Artificial Intelligence++

Antal Péter
antal@mit.bme.hu

Computational Biomedicine (Combine) workgroup
Intelligent Systems group
Department of Measurement and Information Systems,
Budapest University of Technology and Economics

What is AI?

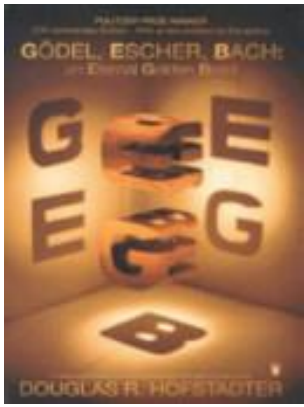
AI approaches can be grouped as follows:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

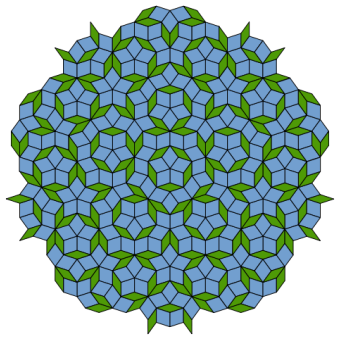
These are individualist approaches, others:

- „Strong" AI: conscious, self-aware AI
- Singularity: self-accelerating AI
- Collective intelligence: multiagent AI approaches
- Automated Science: beyond a singular cognition
- Scientific paradigms in AI: towards data-driven AI

What is AI? Strong AI



- About computation, cognition and art
  - Interview with Douglas R. Hofstadter
    - http://www.americanscientist.org/bookshelf/pub/douglas-r-hofstadter
  - D.R.Hofstadter: Gödel, Escher, Bach, 1979
- About computation and cognition
  - R.Penrose: Emperor's new mind
- About cognition
  - M.Boden:Artificial Intelligence and Natural Man, 1977
  - M.B.: Computer Models of Mind, 1988
  - M.B.: Mind As Machine: a History of Cognitive Science, 2006
- About consciusness
  - **K.R. Popper: The Self and Its Brain, 1977**
  - **D.C. Denett: Consciousness Explained, 1991**
  - http://www.ted.com/talks/dan_dennett_on_our_consciousness?language=en



What is AI? Singularity

- About singularity
  - R. Kurzweil: The singularity is near (when humans transcend biology), 2005
- „Optimal scientist”:
  - J. Schmidhuber: „his main scientific ambition has been to build an optimal scientist through self-improving AI, then retire.” „ build a scientist better than himself (his colleagues claim that should be easy) who will then do the remaining work” http://people.idsia.ch/~juergen/
- Self-improving AI
  - M. Hutter: AIXI
- Neuroprothesis
  - W.Gibson: Neuromancer, 1984
  - Artificial retina, Cochlear implants, Deep brain stimulation
- Simulating neurons
  - IBM’s TrueNorth: neuromorph chip with a 10^6 neuron and 10^8 synapses
  - http://www.research.ibm.com/articles/brain-chip.shtml
- Connecting brains
  - Brainets: Building an organic computing device with multiple interconnected brains
- Synthetic biology
  - International Genetically Engineered Machine (iGEM)
  - http://igem.org/Main_Page

What is AI? Collective intelligence

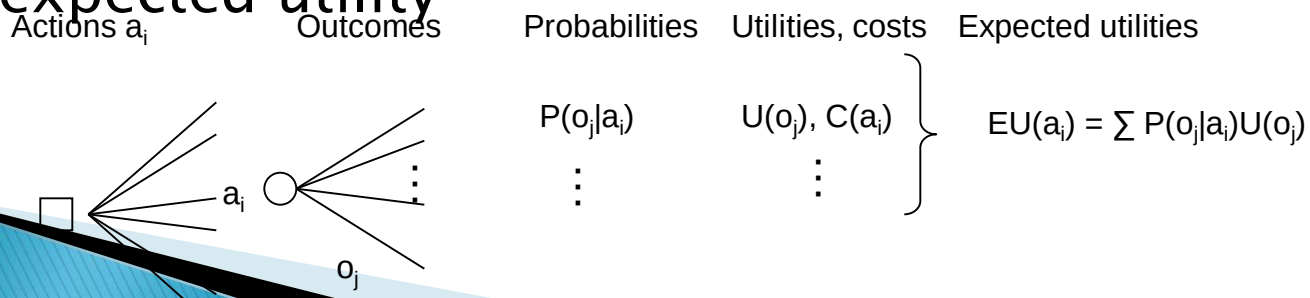
- Game theory
  - Sequential decisions
  - Optimal policy
- Dashboard architecture
- „Democratic” voting schemes
- Protocols for agent communities
- Multiplayer games

What is AI? Automated science

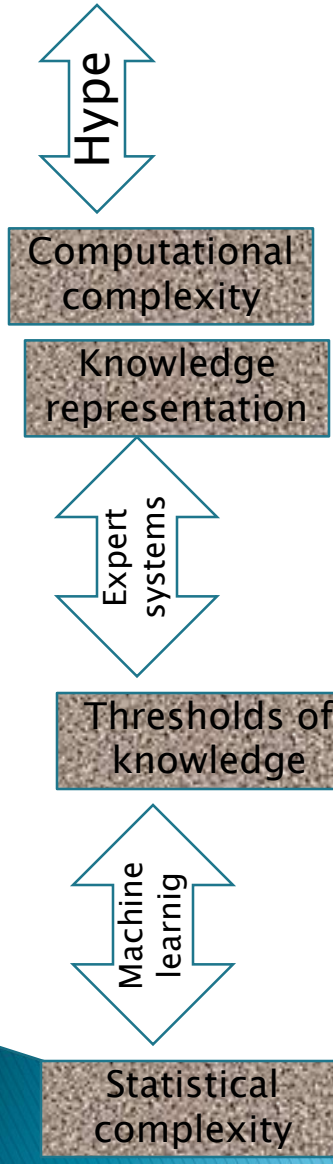
- Classical flow/loop of science
  - Hypothesis/theory➔design➔experiment➔data➔..
  - K.R.Popper: The Logic of Scientific Discovery, 1934
  - K.R.P: Objective Knowledge, 1972
  - T.S.Kuhn: The Structure of Scientific Revolutions, 1962
- Evolution/types of **scientific paradigms.**
  - Experimental
  - Theoretical
  - Computational
  - Data-driven
  - (AI: Brooks: Intelligence without representation)

