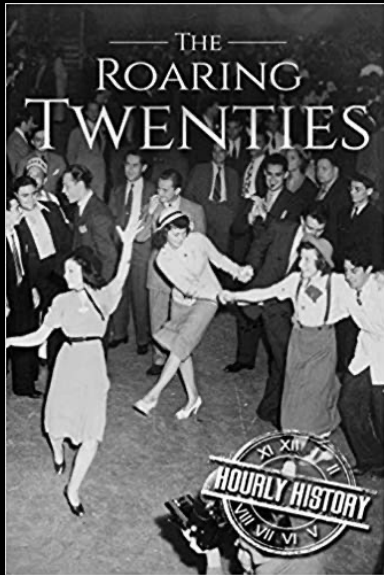


# ICT to support the transformation of Science in the Roaring Twenties

**Cees de Laat**

Systems and Networking Laboratory  
Complex Cyber Infrastructure 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

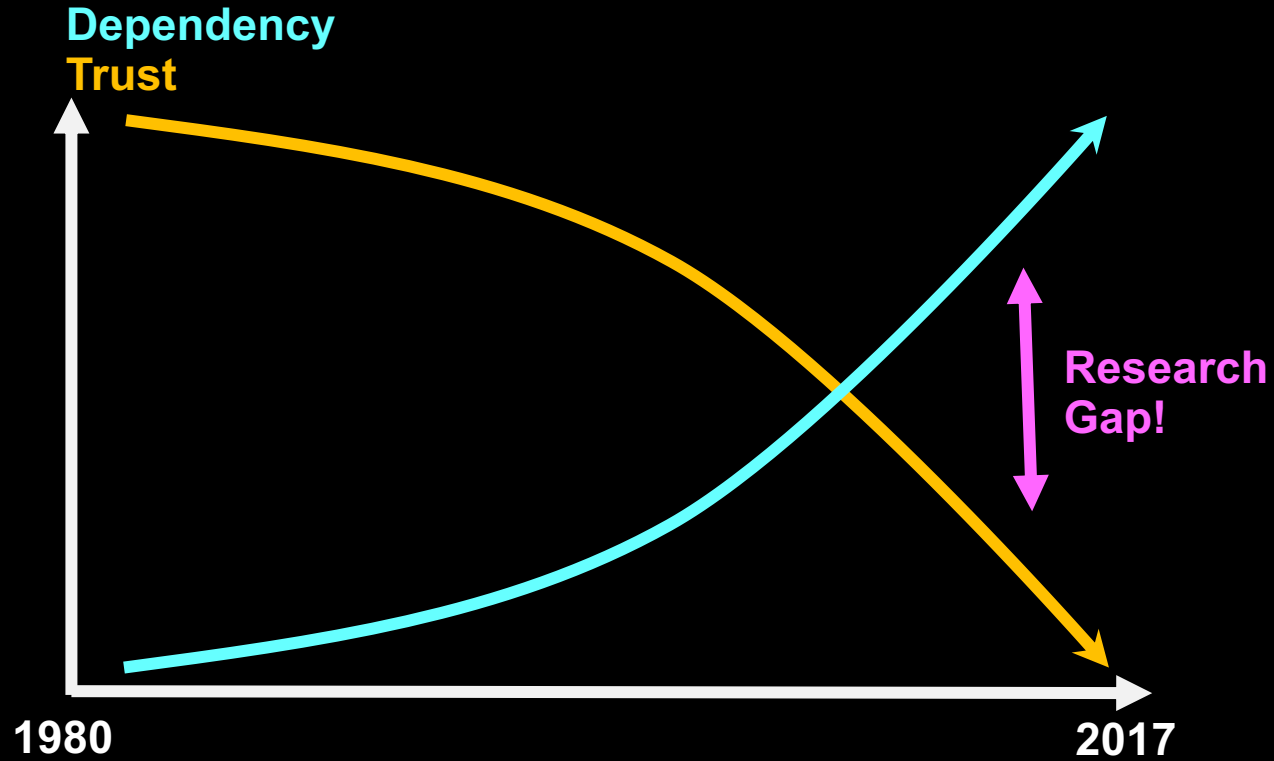
- Internet
- Computing
- Data
- Science

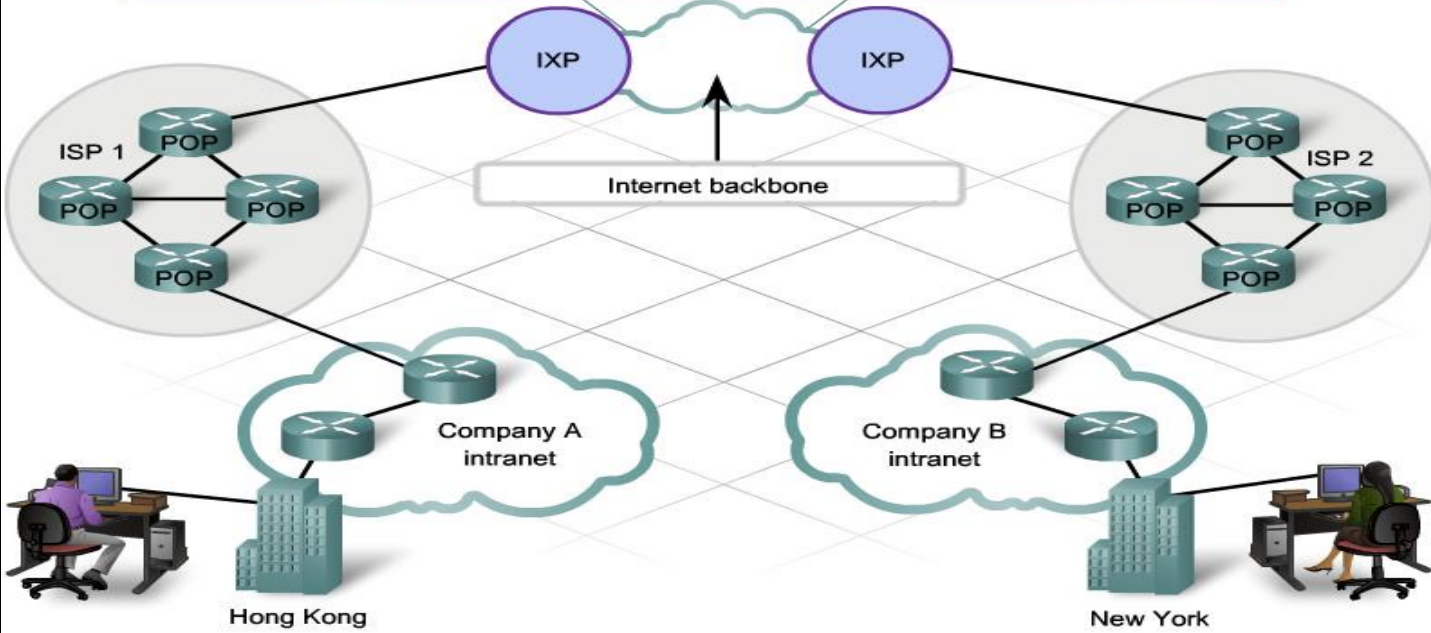
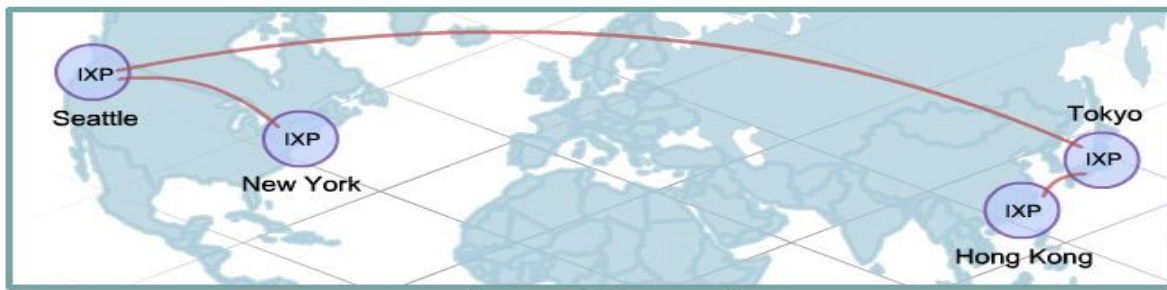
# Transformations

- Internet
  - From end to end to client server bubbles
- Computing
- Data
- Science



