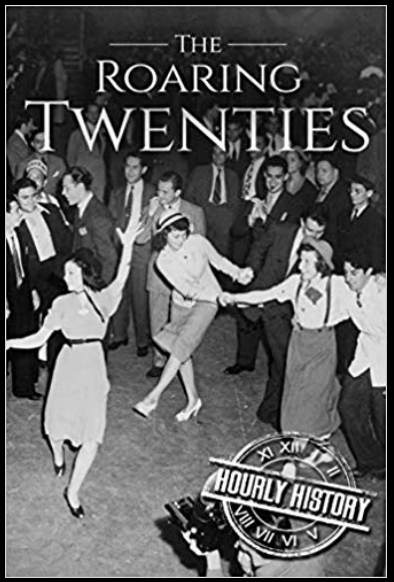
ICT to support the transformation of Science in the Roaring Twenties

Cees de Laat

Systems and Networking Laboratory
Complex Cyber Infrastrure group
University of Amsterdam

ICT to support the transformation of Science in the Roaring Twenties



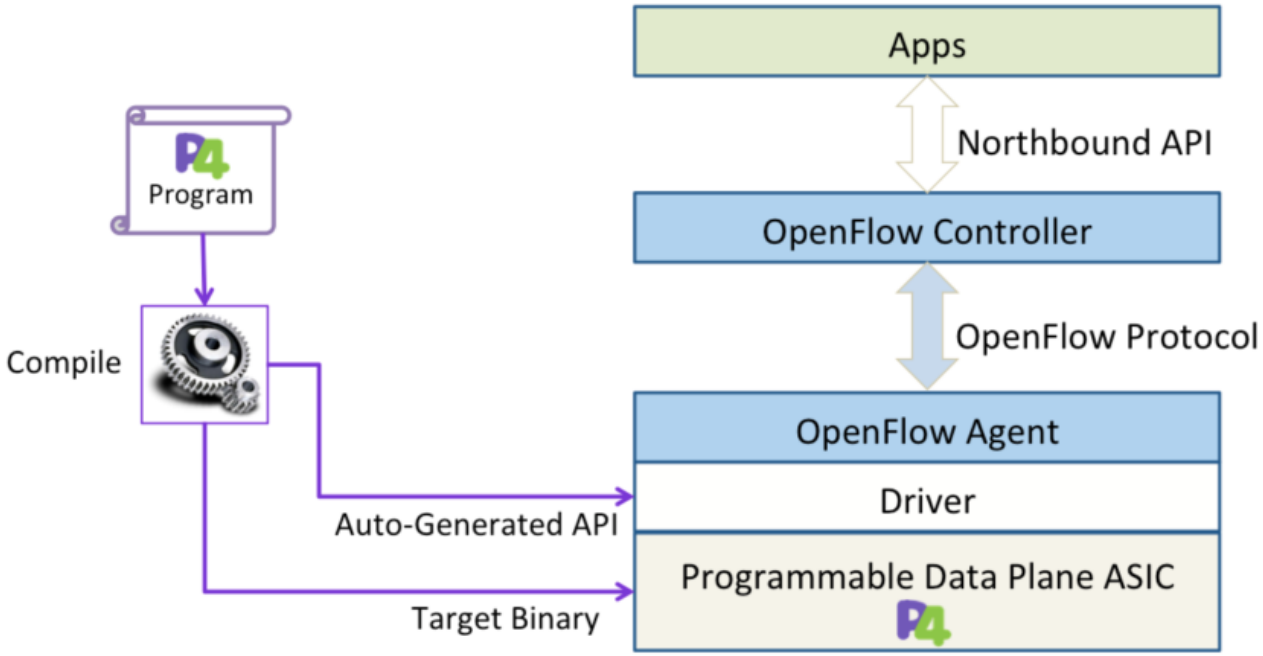
From Wikipedia: The Roaring Twenties refers to the decade of the 1920s in Western society and Western culture. It was a period of economic prosperity with a distinctive cultural edge in the United States and Western Europe, particularly in major cities such as Berlin, Chicago, London, Los Angeles, New York City, Paris, and Sydney. In France, the decade was known as the "années folles" ('crazy years'), emphasizing the era's social, artistic and cultural dynamism. Jazz blossomed, the flapper redefined the modern look for British and American women, and Art Deco peaked….

This period saw the large-scale development and use of automobiles, telephones, movies, radio, and electrical appliances being installed in the lives of thousands of Westerners. Aviation soon became a business. Nations saw rapid industrial and economic growth, accelerated consumer demand, and introduced significantly new changes in lifestyle and culture. The media focused on celebrities, especially sports heroes and movie stars, as cities rooted for their home teams and filled the new palatial cinemas and gigantic sports stadiums. In most major democratic states, women won the right to vote. The right to vote made a huge impact on society.