Decision theory probability theory+utility theory

- Decision situation:
  - Actions a_i
  - Outcomes o_j
  - Probabilities of outcomes $p(o_j \mid a_i)$
  - Utilities/losses of outcomes $U(o_j \mid a_i)$
  - Maximum Expected Utility Principle (MEU) $EU(a_i) = \sum_j U(o_j \mid a_i) p(o_j \mid a_i)$
  - Best action is the one with maximum expected utility $a^* = \arg\max_i EU(a_i)$



Milestones and phases in AI

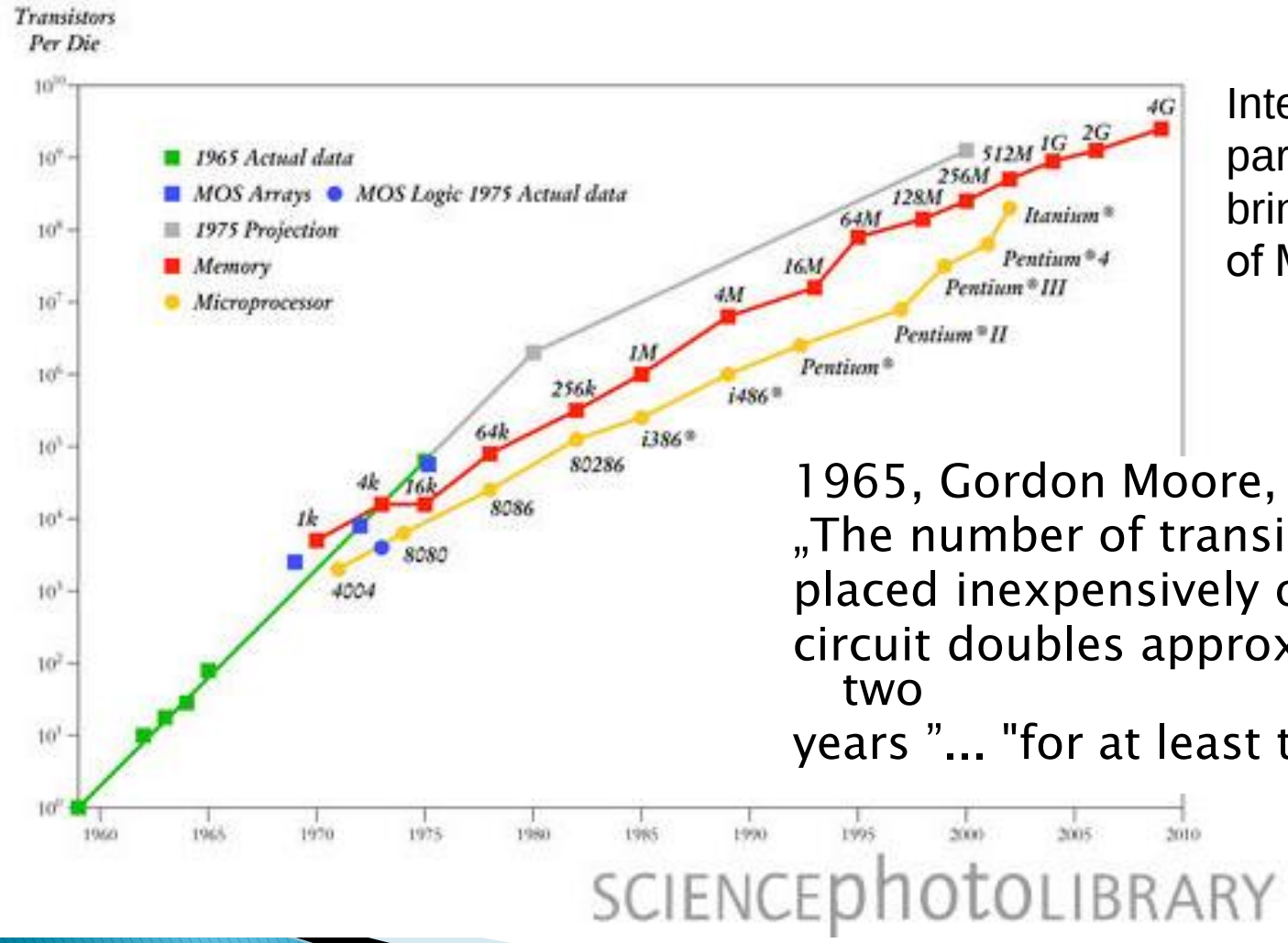


- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1956 Dartmouth meeting: the term "Artificial Intelligence"
- 1950s Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1965 Robinson's complete algorithm for logical reasoning
- 1966—73 AI discovers computational complexity Neural network research almost disappears
- 1969—79 Early development of knowledge-based systems

- 1986-- Neural networks return to popularity
- 1988-- Probabilistic expert systems
- 1995-- Emergence of machine learning

Today: heterogeneous AI, data-intensive science, data and knowledge fusion

Moore's Law (in computation)



Integration and parallelization wont bring us further. End of Moor's law?