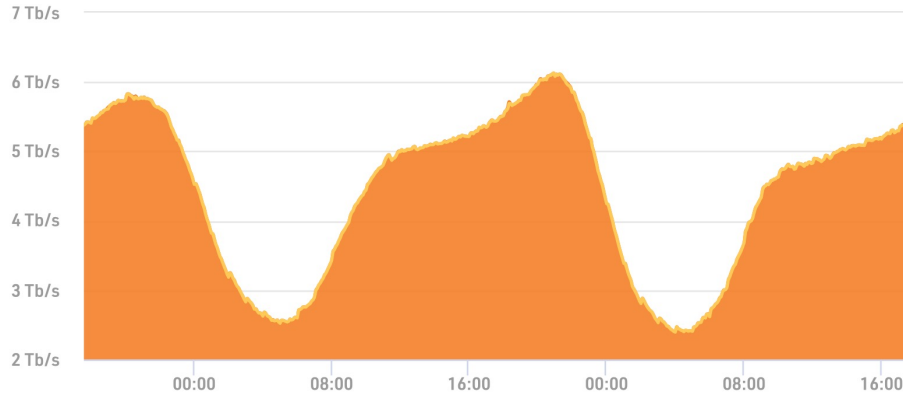
# Fading Trust in Internet





# Internet Backbone

TOTAL DAILY



● Input    ● Output

PEAK IN

**6.127** Tb/s

PEAK OUT

**6.132** Tb/s

AVERAGE IN

**4.431** Tb/s

AVERAGE OUT

**4.431** Tb/s

CURRENT IN

**5.394** Tb/s

CURRENT OUT

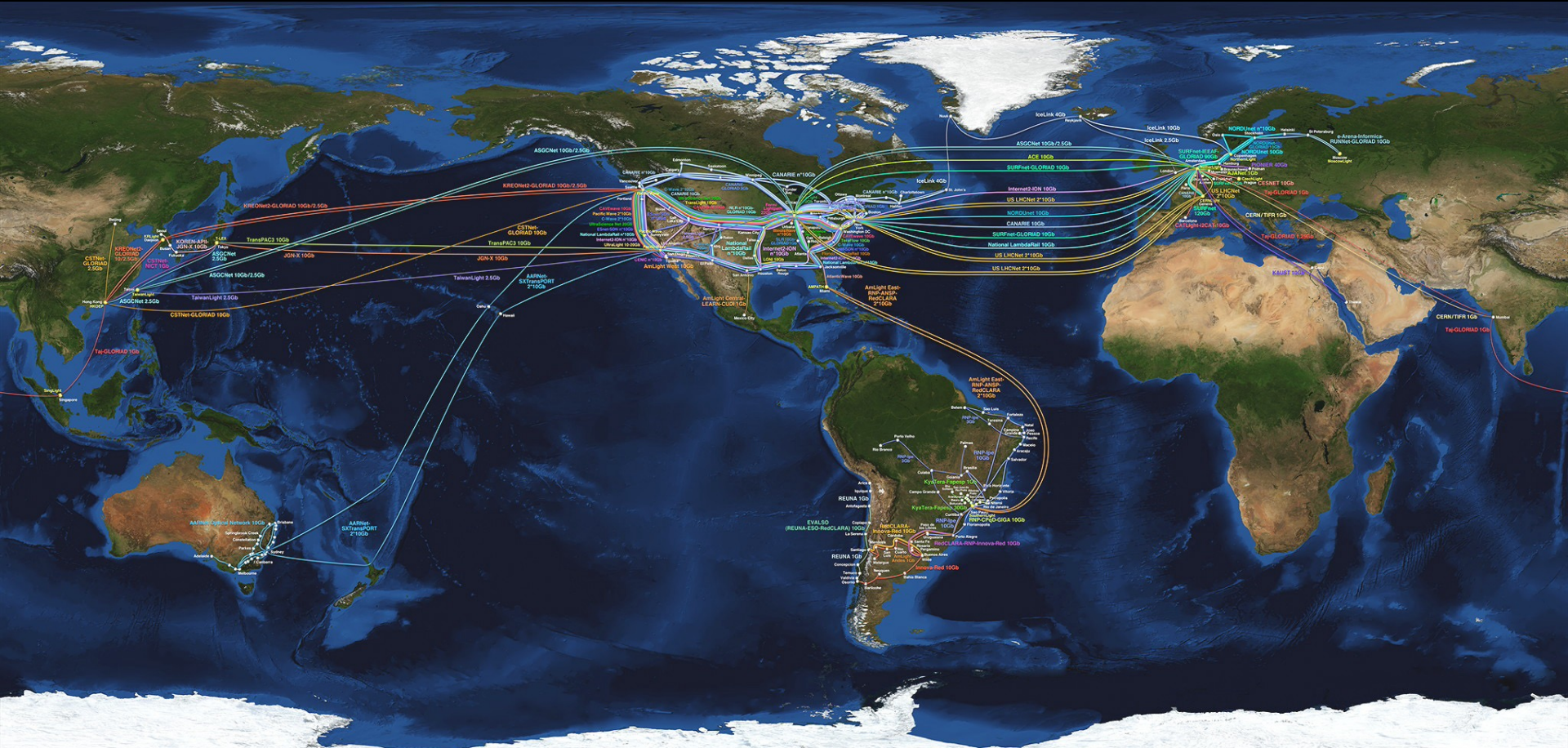
**5.391** Tb/s

**2019**

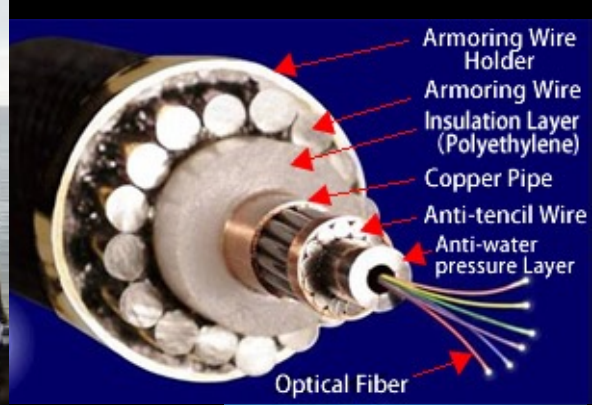




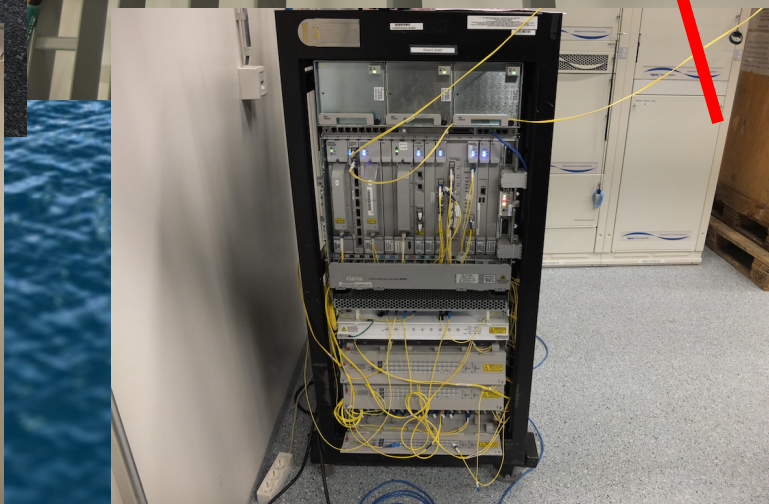
# The GLIF – LightPaths around the World