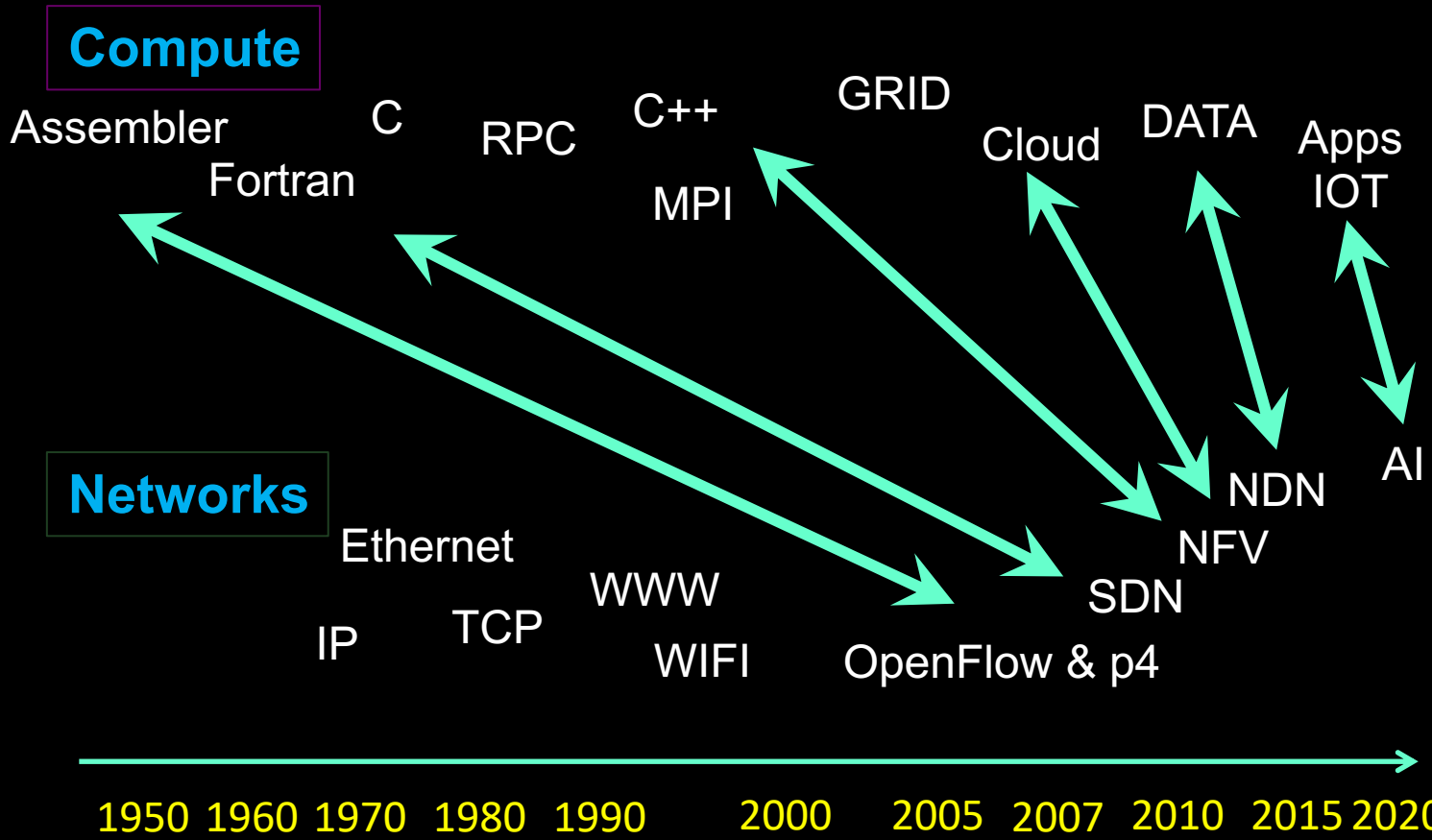
Transformations

- Internet
- Computing
- Data
- Science

P4 & OpenFlow



TimeLine
Network programmability and virtualisation

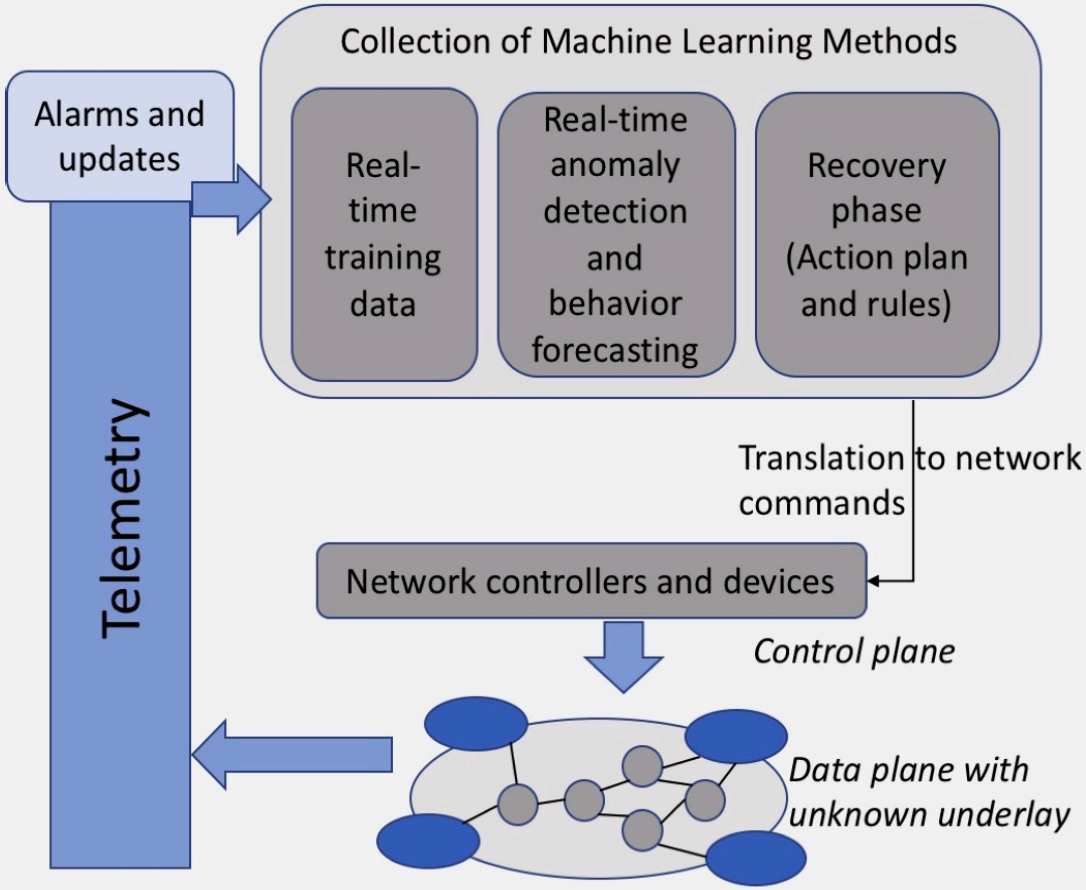


Example 1: Optimizing Network Traffic with Machine Learning

Exascale and increasingly complex science applications are exponentially raising demands from underlying DOE networks, such as traffic management, operational latency, and bandwidth constraints. Networks are the backbone to complex science workflows, ensuring data are delivered securely and on time for important computations to happen. To optimize these distributed workflows, networks are required to understand end-to-end performance ahead of time in advance and be faster, efficient, and more proactive, anticipating bottlenecks before they happen. However, to intelligently route multiple network paths intelligently, various tasks, such as pre-computation and prediction, must be done in near real time. ML provides a collection of algorithms that can add autonomy and assist in decision making to sup-

AI creeping in the control plane

Self driving cars, so why not self driving CI?