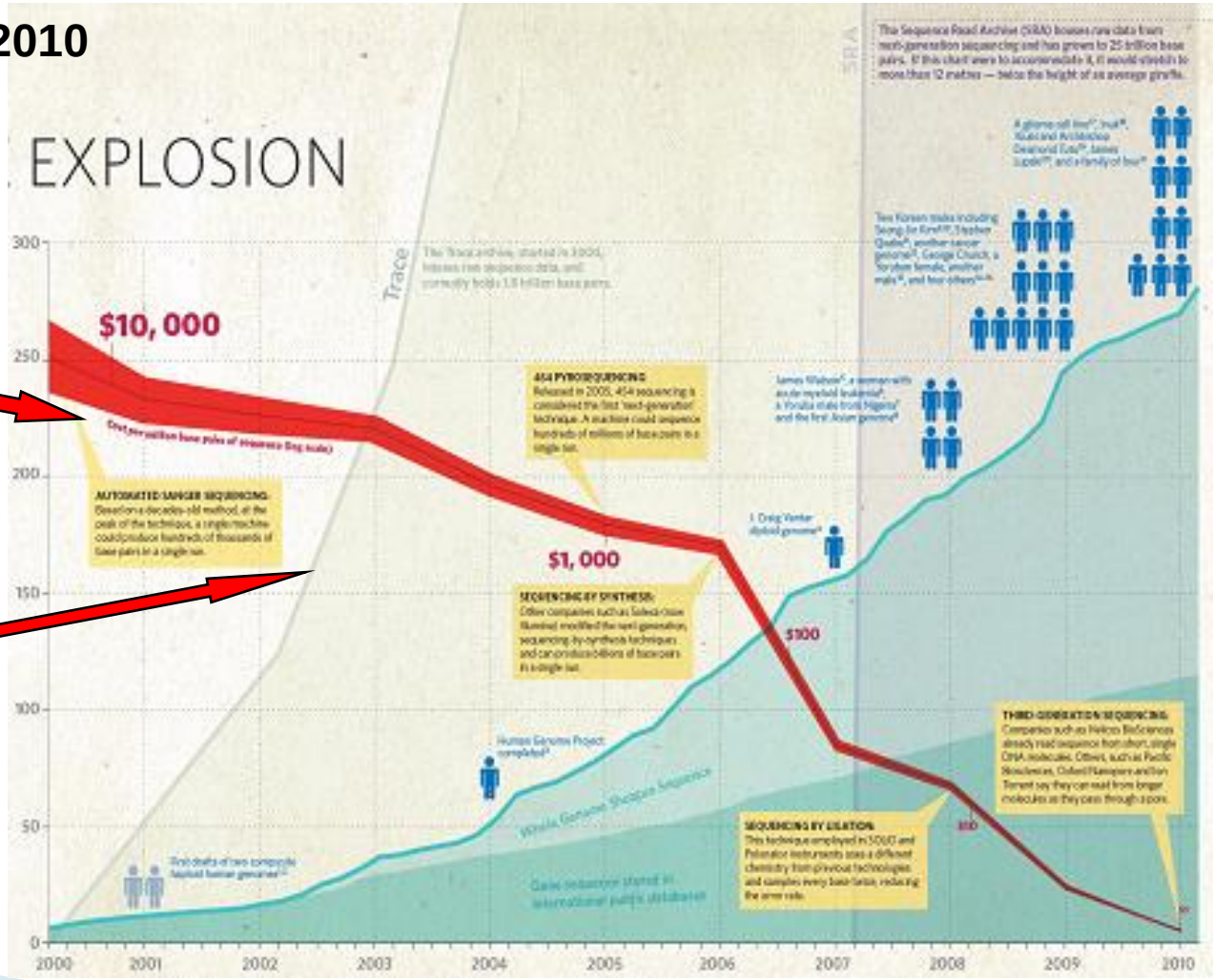
1965, Gordon Moore, founder of Intel: „The number of transistors that can be placed inexpensively on an integrated circuit doubles approximately every two years "... "for at least ten years"