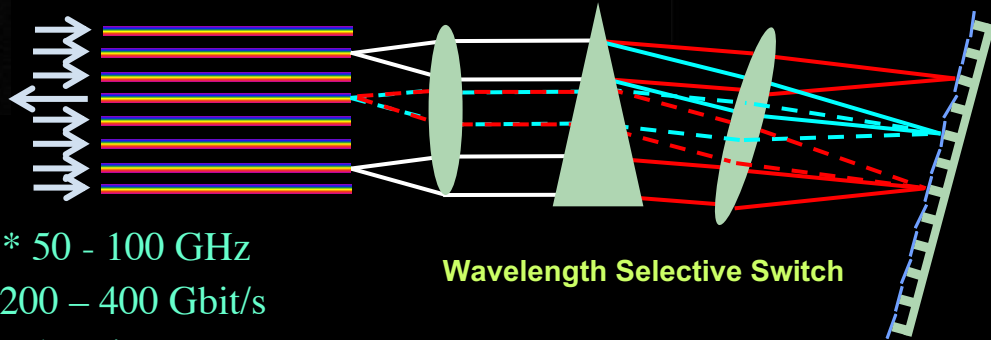
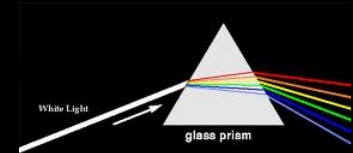
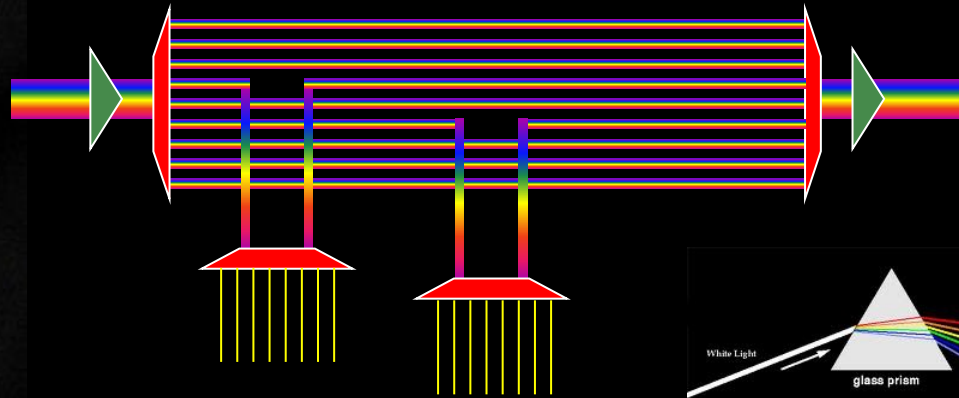
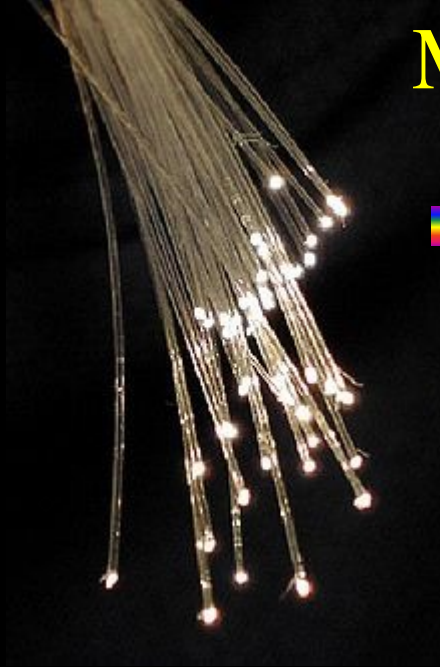


A **cable landing station** may or may not be required, depending on whether, for example, the submarine cable requires power to power submarine repeaters or amplifiers. The voltages applied to the cables can be high **3,000 to 4,000 volts** for a typical trans-Atlantic telecommunications cable system, and 1,000 volts for a cross-channel telecommunications cable system. Submarine power cables can operate at many kilovolts: for example, the [Fenno-Skan power cable operates at 400 kV DC](#).



# Undersea Cable HV

# Multiple colors / Fiber



Per fiber: ~ 80-100 colors \* 50 - 100 GHz  
Per color: 10 - 40 - 100 - 200 - 400 Gbit/s  
Max total: ~20 Tbit/s = ~2 Tbyte/s

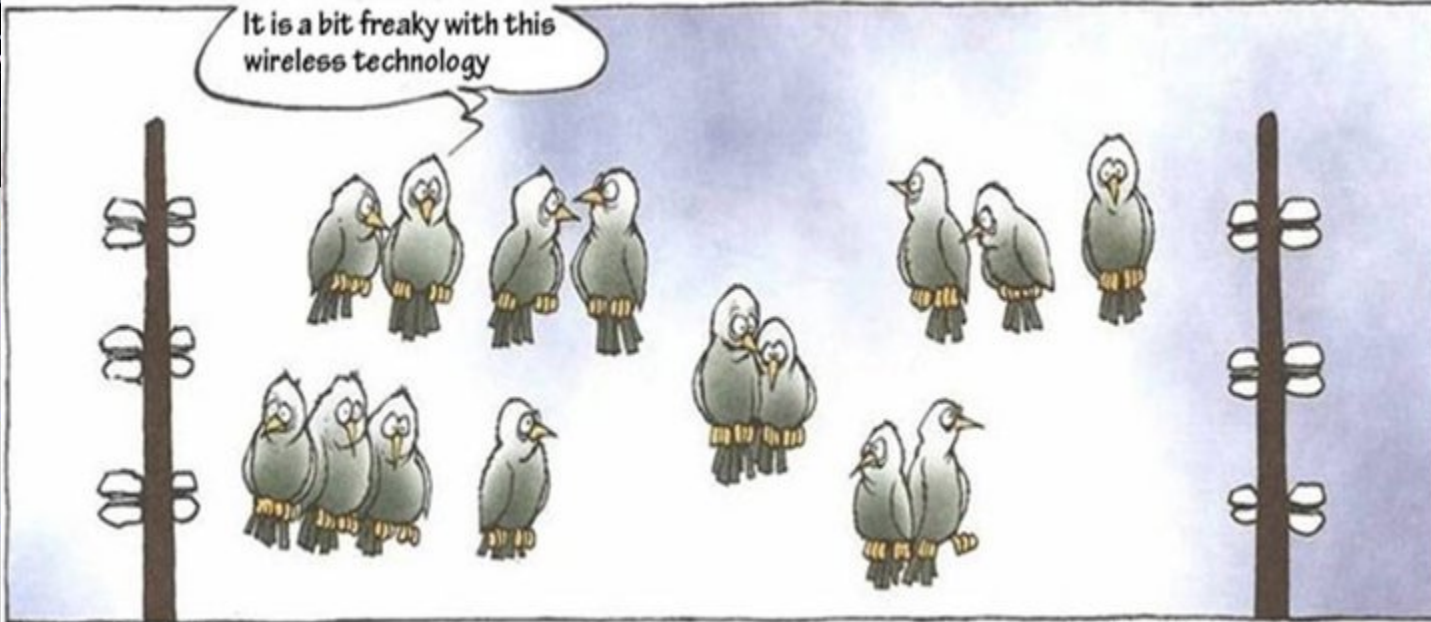
New: Hollow Fiber!

→ less RTT!





# Draadloze netwerken



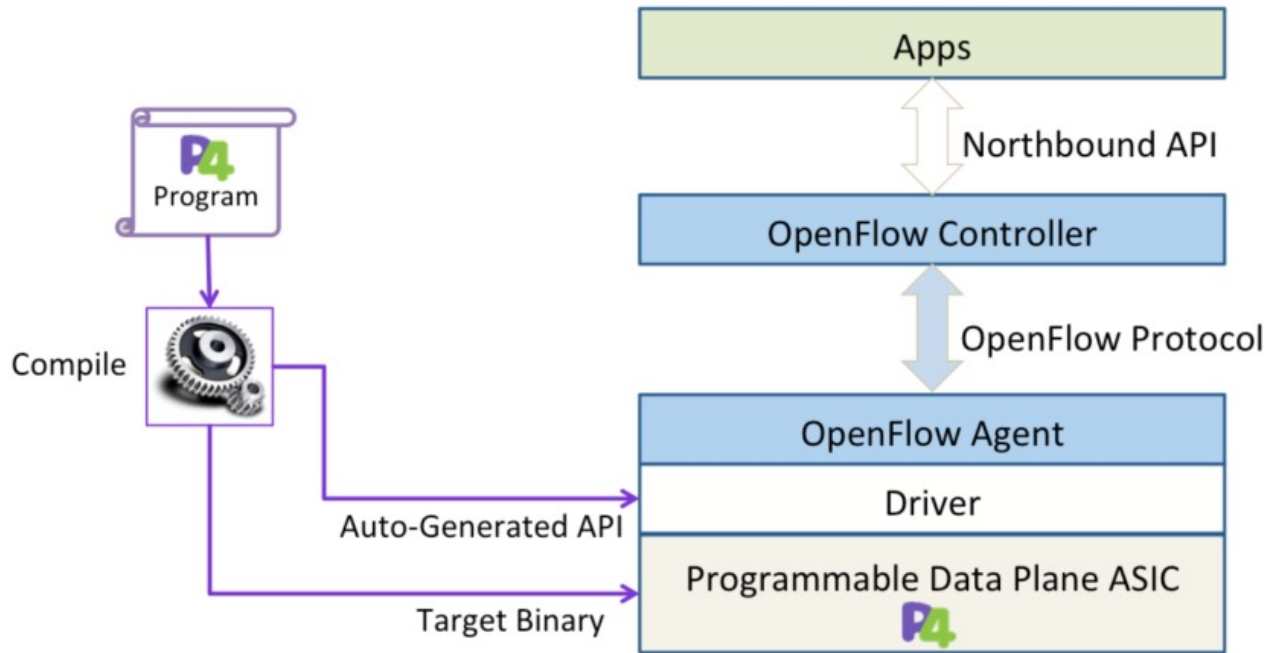
COPYRIGHT : MORTEN INGEMANN

protocol LAN due to the easy comparison and convenience in the **digital home**. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.