The Trend

- Internet used to be end user to end user or service
  - Meshed network
  - Internet exchanges
  - Net Neutrality
- It is becoming end user to data center
  - Internal data center “meet me” rooms
  - Data centers interconnect based on business
  - Less and less data via Internet exchanges
  - Neutrality may get violated by filtering, policing
- And we are back where we started, a bundled phone system.

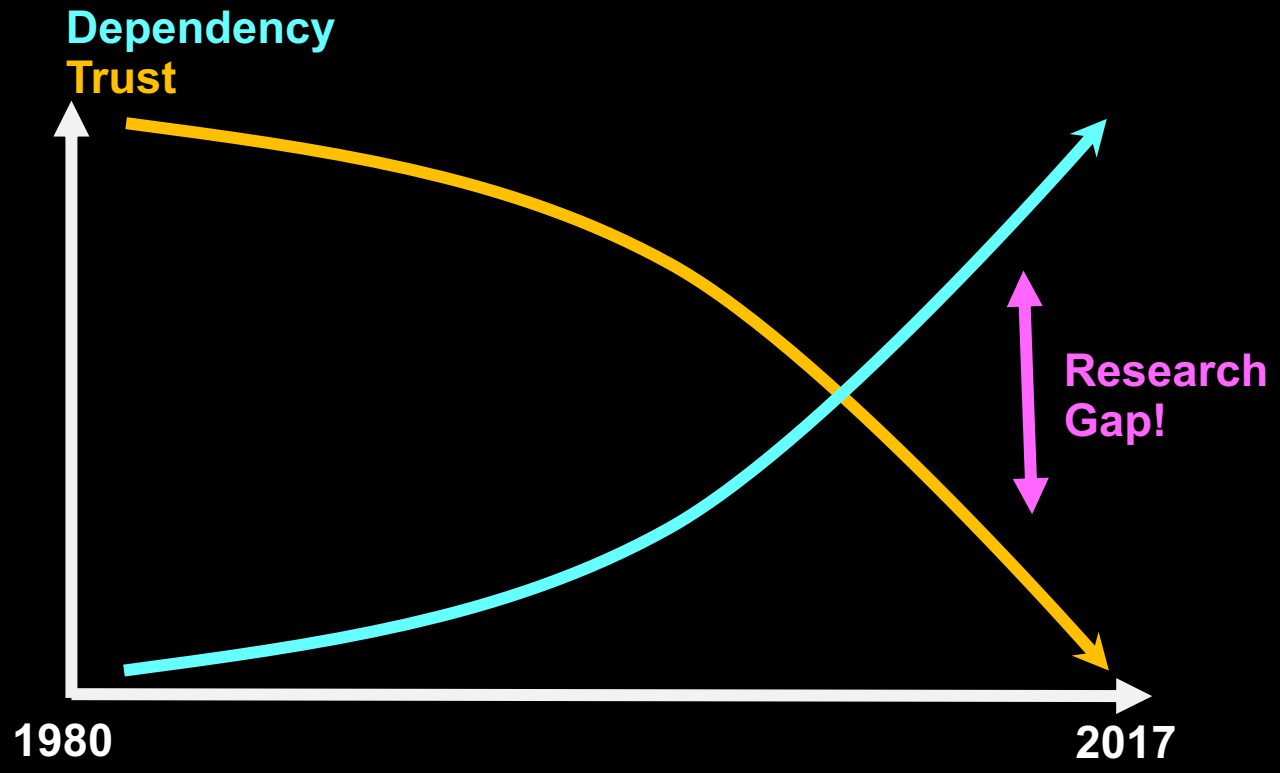
Internet moves from IXP's into datacenters

Rapidly loosing internet transparency

A Responsible Internet to Increase Trust in the Digital World

Cristian Hesselman, Paola Grosso, Ralph Holz, Fernando Kuipers, Janet Hui Xue, Mattijs Jonker, Joeri de Ruiter, Anna Sperotto, Roland van Rijswijk-Deij, Giovane C. M. Moura, Aiko Pras, Cees de Laat, "A Responsible Internet to Increase Trust in the Digital World", Journal of Network and Systems Management, (JNSM), special issue on "Future of Network and Service Operations and Management: Trends, Developments, and Directions", 28, 882–922 (2020). https://doi.org/10.1007/s10922-020-09564-7.

