# CLOUD COMPUTING APPLICATIONS

BIG DATA

Roy Campbell & Reza Farivar

# Big Data (a Singular Phrase)!

- A collection of data sets so large and complex, it's impossible to process it on one computer with the usual databases and tools

- Because of its size and complexity, Big Data is hard to capture, store, copy, delete (privacy), search, share, analyze, and visualize
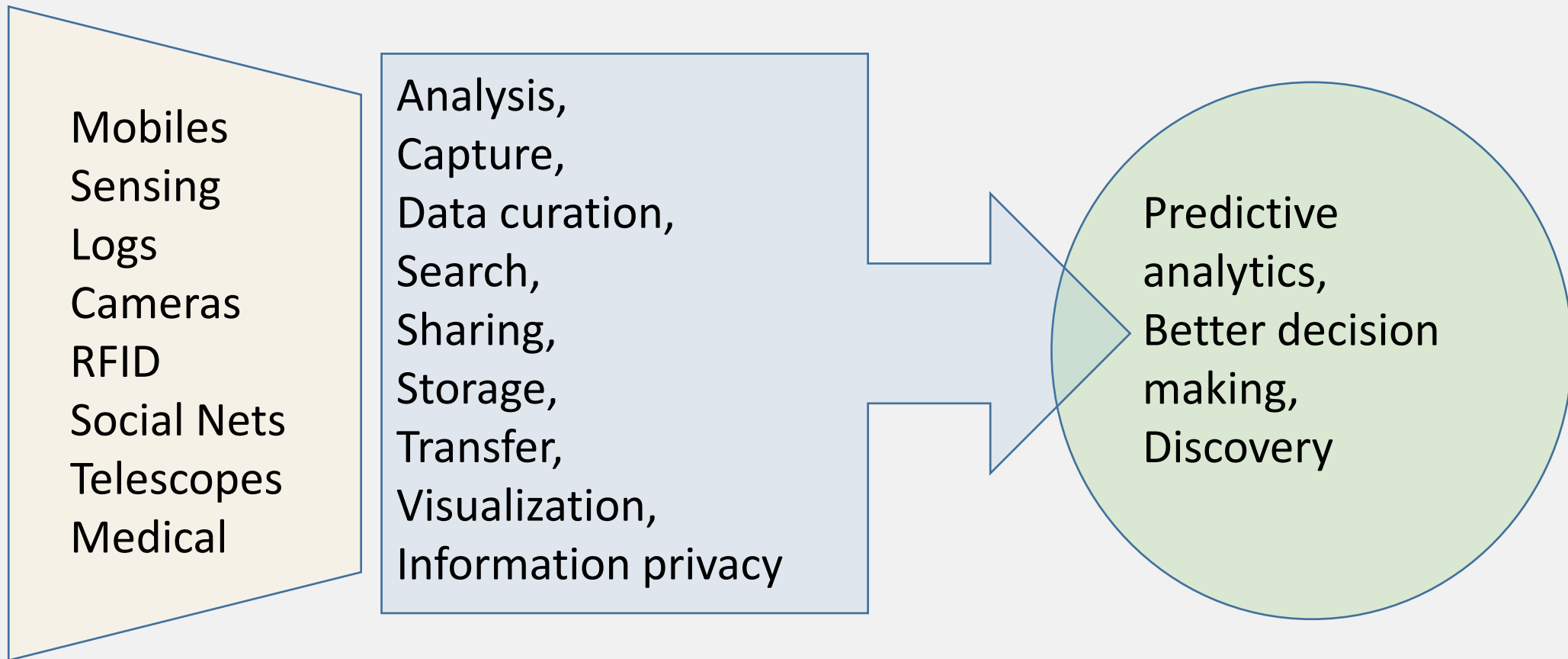
# Big Data

Big Data represents the information assets characterized by such high
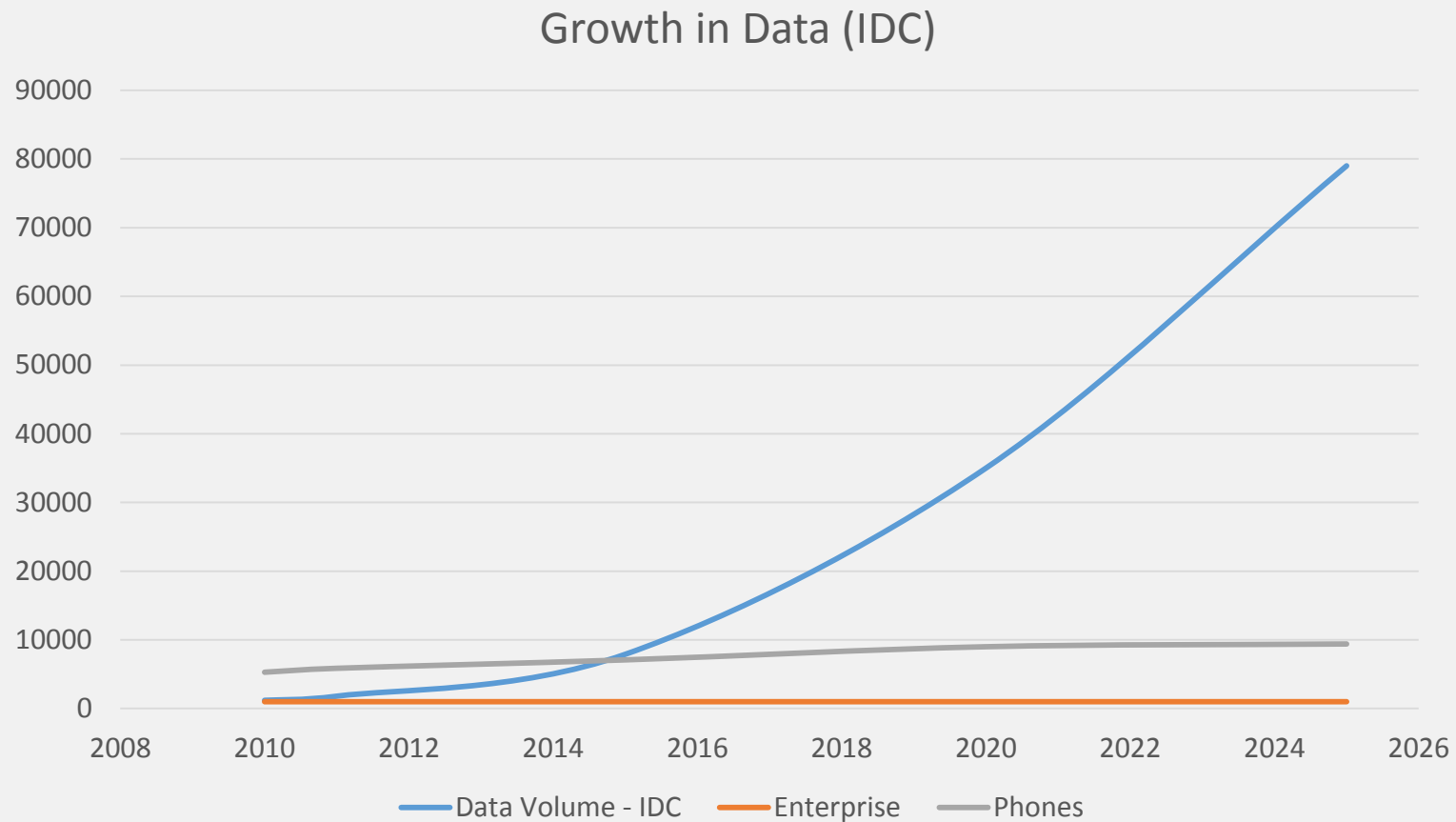
- Volume,
- Velocity, and
- Variety

as to require specific technology and analytical methods for its transformation into
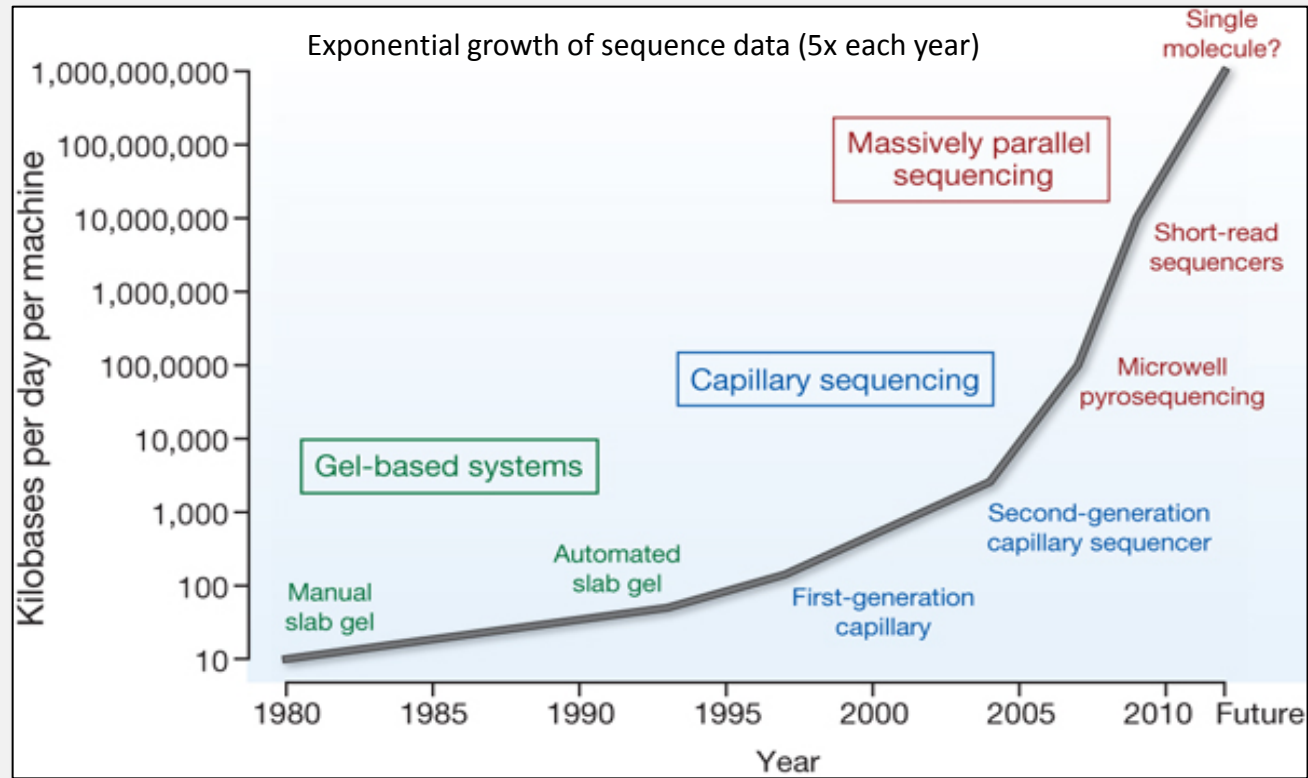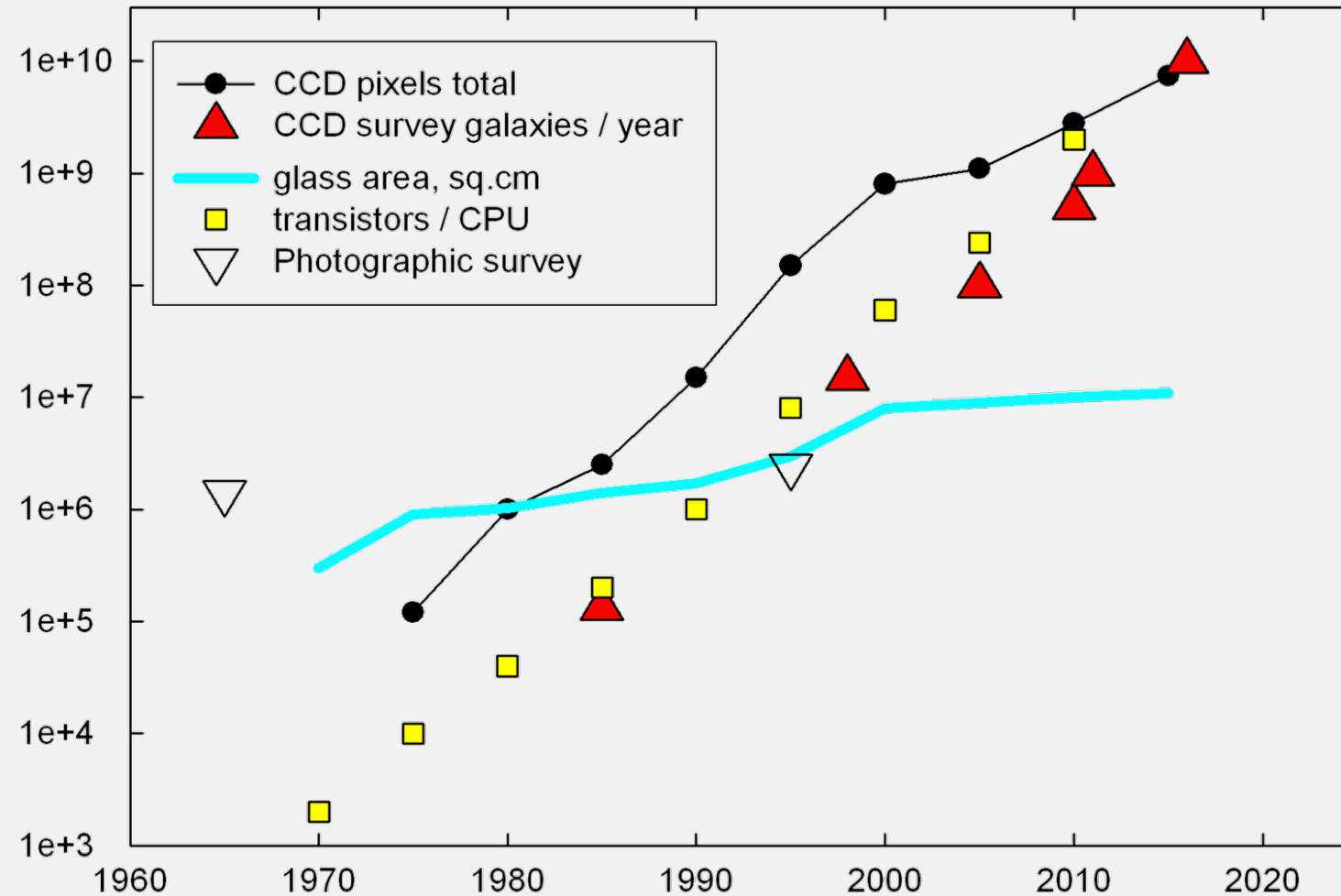
- Value

# Challenges

Mobiles
Sensing
Logs
Cameras
RFID
Social Nets
Telescopes
Medical

Analysis,
Capture,
Data curation,
Search,
Sharing,
Storage,
Transfer,
Visualization,
Information privacy

Predictive analytics,
Better decision making,
Discovery

# Timeline



Growth in Data (IDC)

# Example: Bioinformatics
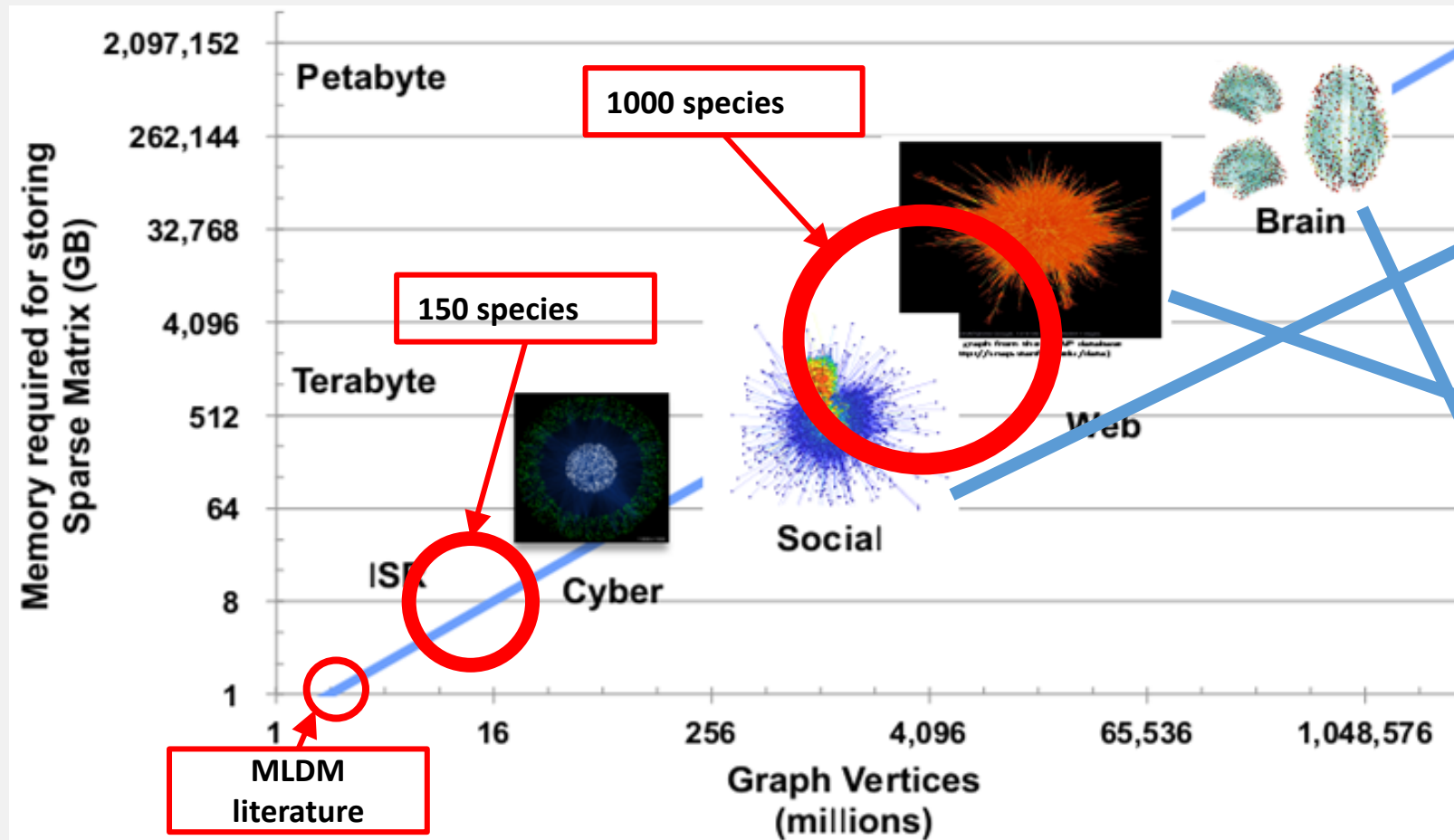


Exponential growth of sequence data (5x each year)

Michael R. Stratton, Peter J. Campbell & P. Andrew Futreal
*Nature* **458, 719-724(9 April 2009)**