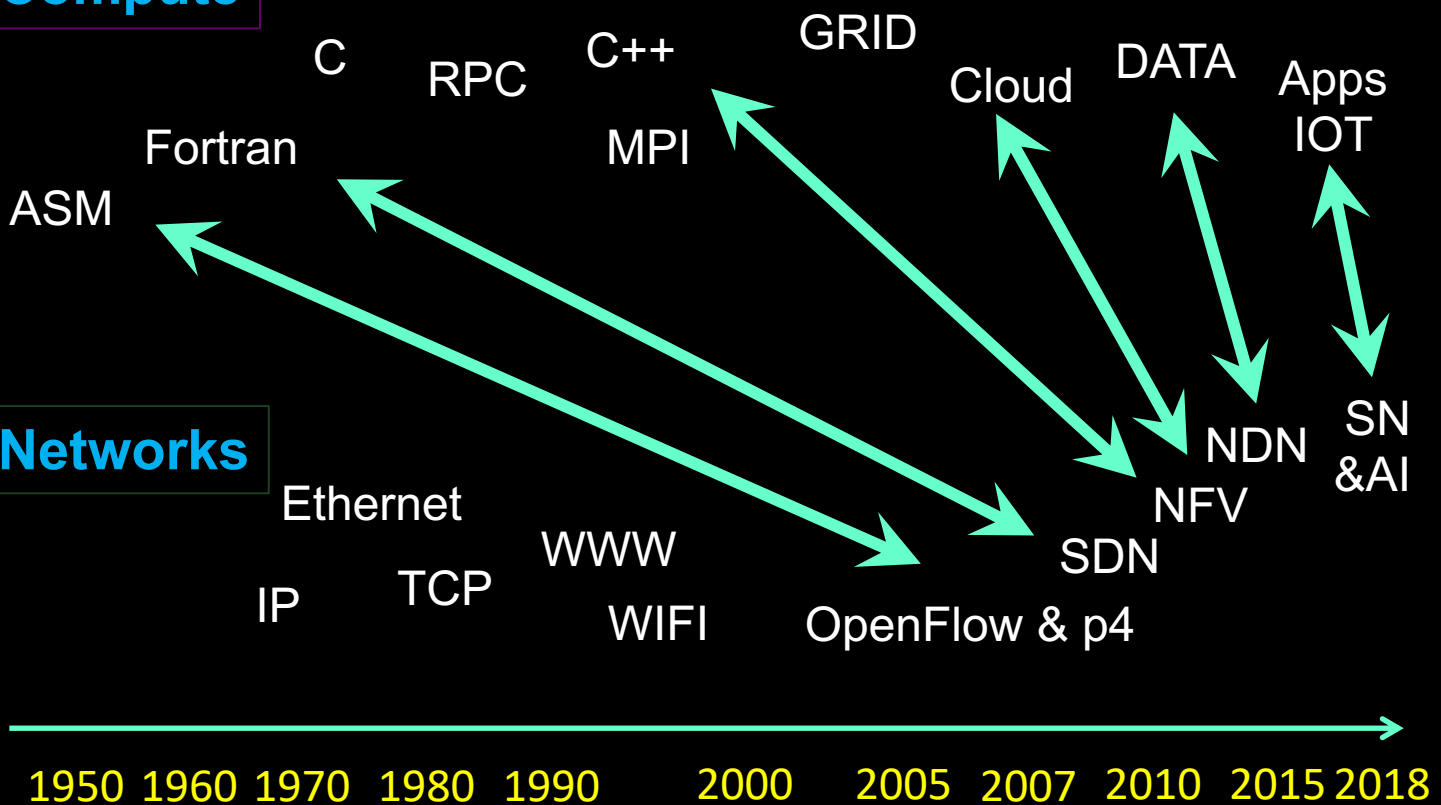
# P4 & OpenFlow



# TimeLine

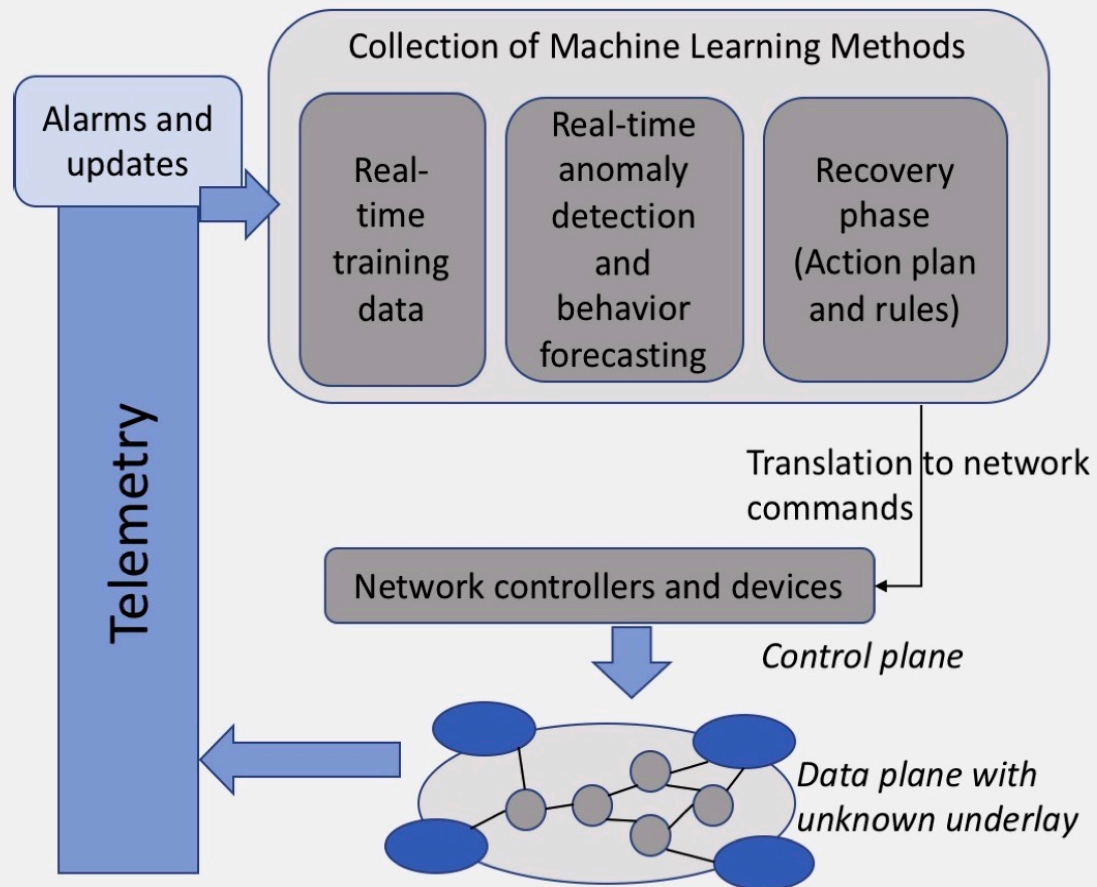
**Compute**

**Networks**



# 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, operation scale, and reliability 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 needs in advance and be faster, efficient, and more proactive, anticipating bottlenecks before they happen. However, to manage 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-



# Internet moves from IXP's into datacenters

LIMITED TIME OFFERS

FIRST MONTH FREE  
CHOOSE FROM 3 PACKAGES

Bundles

&

Fragmentation

THE FILTER  
BUBBLE

EXPLOSIVE  
What the  
Internet  
Is Hiding  
from You

ELI PARISER



# 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.

# Transformations

- Internet
  - From end to end to client server bubbles
- Computing
  - From Dinosaurs to Ant Colonies
- Data
- Science