Fading Trust in Internet

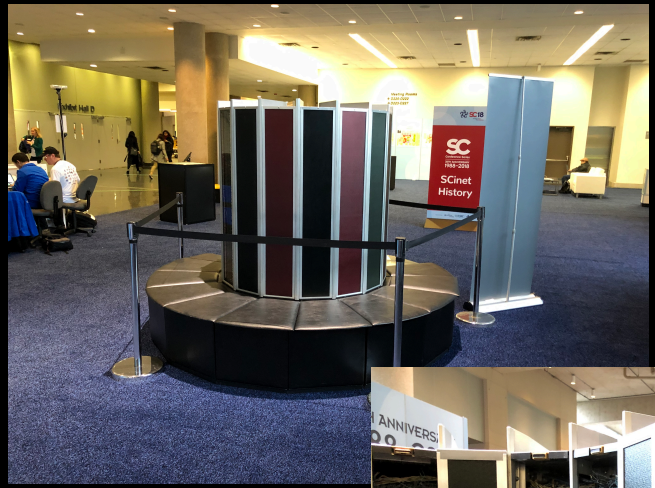


Some progress

2018

540 MHz
~ 1 GFlops
1000 MByte memory
16000 MByte ssd
0,0012 kWh – 18 h

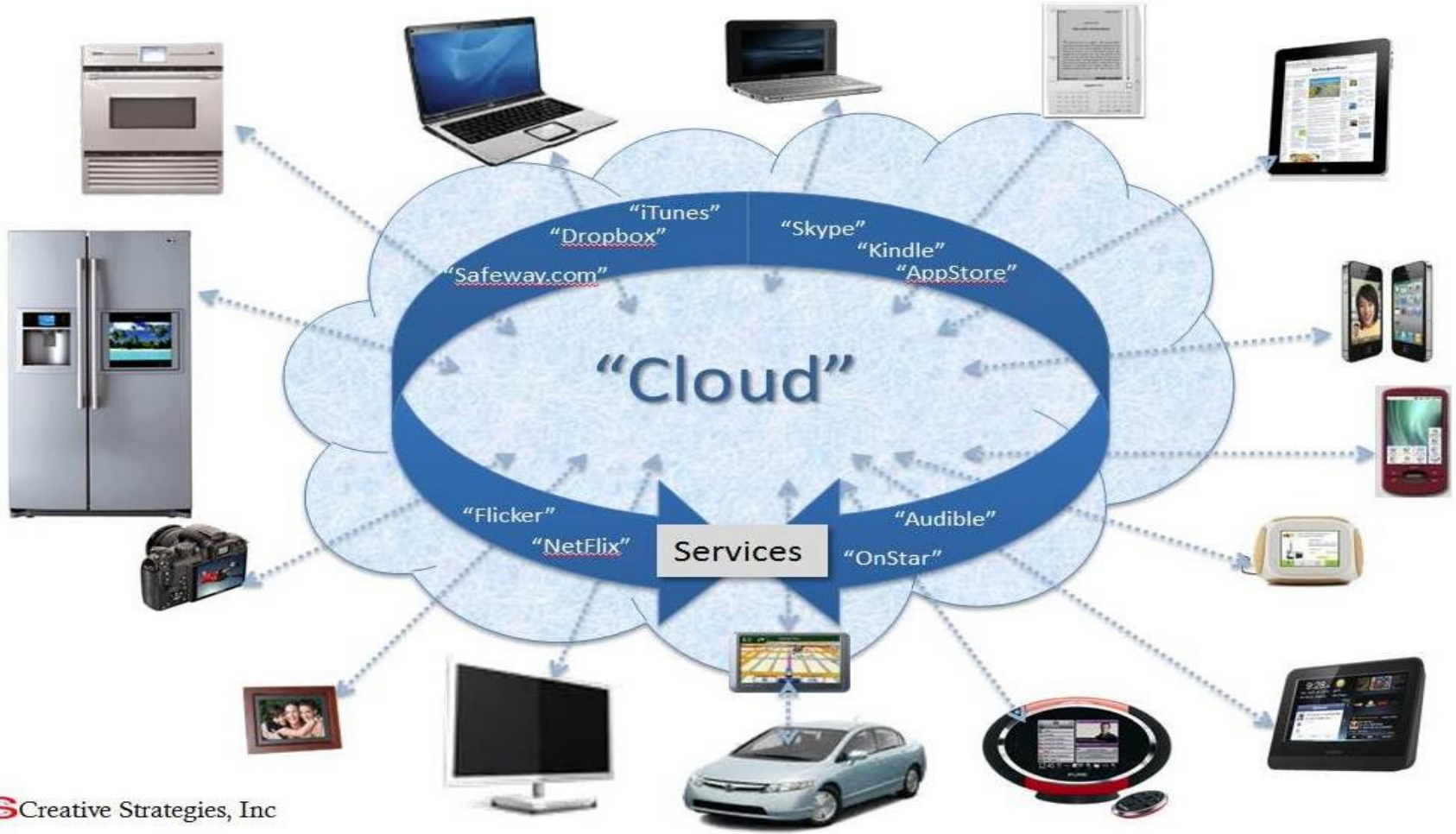
= ~7 ✖

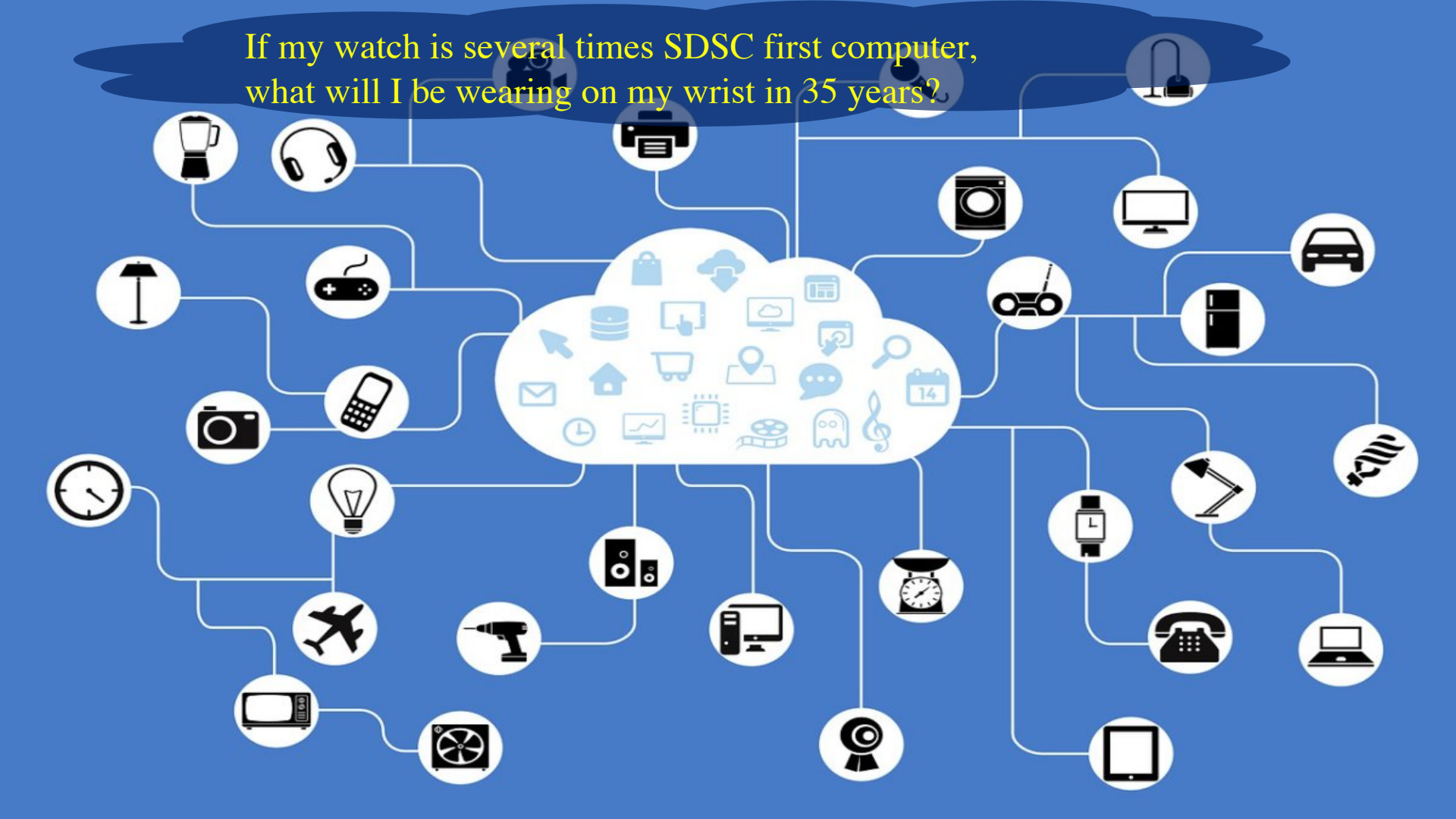


1976

80 MHz
160 MFlops
8 MByte memory
300 MByte disks
120 kW

Internet of Things



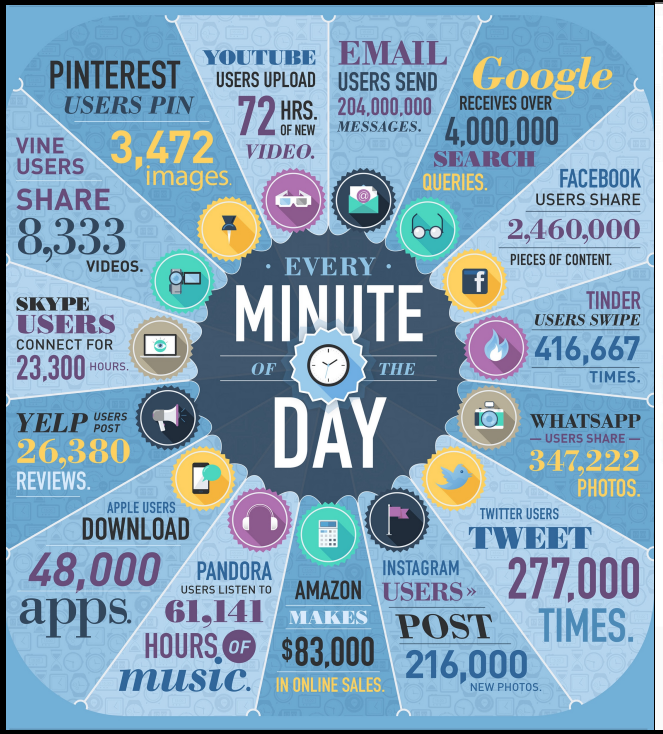

If my watch is several times SDSC first computer,
what will I be wearing on my wrist in 35 years?

Change in computing

- Early days a few big Supercomputers
  - Mostly science domain