Carlson's Law for Biological Data

NATURE, Vol 464, April 2010

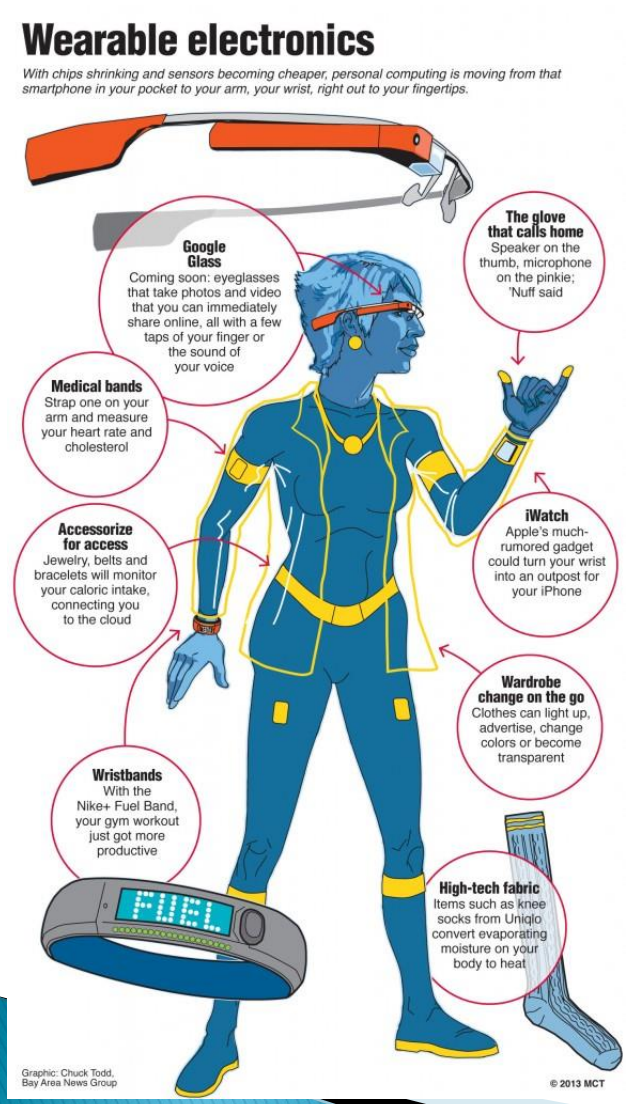
Sequencing costs per mill. base

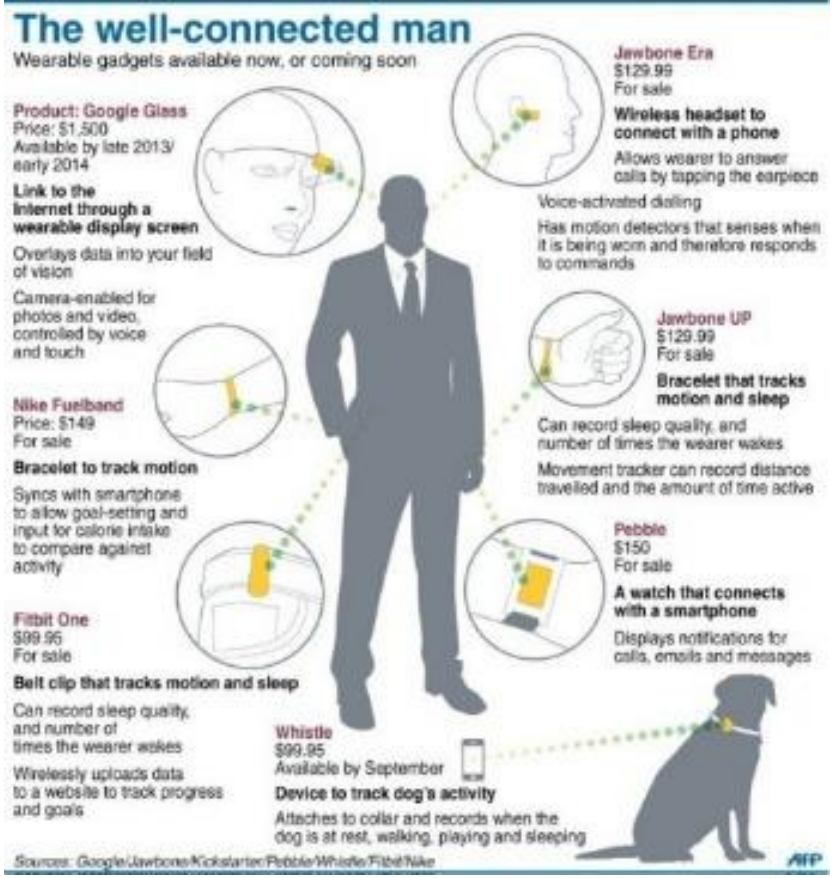
Publicly available genetic data

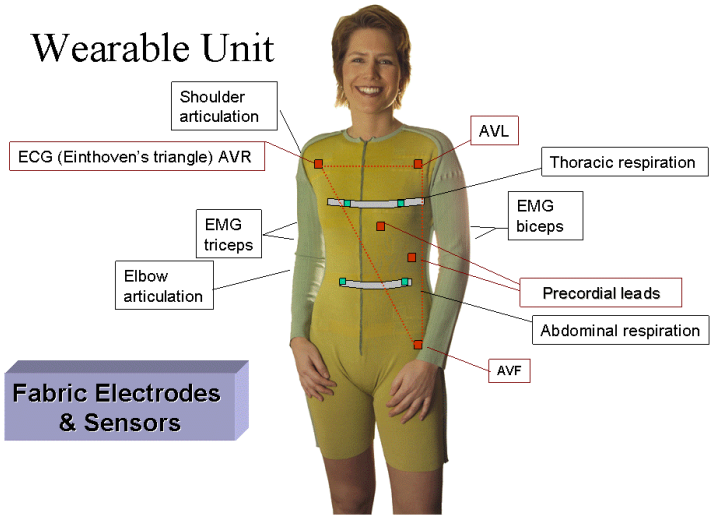


- x10 every 2-3 years
- Data volumes and complexity that IT has never faced before…