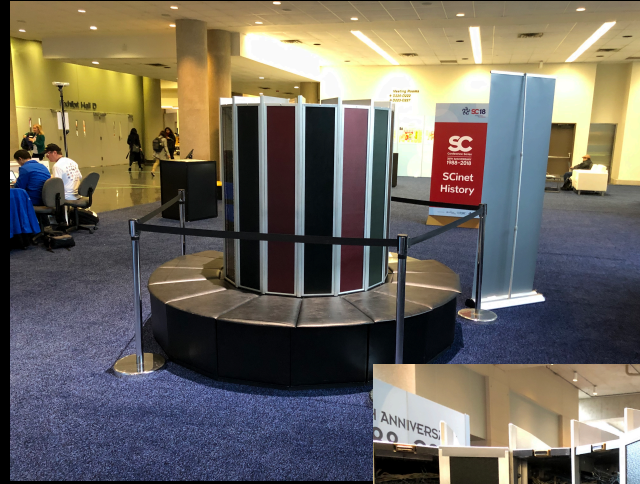
# Some progress



2018

= ~7x

? 540 MHz  
? MFlops  
1000 MByte memory  
16000 MByte ssd  
0,0012 kWh – 18 h



1976



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



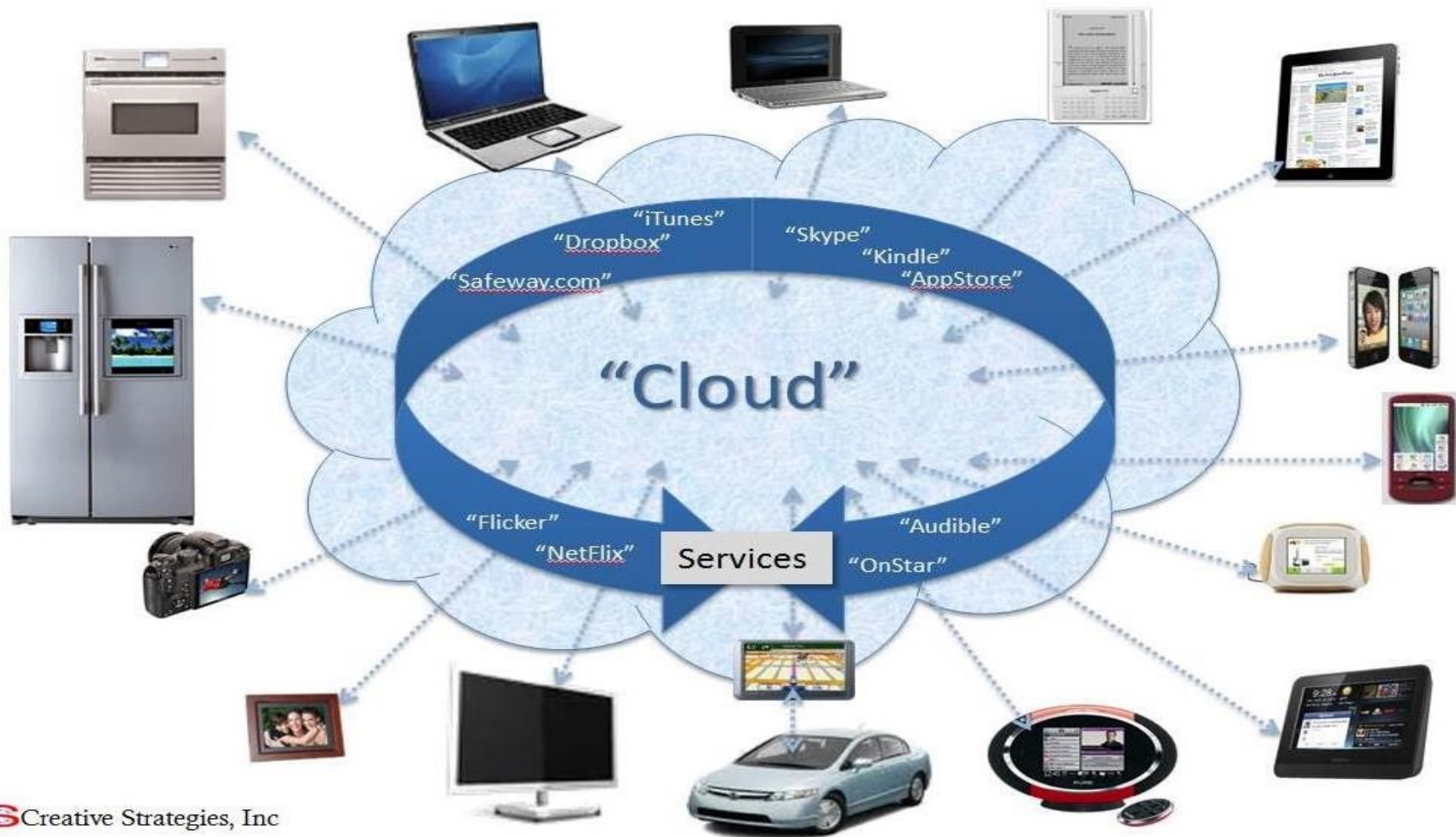


# Change in computing

- 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



# Internet of Things



# Transformations

- Internet
  - From end to end to client server bubbles
- Computing
  - From Dinosaurs to Ant Colonies
- Data
  - From putting Data in the Cloud to getting it back
- Science

# Now, how do we get and use data?

## 2019 *This Is What Happens In An Internet Minute*



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

# Harvard Business Review



Harvard Business Review


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 COMMENT H 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

