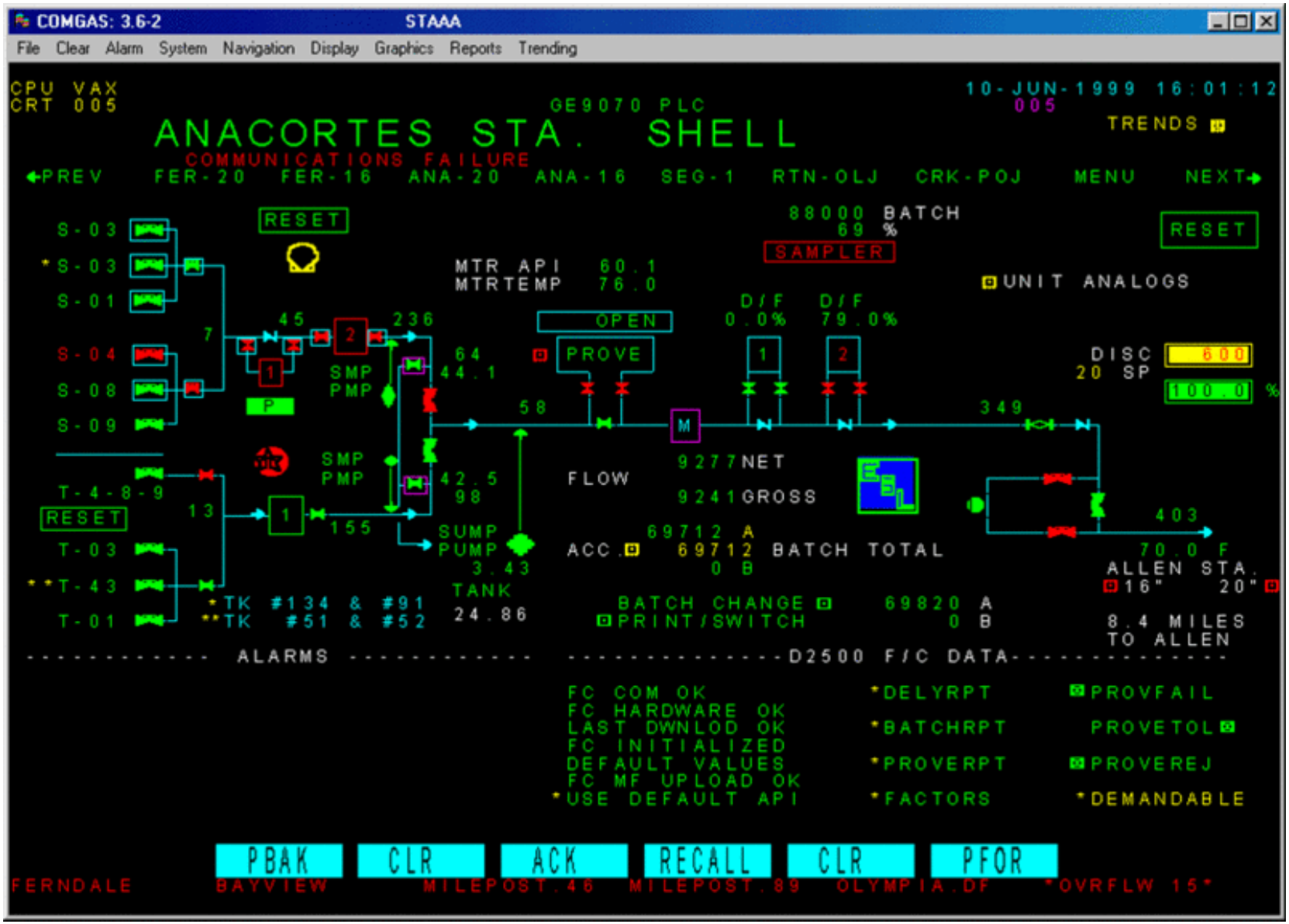
Drivers of change

- Mobile space
  - 4G and soon 5G
  - Ubiquitous location tracking
  - Extensive use of apps
- Internet of things
  - RFIDs embedded everywhere
  - Smart power grid, smart appliances
  - Robotics

Internet of Things



Control Systems



License Plate Readers

Self-Driving Cars

Drivers of change

- The human body
  - Genetic data
  - Biometrics (Fitbit, Apple Watch)
  - Implants
  - Aging population
- Data-mining tools
  - Machine learning, pattern recognition
  - AI

Drivers of change

Economic drivers of change:

- The cost of keeping data is now lower than the cost of deleting it
- The cost of collecting everything you can is now lower than the cost of being selective

Changes in “Data”

- From questions of *creation* to questions of *access* and *attribution*
- The death of *practical obscurity* and *automatic anonymity*
- Routinized, continuous surveillance
- Data that knows more than we do

What do we want?

- **secrecy** or **privacy** — keeping your data unknown (hiding the content)
- **anonymity** — keeping your data unattributed (hiding the author/attribution)
- **autonomy** — keeping your data from constraining your future opportunities (keeping data from being used against you without due process)

What are the threats?

- **identification** – linking data to a person, impacts **autonomy** but doesn't necessarily impact **secrecy** (without a breach)
- **surveillance** – impacts **privacy,** but doesn't necessarily impact **anonymity** (without identification)

Big Data and Science

From boyd and Crawford:

1. Automating research changes the definition of knowledge itself
2. Claims to *objectivity* and *accuracy* based on data can often be misleading
3. Bigger data aren't always better data
4. Not all data are equivalent
5. Just because data are accessible doesn't mean they're ethically accessible
6. Limited access to that data creates new (scholarly) digital divides

How did this happen?

Zuboff's guiding question: **Even if** all of the technological and economic drivers are there, we still **don't have to** structure society around this kind of data-collection and use. Why did we do it?

Her claim: Big data is the **product** of a new kind of economic order, which she calls **surveillance capitalism**

Surveillance Capitalism

A few important features:

- **formal indifference** — Google doesn't care what people say or do, as long as it can be tracked and monetized
- A relationship that is fundamentally **one-way,** because you don't get anything for it