- Via grid to commercial cloud
  - AWS, Azure, Google Cloud, IBM, Salesforce
  - The big five: Apple, Alphabet, Microsoft, Facebook and Amazon
  - Computing has transformed into an utility
- Data => Information is the key

Now, how do we get and use data?

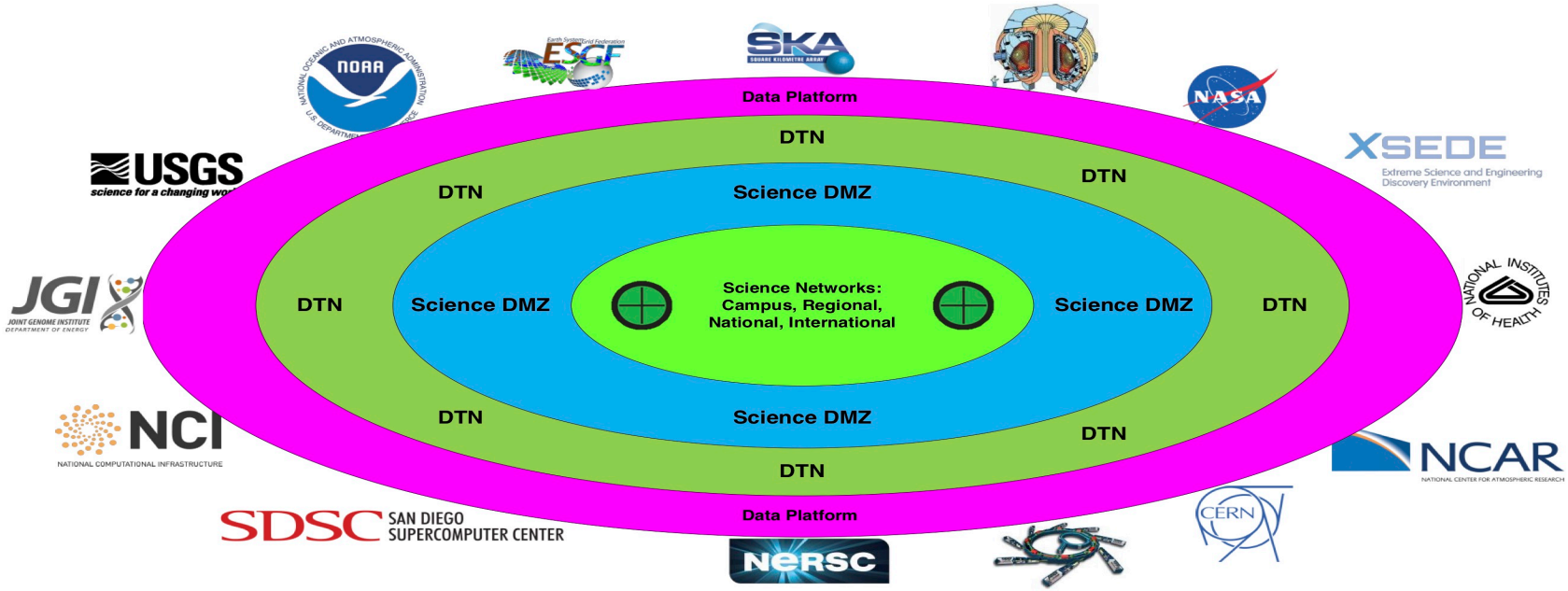
2014



- Move towards streaming
  - Netflix
  - youtube
- Same in science world
  - SKA / LOFAR
  - Light Source
  - Environmental (Marine, Meteorology)
- Data is not always huge
  - Often it is complexity
  - Some example:
    - biodiversity

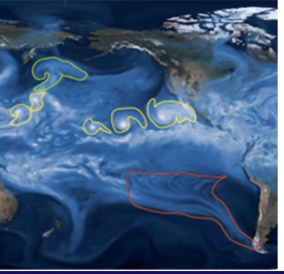
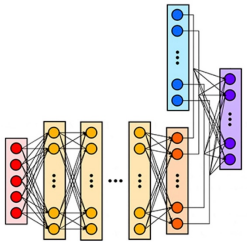
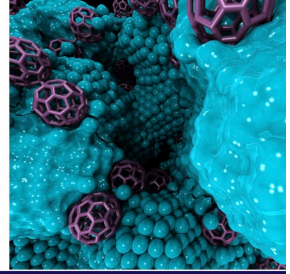
https://www.domo.com/blog/data-never-sleeps-4-0/

Data Ecosystem – Concentric View



ESnet

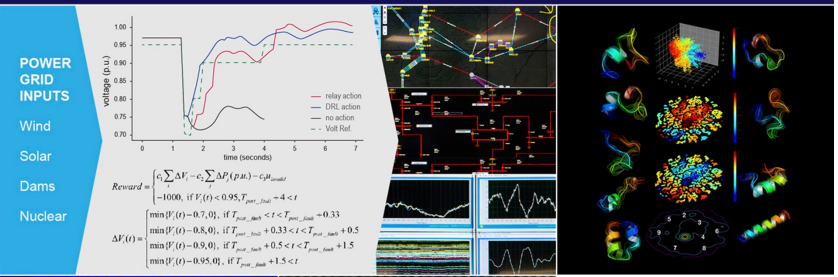
Courtesy Eli Dart, ESnet

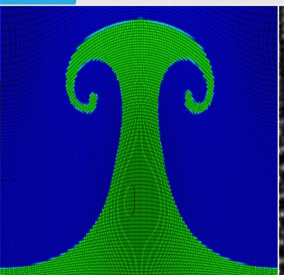
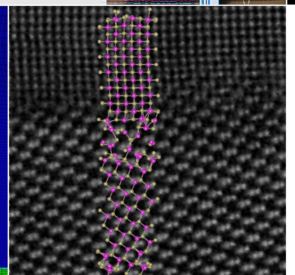




BASIC RESEARCH NEEDS FOR

Scientific Machine Learning

Core Technologies for Artificial Intelligence



Prepared for U.S. Department of Energy Advanced Scientific Computing Research

U.S. DEPARTMENT OF ENERGY

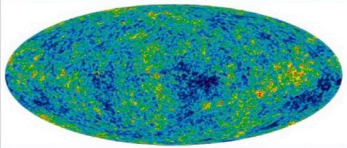
Scientific Machine Learning & Artificial Intelligence

Scientific progress will be driven by
- Massive data: sensors, simulations, networks
- Predictive models and adaptive algorithms
- Heterogeneous high-performance computing

Trend: Human-AI collaborations will transform the way science is done.

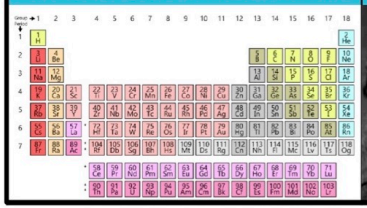
EXEMPLARS OF SCIENTIFIC ACHIEVEMENT

Cosmic Microwave Background

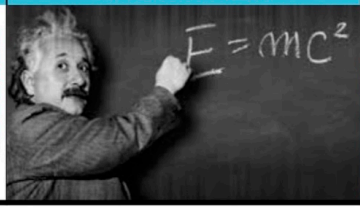


DNA Structure

Periodic Table of the Elements



Special Relativity



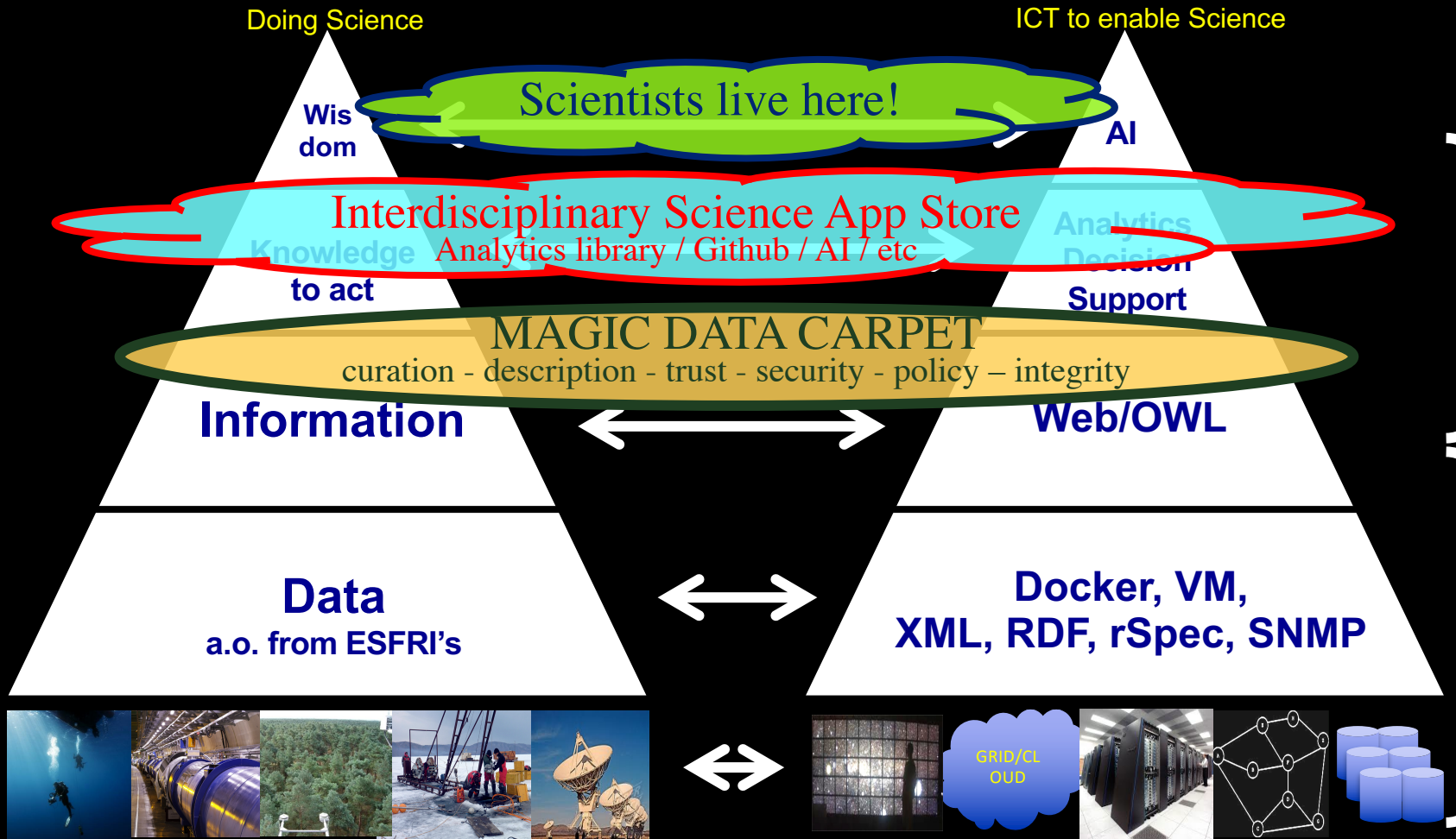
Human-AI insights enabled via scientific method, experimentation, & AI reinforcement learning.