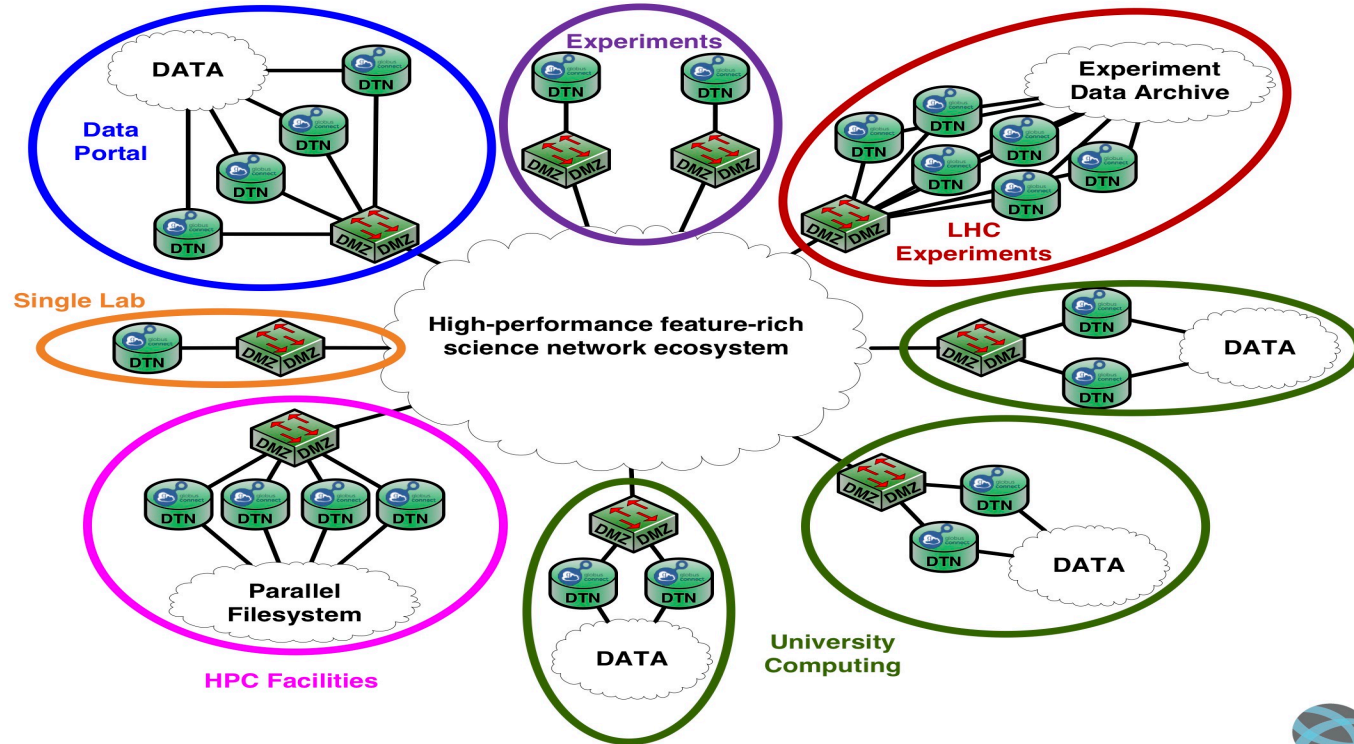


Sound Market  
principles

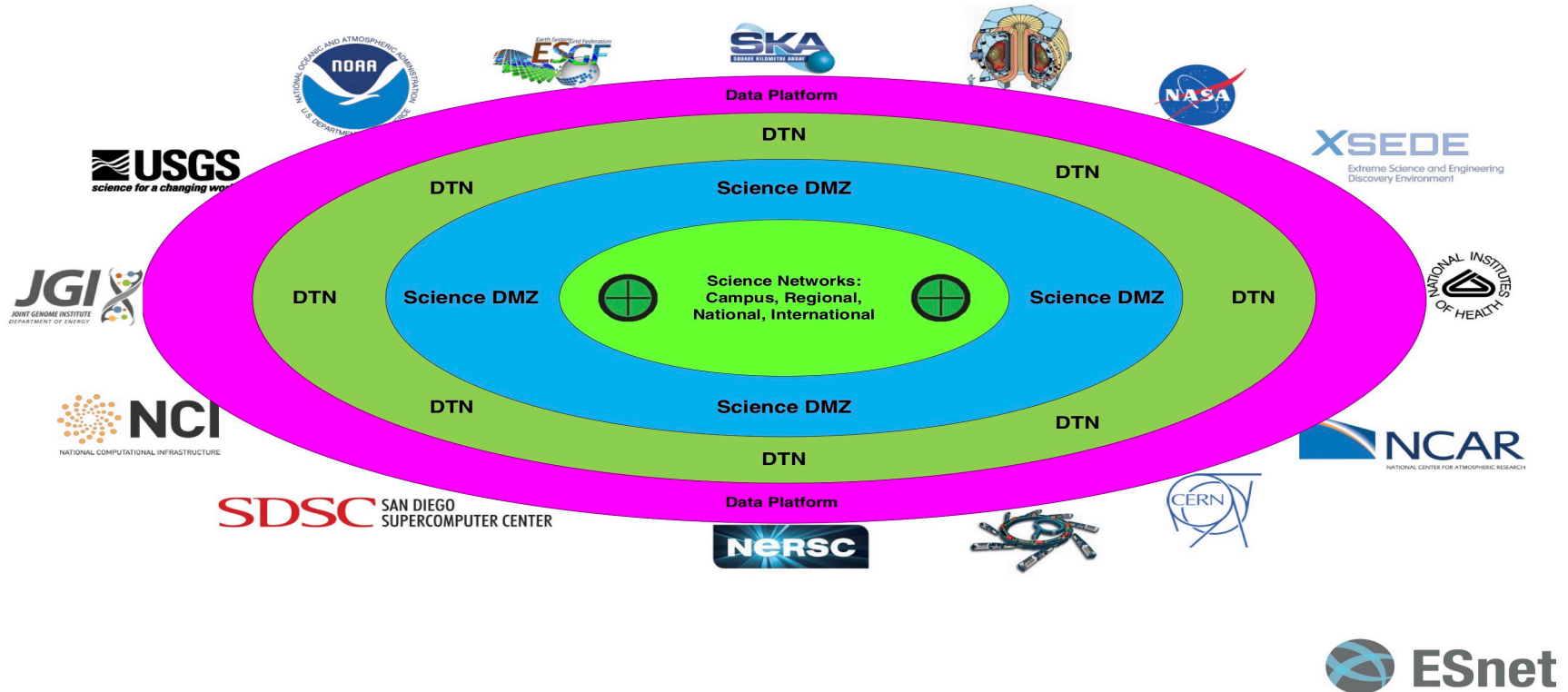
<https://hbr.org/2017/09/managing-our-hub-economy>



# Science DMZs for Science Applications



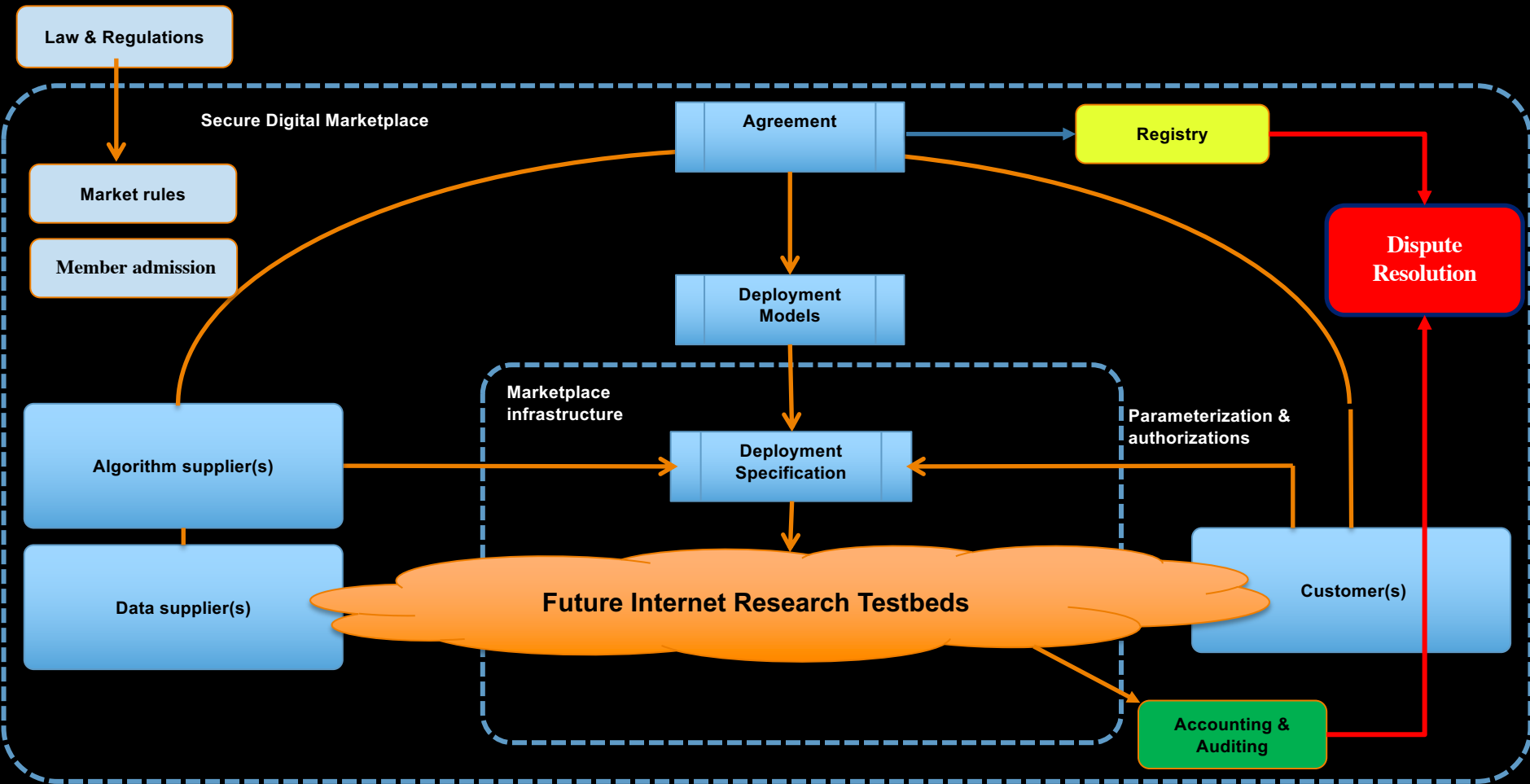
# Data Ecosystem – Concentric View



# Data Sharing: Main problem statement

- Organizations that normally compete have to bring data together to achieve a common goal!
- The shared data may be used for that goal but not for any other!
- Data or Algorithms may have to be processed in foreign data centers.
  - How to organize alliances?
  - How to translate from strategic via tactical to operational level?
  - How to enforce policy using modern Cyber Infrastructure?
  - What are the different fundamental data infrastructure models to consider?

# Secure Digital Market Place Research



# Big Data Sharing use cases placed in airline context



Global Scale



Aircraft Component Health  
Monitoring (Big) Data  
NWO **CIMPLO** project  
4.5 FTE

National Scale



Cargo Logistics Data  
(C1) DaL4LoD  
(C2) Secure scalable  
policy-enforced  
distributed data  
Processing  
(using blockchain)



Cybersecurity Big Data  
NWO **COMMIT/**  
**SARNET** project  
3.5 FTE

Campus /  
Enterprise Scale

NLIP **iShare** project

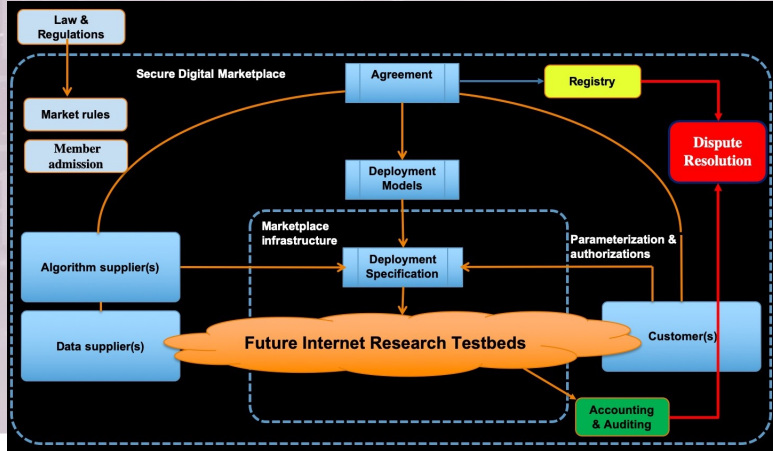
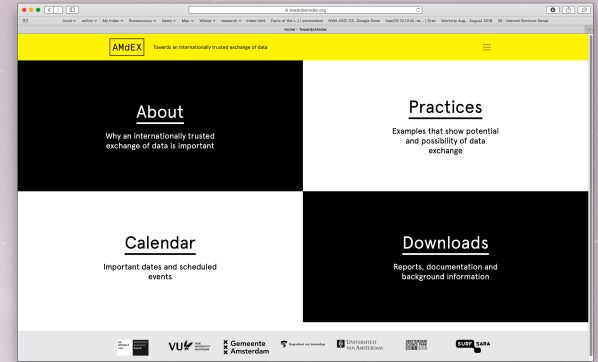