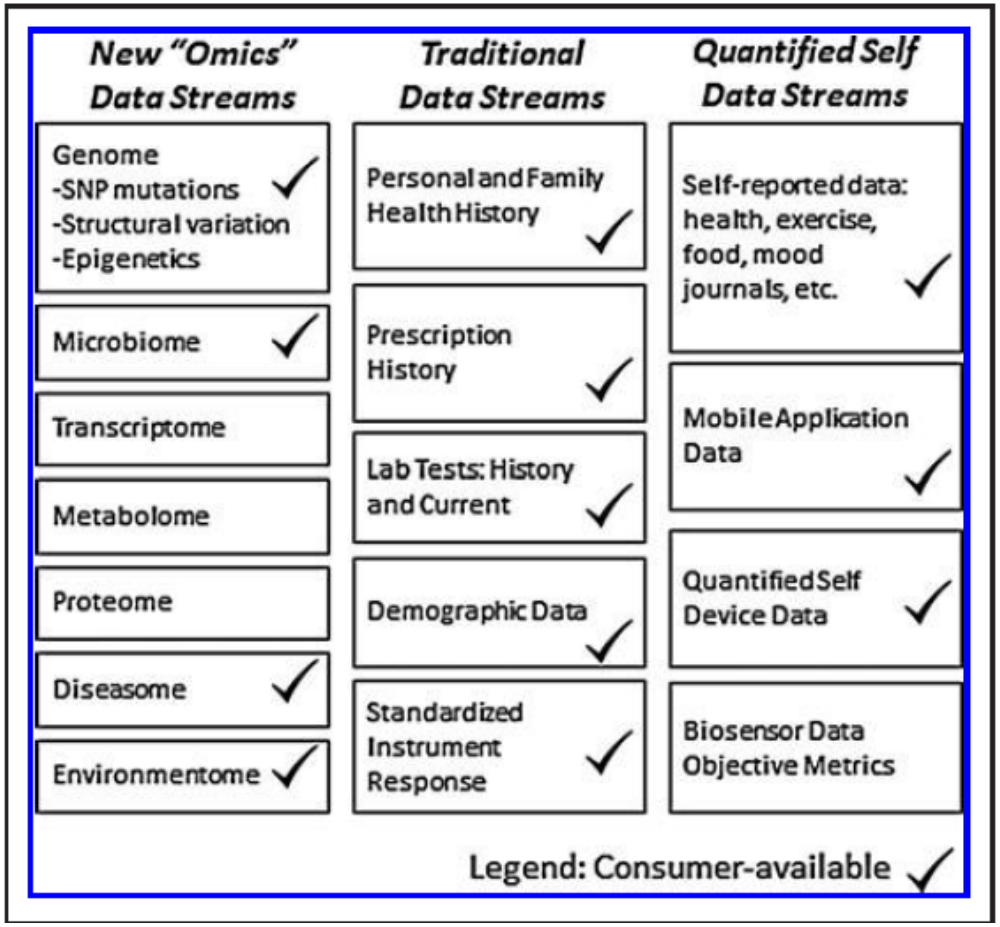
Quantified self





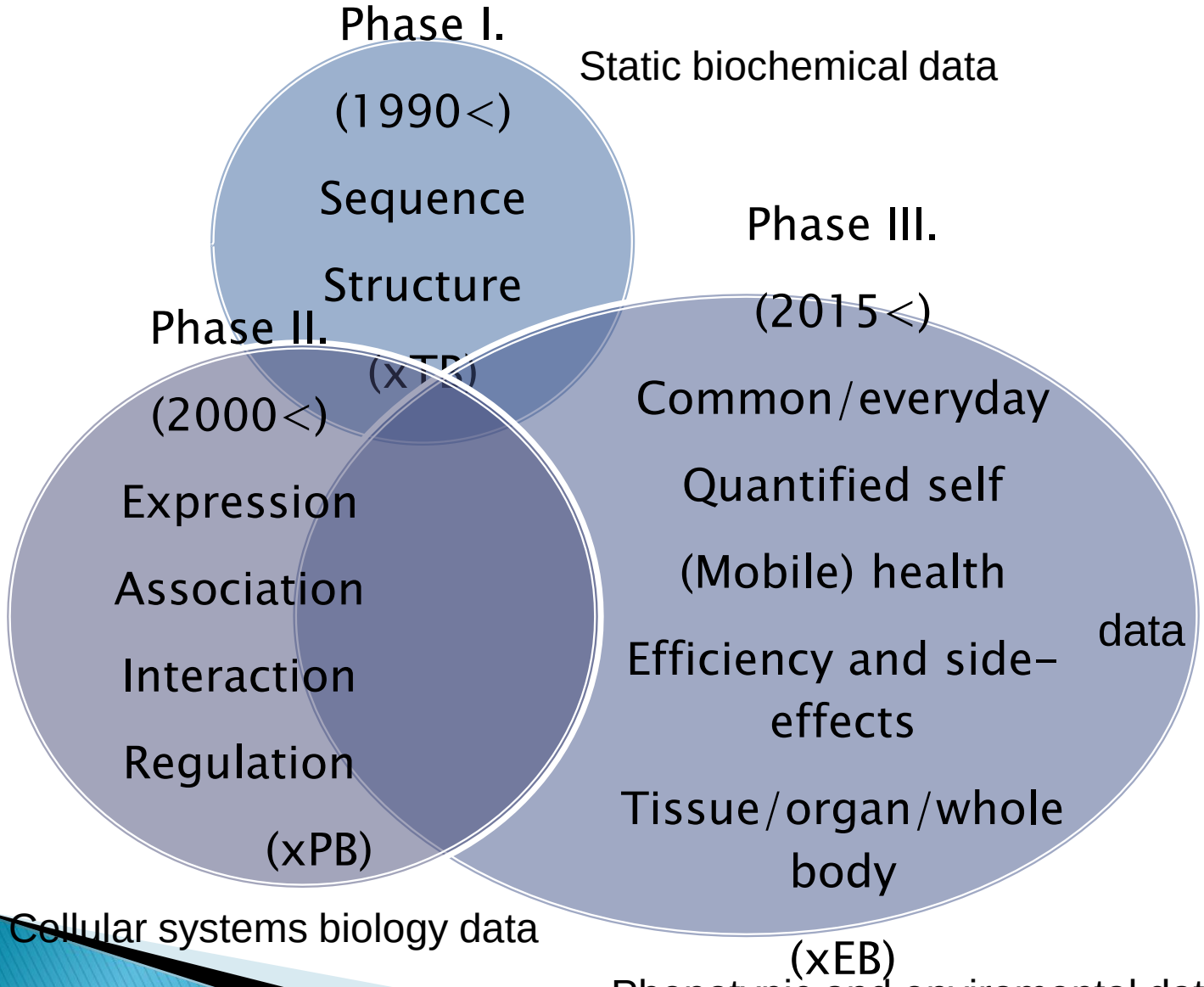


Big health data streams



M.Swan: THE QUANTIFIED SELF: Fundamental Disruption in Big Data Science and Biological Discovery, Big data, Vol 1., No. 2., 2013

(Big) Data floods in biomedicine



The data-intensive science, data science, data engineering

- Data analysis and knowledge fusion is more important than simulation of „simple” laws.
- 20th century: Physics vs. 21st century: Biology.
  - Tony Hey, Stewart Tansley, and Kristin Tolle: **The fourth paradigm (Data-Intensive Scientific Discovery)**, http://research.microsoft.com/en-us/collaboration/fourthparadigm/, 2009
  - Gordon Bell, Tony Hey, Alex Szalay: **Beyond the Data Deluge**, Science, 323, pp 1297-1298, 2009