U.S. DEPARTMENT OF ENERGY | Office of Science

DOE Applied Mathematics Research Program
Scientific Machine Learning Workshop (January 2018)

Workshop report:
https://www.osti.gov/biblio/1478744

The Big Data Challenge



Different ways of using and sharing data

Internet of Food Things Network Plus

- Individual self-resourcing
- How most organisations do it

- Market, sharing and exhange
- Social networks

	‘Gated community’	Open market
Data storage 	Platform limited	Free choice.
Data services 	Forced shopping. No services from others.	Purchase from any service provider.
Data transactions 	Only within the platform. No interoperabiliteit with other platforms.	Peer-to-peer transactions and platform independent.
Data control 	No exclusieve control on sovereign data.	Full control on sovereign data.

Harvard Business Review

Harvard Business Review

Subscribe | Sign In | Register

ECONOMY

Managing Our Hub Economy

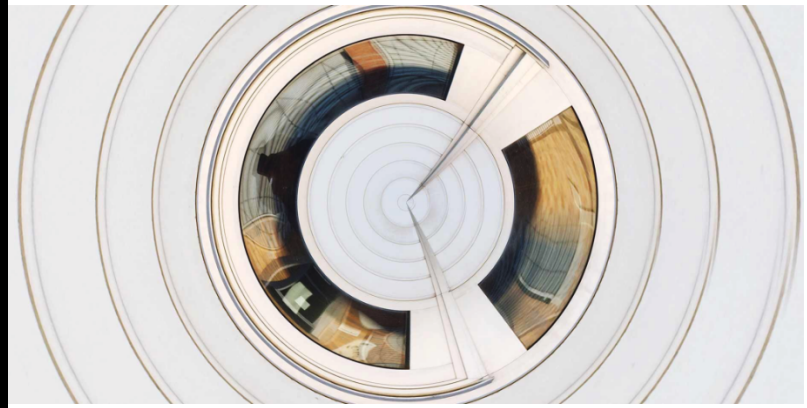
by Marco Iansiti and Karim R. Lakhani

FROM THE SEPTEMBER–OCTOBER 2017 ISSUE

WHAT TO READ NEXT

The IT Transformation Health Care Needs

SUMMARY SAVE SHARE 3 COMMENT TEXT SIZE PRINT $8.95 BUY COPIES



THOMAS M. SCHEER/EYEEM/GETTY IMAGES

I. The Problem

The global economy is coalescing around a few digital superpowers. We see unmistakable evidence that a winner-take-all world is emerging in which a small number of "hub firms"—including Alibaba, Alphabet/Google, Amazon, Apple, Baidu, Facebook, Microsoft, and Tencent—occupy central positions. While creating real value for users, these companies are also capturing a disproportionate and expanding share of the value, and that's shaping our collective economic future. The very same technologies that promised to democratize business are now threatening to make it more monopolistic.