# Example: Astronomy

# Example: Graphs

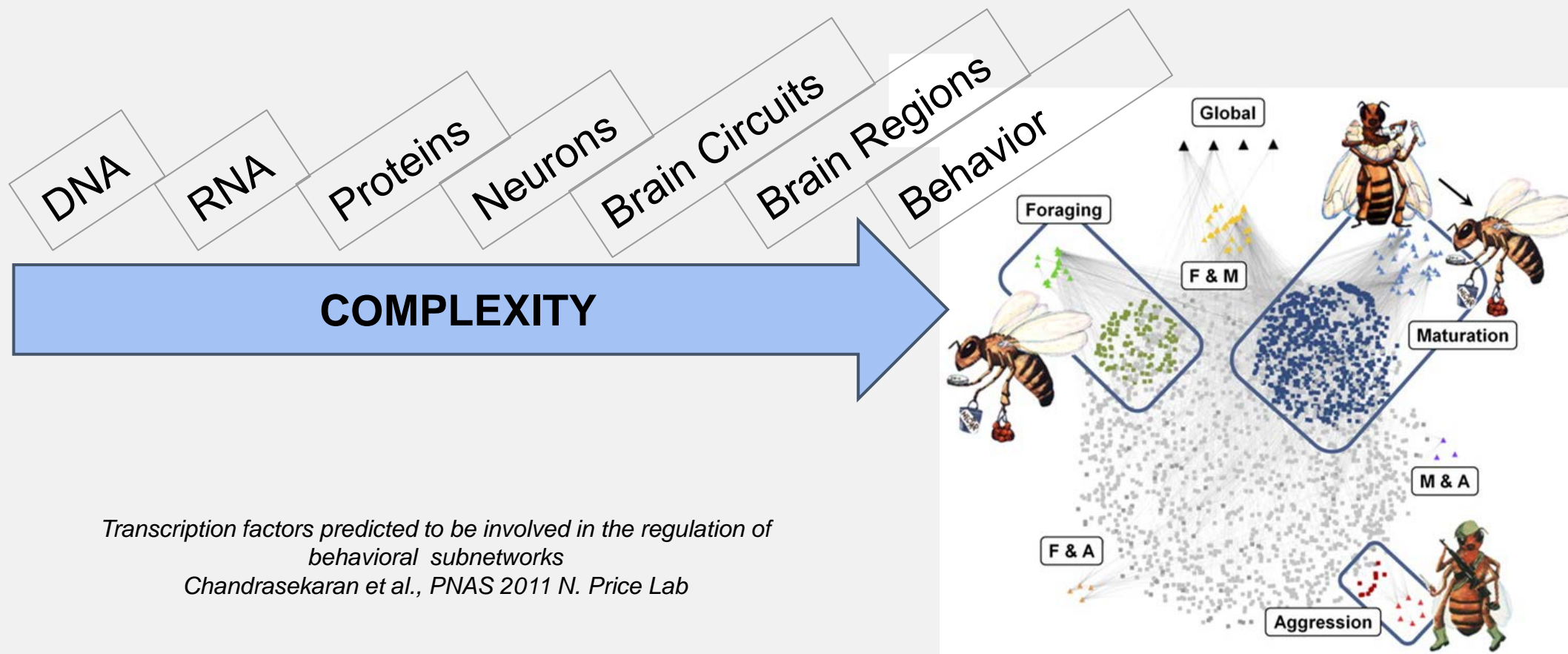# Future: Scale of Real-World Graphs

From genes to brains and social behavior…



*Transcription factors predicted to be involved in the regulation of behavioral subnetworks*
*Chandrasekaran et al., PNAS 2011 N. Price Lab*

# Scale of Real-World Graphs: Today
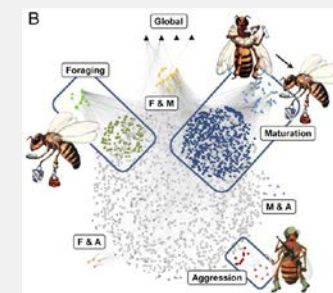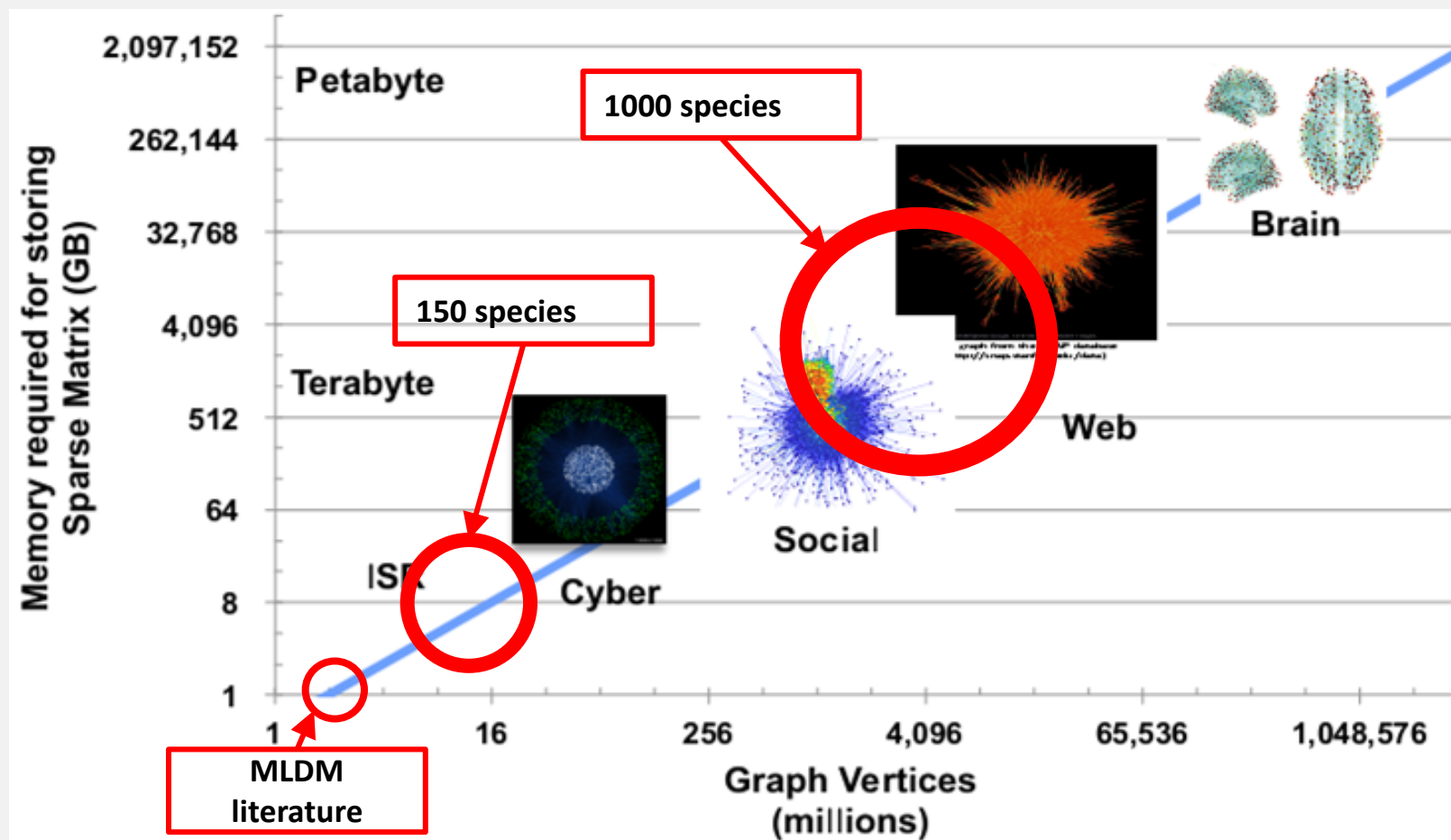


Genomic and brain to social behavior