Besides data: expert knowledge

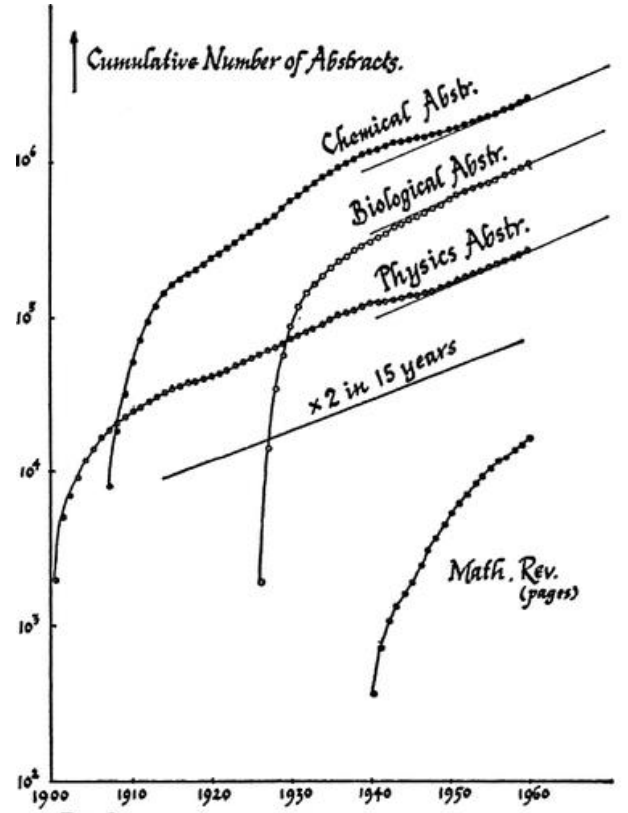
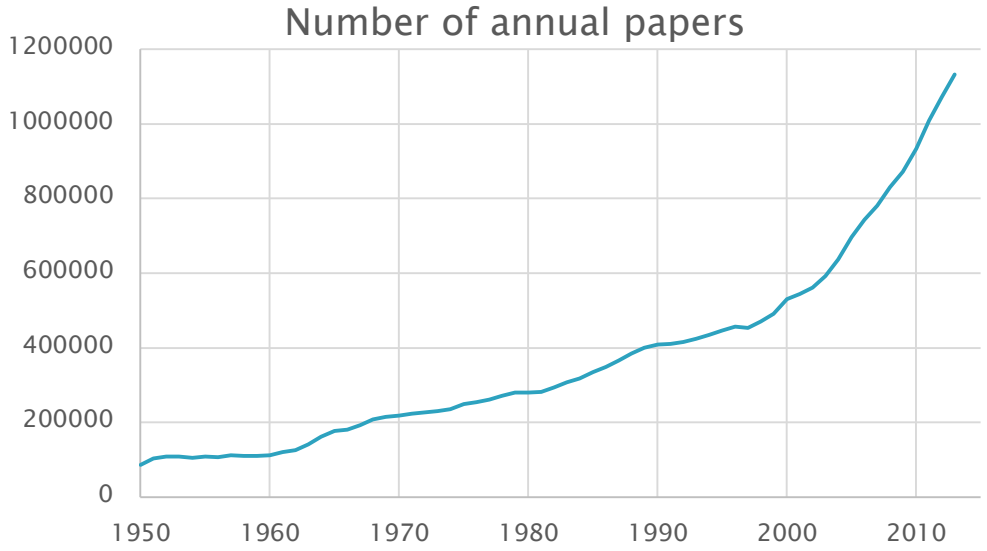
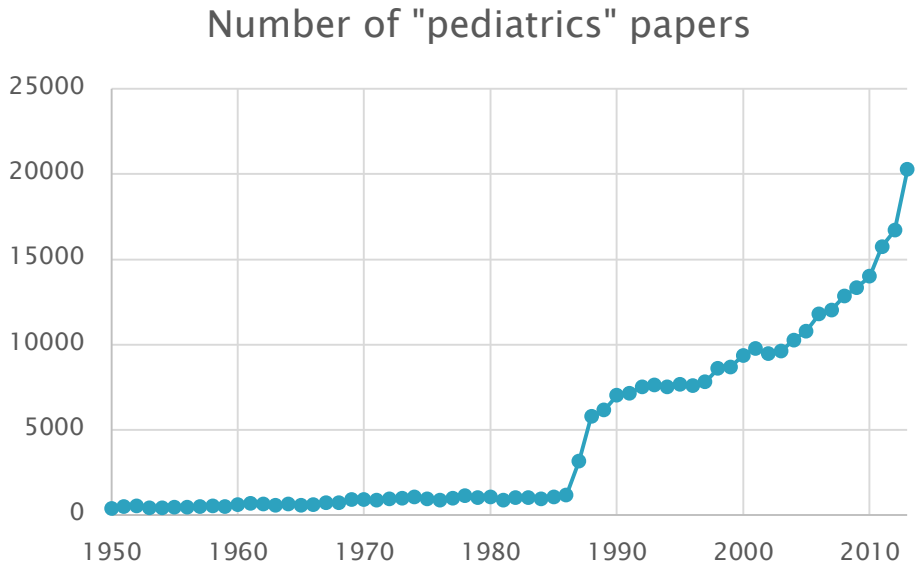


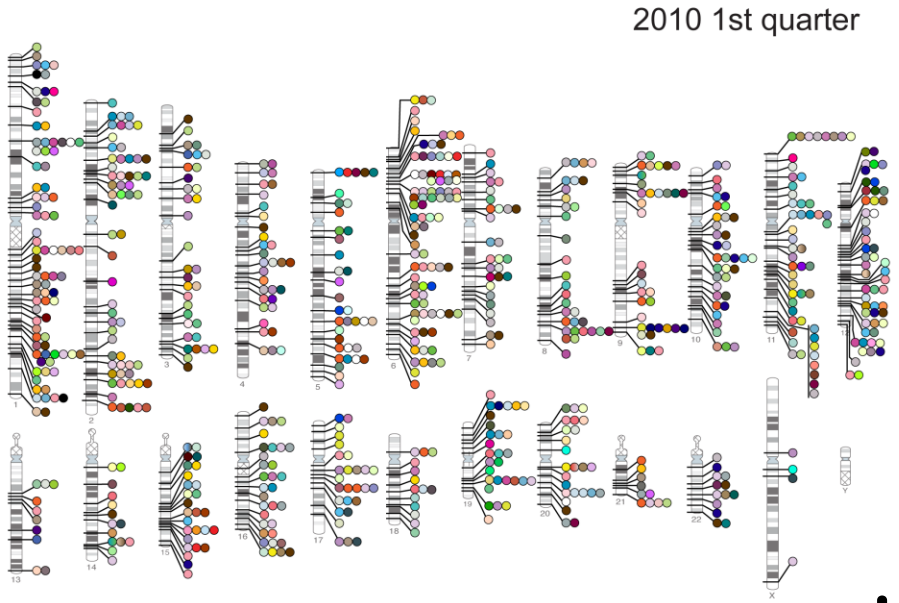
Fig. 2. CUMULATIVE NUMBER OF ABSTRACTS IN VARIOUS SCIENTIFIC FIELDS, FROM THE BEGINNING OF THE ABSTRACT SERVICE TO GIVEN DATE

It will be noted that after an initial period of rapid expansion to a stable growth rate, the number of abstracts increases exponentially, doubling in approximately 15 years.