Data value creation monopolies

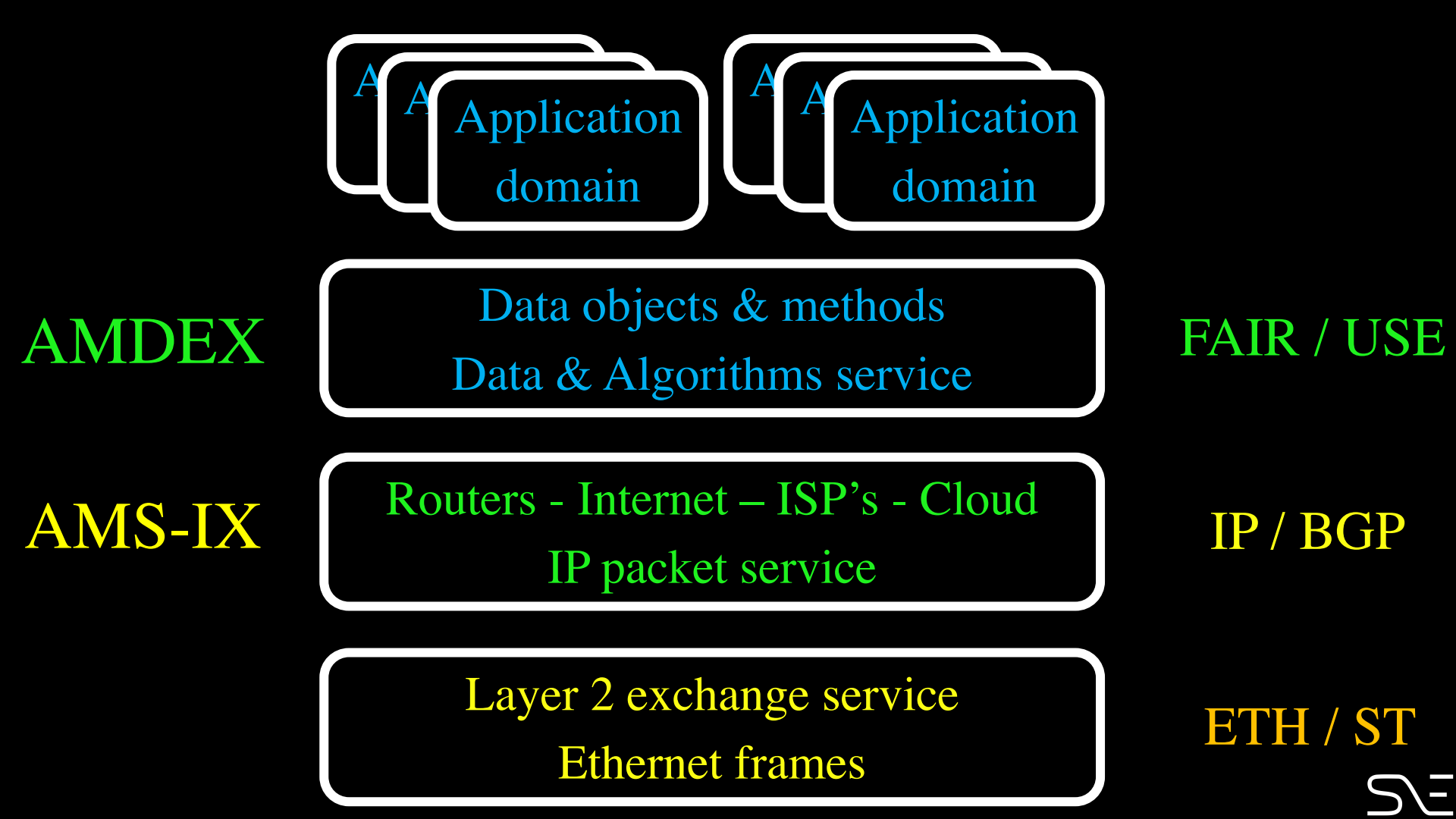
➔

Create an equal playing field

➔

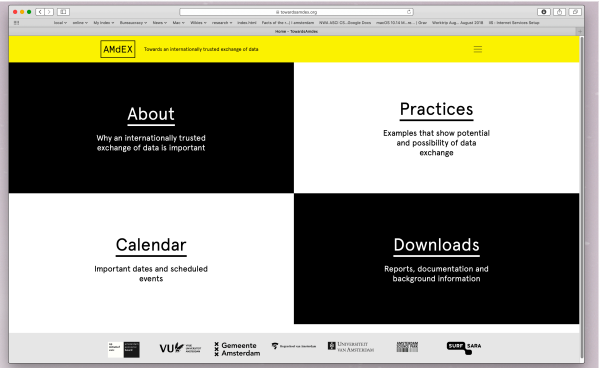
Sound Market principles

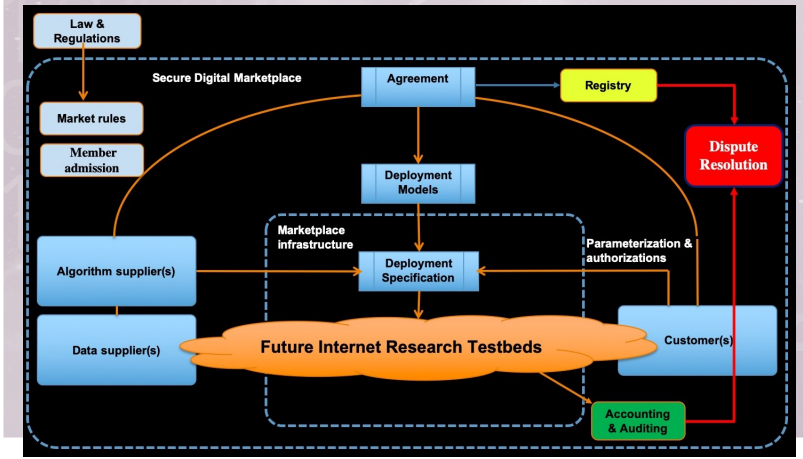
https://hbr.org/2017/09/managing-our-hub-economy


Application domain
Application domain
Data objects & methods
Data & Algorithms service
AMDEX
FAIR / USE
Routers - Internet – ISP's - Cloud
IP packet service
AMS-IX
IP / BGP
Layer 2 exchange service
Ethernet frames
ETH / ST

AMdEX.eu

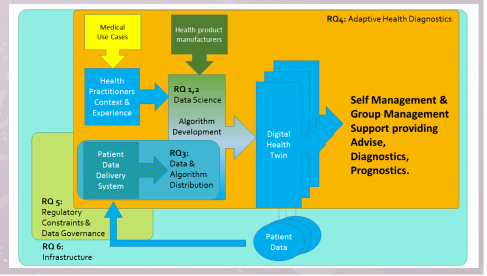
- Competing organisations, share data for common benefit
- Trust, Risk, data ownership & control
  - Industry: AF-KLM, Health, etc
  - Science: European Open Science Cloud
  - Society: Amsterdam Economic Board





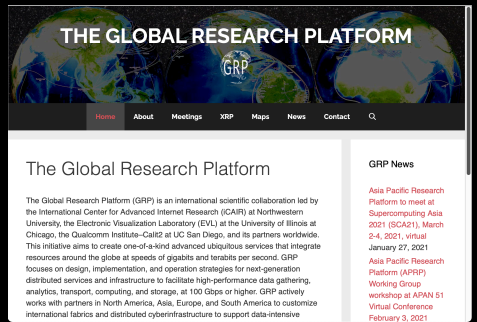
Aircraft Maintenance AF-KLM

Health: Enabling Personal Interventions



The Roaring Twenties!

- In the 90's the Internet was running on top of the telco's
- We freed it in the 2000's with GLIF and the *Lights
- We developed the computer science for virtualization of CI
- Networking is (almost) not the problem anymore (DMC2022...)
- Data and algorithms & apps and services are now in the cloud
- Just a few large players emerge with an almost monopoly
- Roaring 20's to free the Data with initiatives such as GRP!



Conclusions, Info, Acknowledgements, Q&A

- Data hindered by risk of unexpected use, lack of trust
- Using market principles, enforcement and determining incentives and value in the data life cycle to make data flow
- More information:
  - http://delaat.net/dl4ld http://delaat.net/epi
  - https://www.esciencecenter.nl/project/secconnet
  - https://towardsamdex.org