**iSHARE**  
powered by NLIP



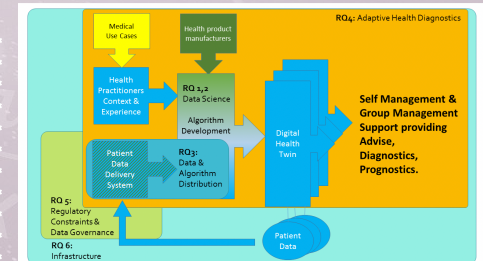
## AMdEX.eu

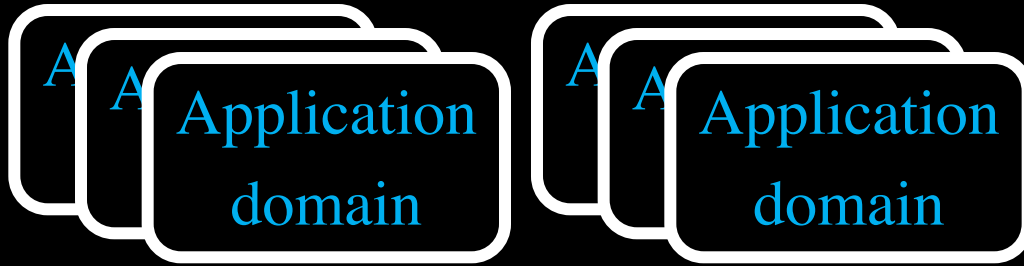
- Competing organisations, share data for common benefit
- Trust, Risk, data ownership & control
  - Industry: AF-KLM, Health, etc
  - Science: European Open Science Cloud
  - Society: Amsterdam Economic Board



Aircraft Maintenance AF-KLM

Health:  
Enabling  
Personal  
Interventions





AMDEX

Data objects & methods  
Data & Algorithms service

FAIR / USE

AMS-IX

Routers - Internet – ISP's - Cloud  
IP packet service

IP / BGP

Layer 2 exchange service  
Ethernet frames

ETH / ST

# Transformations

- Internet
  - From end to end to client server bubbles
- Computing
  - From Dinosaurs to Ant Colonies
- Data
  - From putting Data in the Cloud to getting it back
- Science
  - Pulling it all together



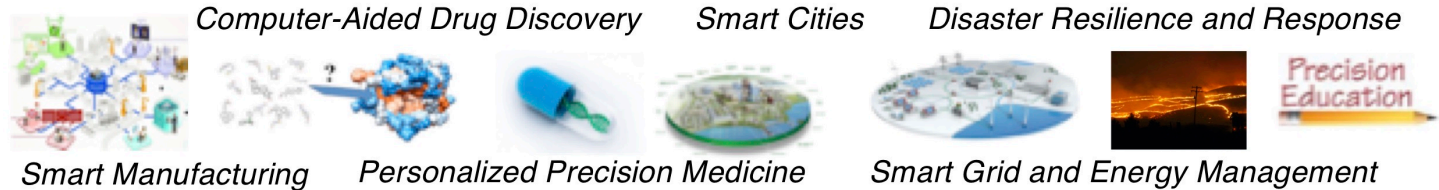
**In most applications, utilization of Big Data often needs to be combined with Scalable Computing.**