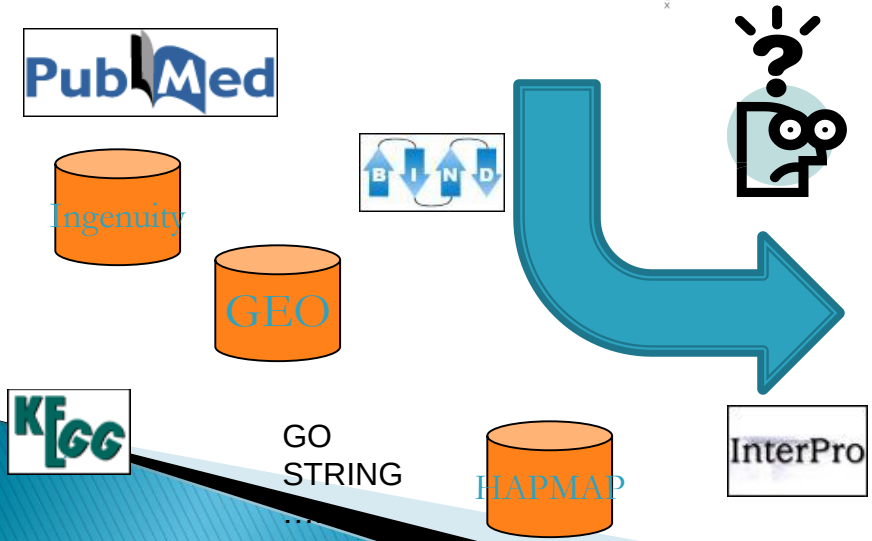


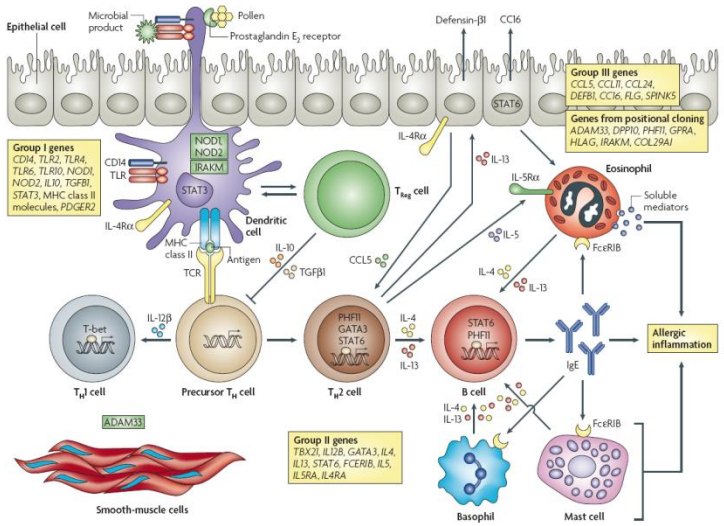


The hypothesis-free research



- Hypothesis-free measurement
- Hypothesis-free data analysis
- Interpretational/translational bottleneck
- Dissemination and reuse of significantly weak results?

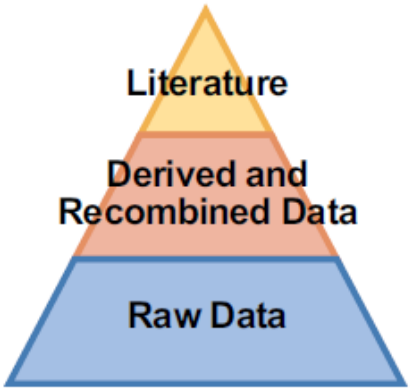


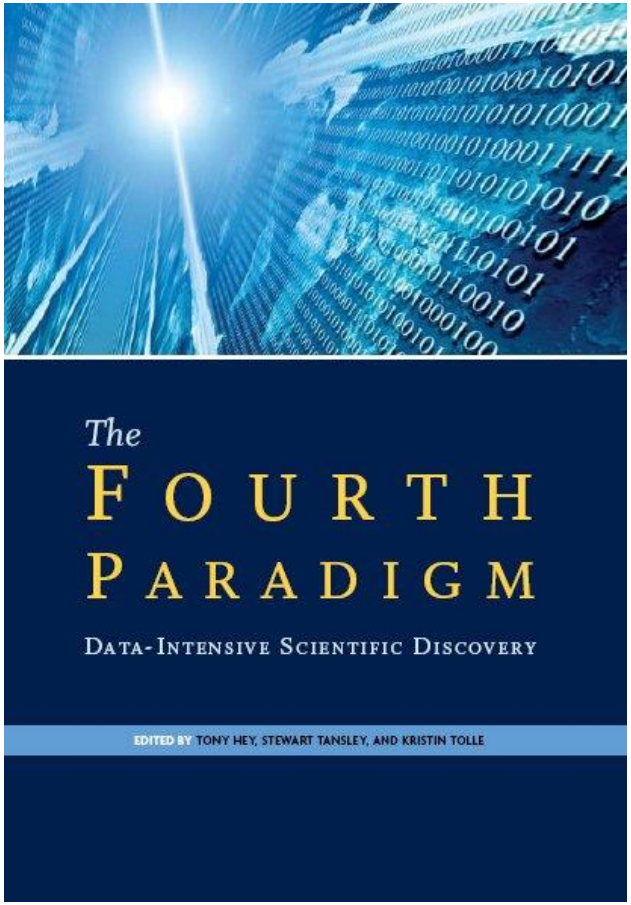


E-science, data-intensive science, the fourth paradigm

All Scientific Data Online

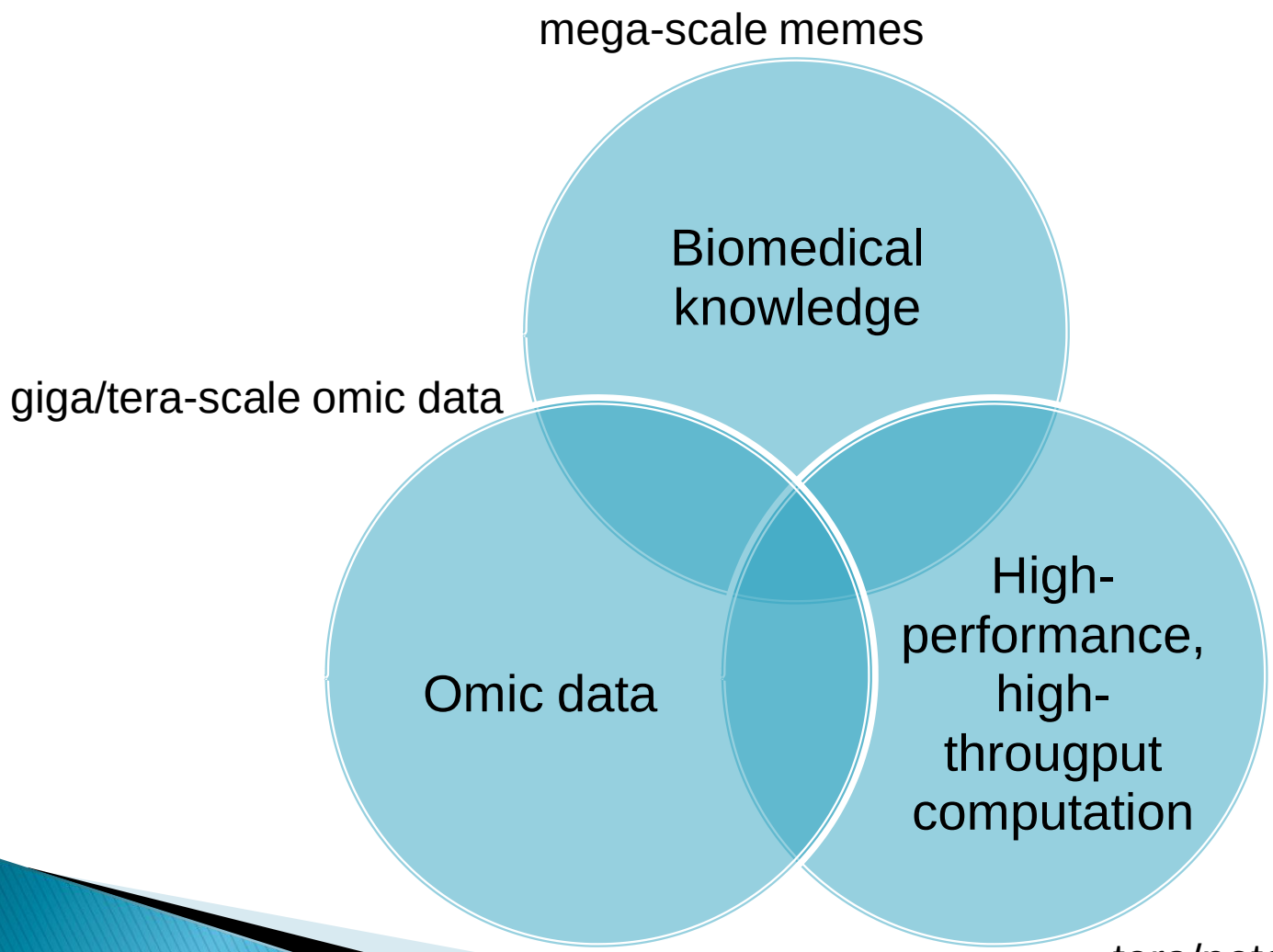
- Many disciplines overlap and use data from other sciences
- Internet can unify all literature and data
- Go from literature to computation to data back to literature
- Information at your fingertips for everyone-everywhere
- Increase Scientific Information Velocity
- Huge increase in Science Productivity



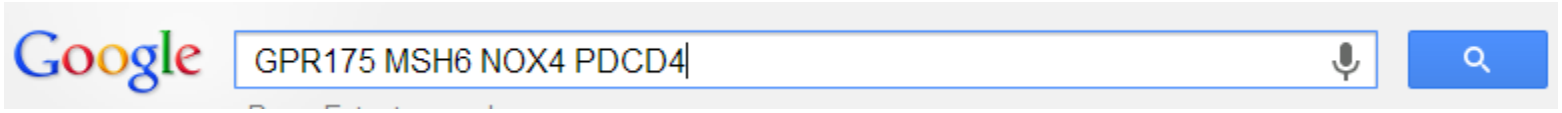


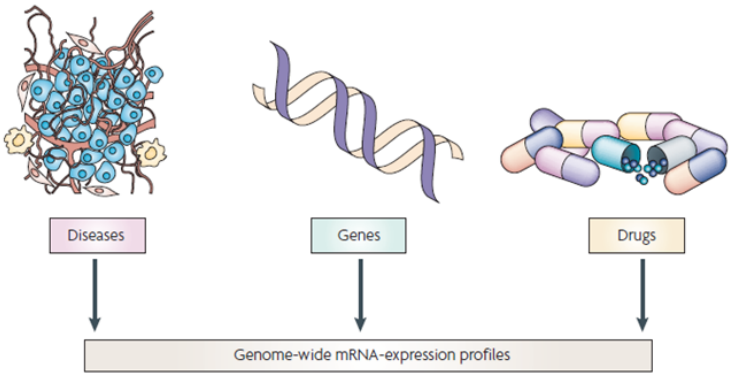
http://research.microsoft.com/en-us/collaboration/fourthparadigm/

Vision: the fusion dream



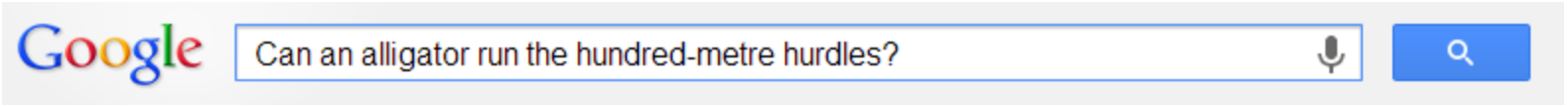
“Real” artificial intelligence?





“Molecular google”

J.Lamb: The Connectivity Map, Nature, 2007



“The Science Behind an Answer”

Turing test and prediction, Computing Machinery and Intelligence, 1950

Automated discovery systems

- Langley, P. (1978). Bacon: A general discovery system. Proceedings of the Second Biennial Conference of the Canadian Society for Computational Studies of Intelligence (pp. 173-180). Toronto, Ontario.
- …
- Chrisman, L., Langley, P., & Bay, S. (2003). Incorporating biological knowledge into evaluation of causal regulatory hypotheses. Proceedings of the Pacific Symposium on Biocomputing (pp. 128-139). Lihue, Hawaii.
- (Gene prioritization…)
- R.D.King et al.: The Automation of Science, Science, 2009

Summary

- Besides individualist approaches to AI others:
  - Self-aware AI
  - Self-accelerating AI
  - Cooperative/collective AI
  - Data-driven AI
  - Automated Science

- Suggested reading:
  - Boden: Computer models of creativity, 2009
  - King: The Automation of Science, Science, 2009