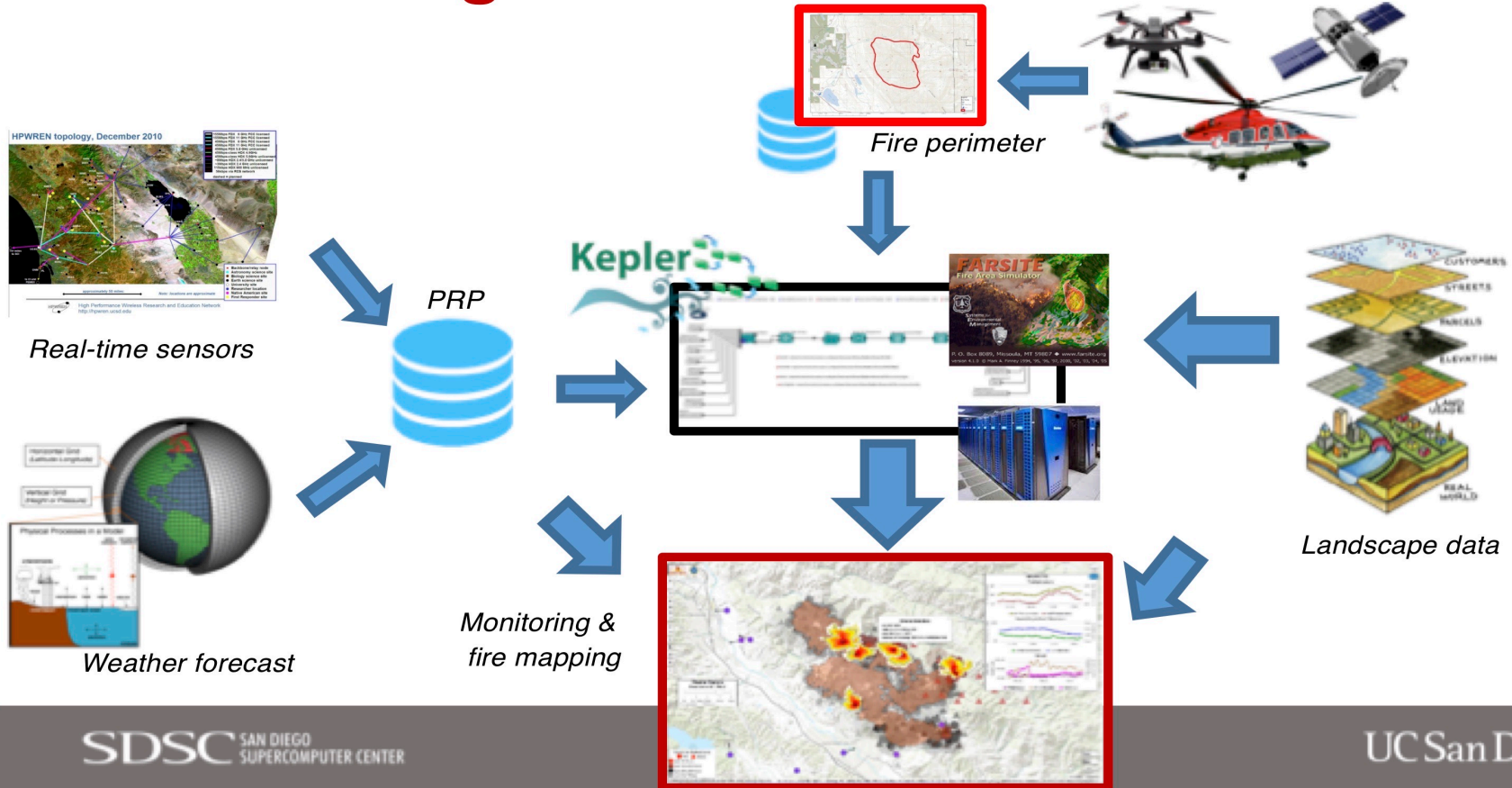
*COMPUTING AT DIVERSE SCALES*

*“BIG” DATA*

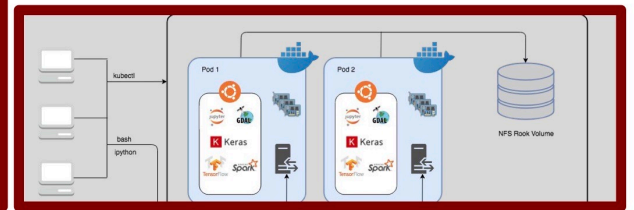
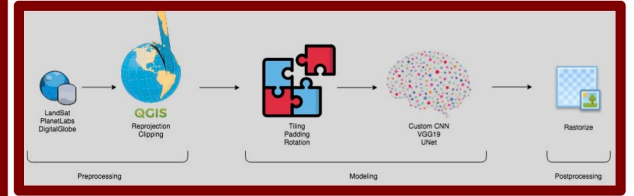
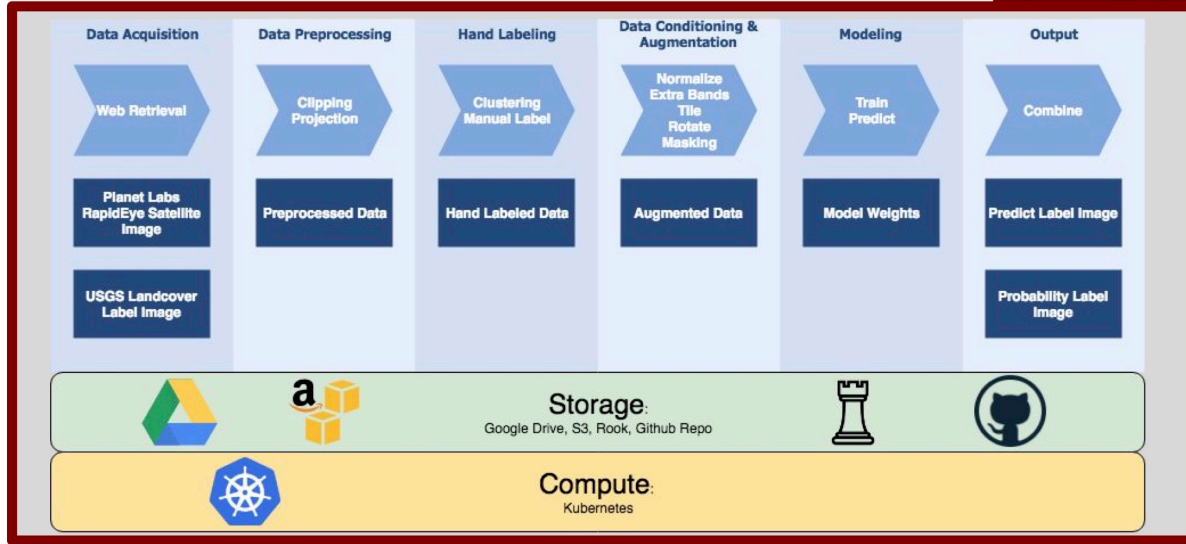
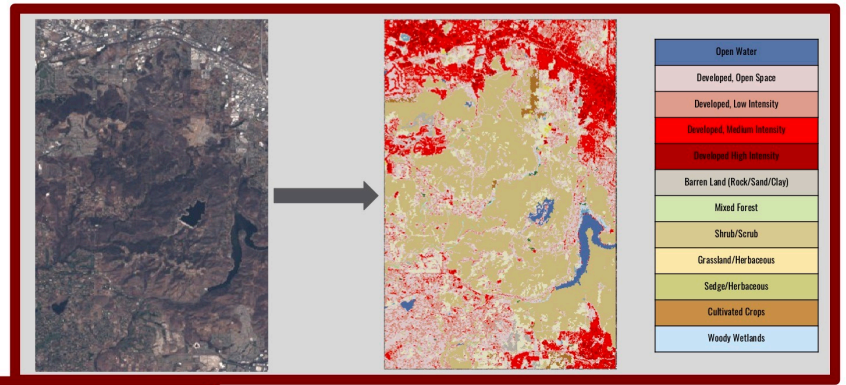
*Enables dynamic data-driven applications*



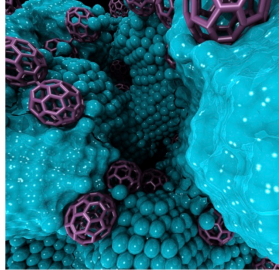
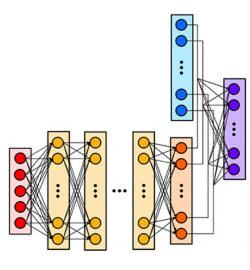
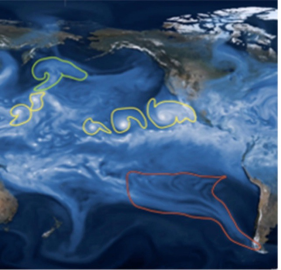
# Fire Modeling Workflows in WIFIRE



# One Piece of the Puzzle: Vegetation Classification using Satellite Imagery







# Scientific Machine Learning & Artificial Intelligence

Scientific progress will be driven by

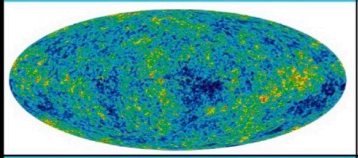
- Massive data: sensors, simulations, networks
- Predictive models and adaptive algorithms
- Heterogeneous high-performance computing

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

## BASIC RESEARCH NEEDS FOR Scientific Machine Learning Core Technologies for Artificial Intelligence

### EXEMPLARS OF SCIENTIFIC ACHIEVEMENT

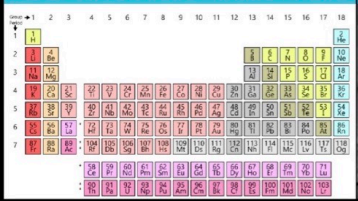
Cosmic Microwave Background



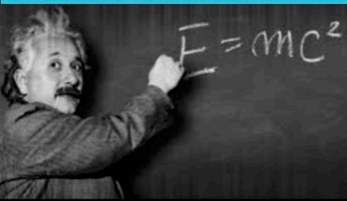
DNA Structure



Periodic Table of the Elements



Special Relativity

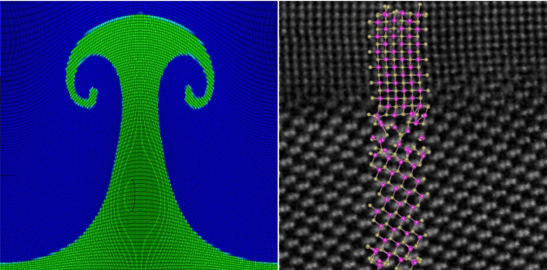


Human-AI insights enabled via scientific method, experimentation, & AI reinforcement learning.

**POWER GRID INPUTS**

Wind  
Solar  
Dams  
Nuclear

$\begin{aligned}
 \text{Reward} &= \begin{cases} c_1 \sum \Delta V_i - c_2 \sum \Delta P_i(P_i, R_i) - c_3 P_{\text{total}} \\ -1000, \text{ if } V_i(t) < 0.95, T_{\text{total}} = 4 < t \\ \min \{P_i(t) - 0.7, 0\}, \text{ if } T_{\text{total}} < t < T_{\text{total}} + 0.33 \\ \min \{P_i(t) - 0.8, 0\}, \text{ if } T_{\text{total}} + 0.33 < t < T_{\text{total}} + 0.5 \\ \min \{P_i(t) - 0.9, 0\}, \text{ if } T_{\text{total}} + 0.5 < t < T_{\text{total}} + 1.5 \\ \min \{P_i(t) - 0.95, 0\}, \text{ if } T_{\text{total}} + 1.5 < t \end{cases} \\
 \Delta V_i(t) &= \dots
 \end{aligned}$



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



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

**Workshop report:**  
<https://www.osti.gov/biblio/1478744>

# The Big Data Challenge

Doing Science

ICT to enable Science

Wisdom

AI

Knowledge to act

Analytics  
Decision Support

Information

Web/OWL

Data  
a.o. from ESFRI's

Docker, VM,  
XML, RDF, rSpec, SNMP



# The Big Data Challenge

Doing Science

ICT to enable Science

Wisdom

AI

Scientists live here!

Interdisciplinary Science App Store

Analytics library / Github / etc

Knowledge to act

Analytics Decision Support

MAGIC DATA CARPET

curation - description - trust - security - policy - integrity

Information

Web/OWL

Data

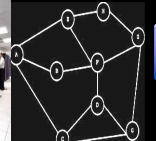
a.o. from ESFRI's

Docker, VM, XML, RDF, rSpec, SNMP

DSC eScience

RDM/DANS

ICT/SURF



Why?



**Because we can!**



# Questions?



<http://delaat.net>

<http://sne.science.uva.nl>

<http://www.os3.nl/>

<http://sne.science.uva.nl/openlab/>

<http://pire.opensciencedatacloud.org>

<http://staff.science.uva.nl/~delaat/pire/>

<https://rd-alliance.org>

<http://envri.eu>



Amsterdam Data Science



Supported by: