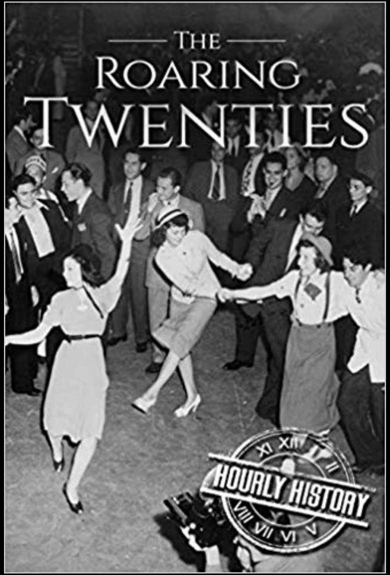
ICT to support the transformation of Science in the Roaring Twenties

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.

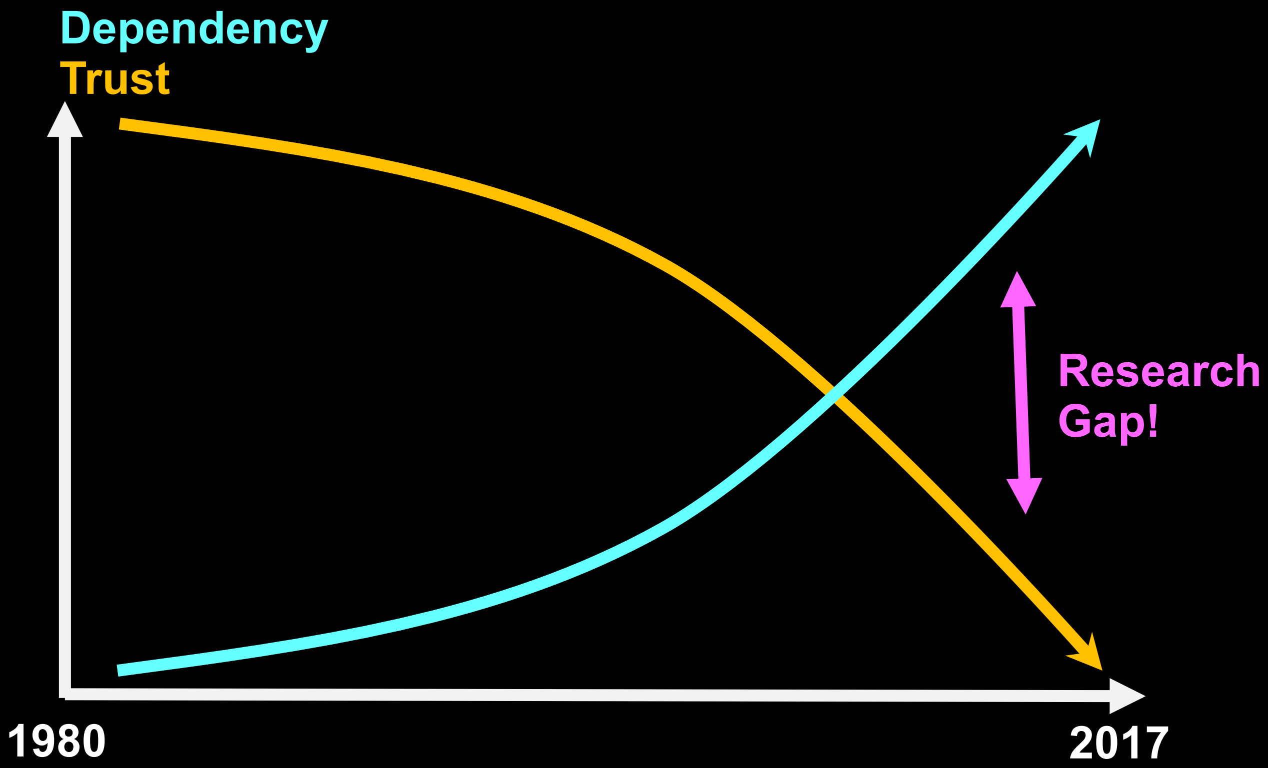
Transformations

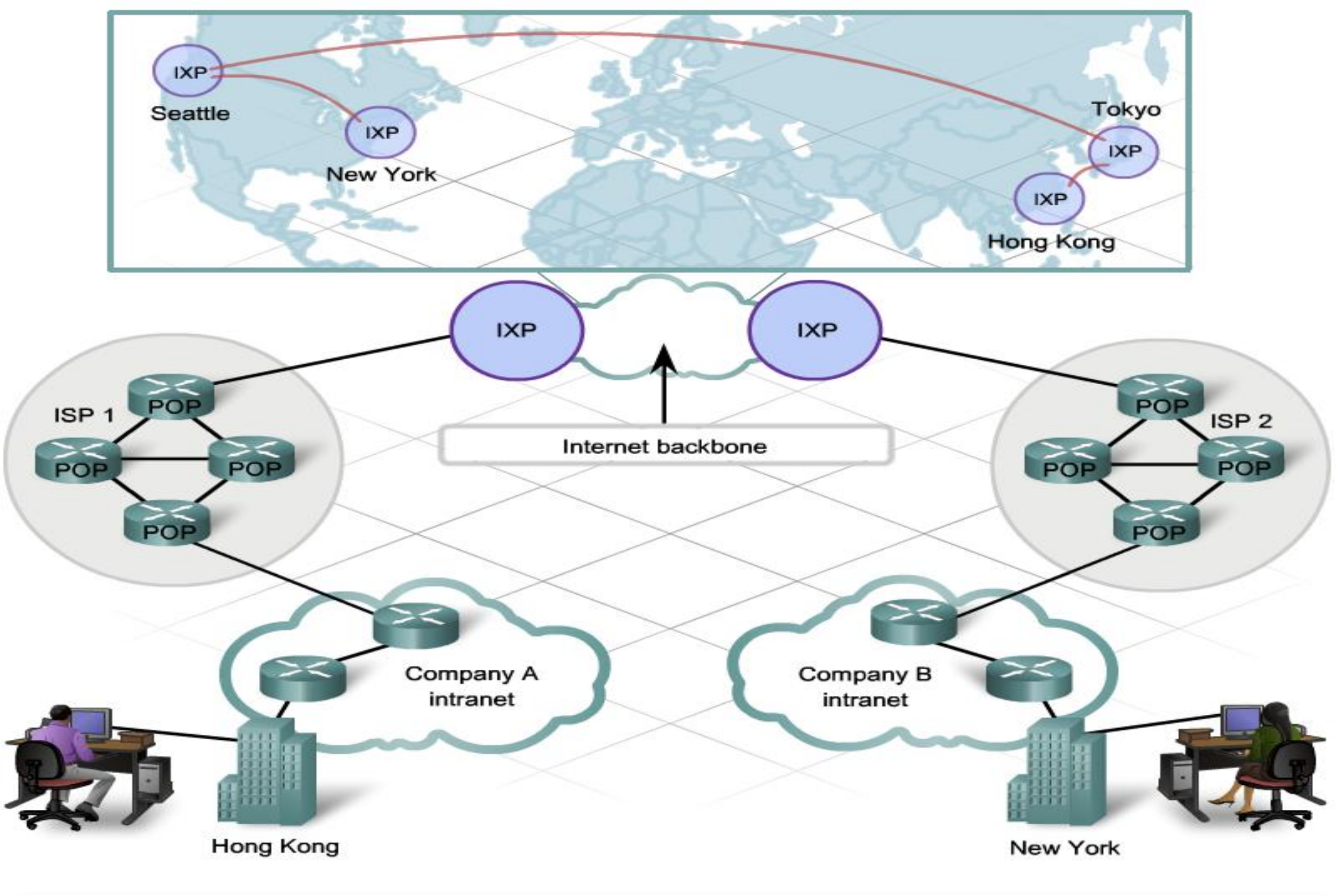
- Internet
- Computing
- Data
- Science

Transformations

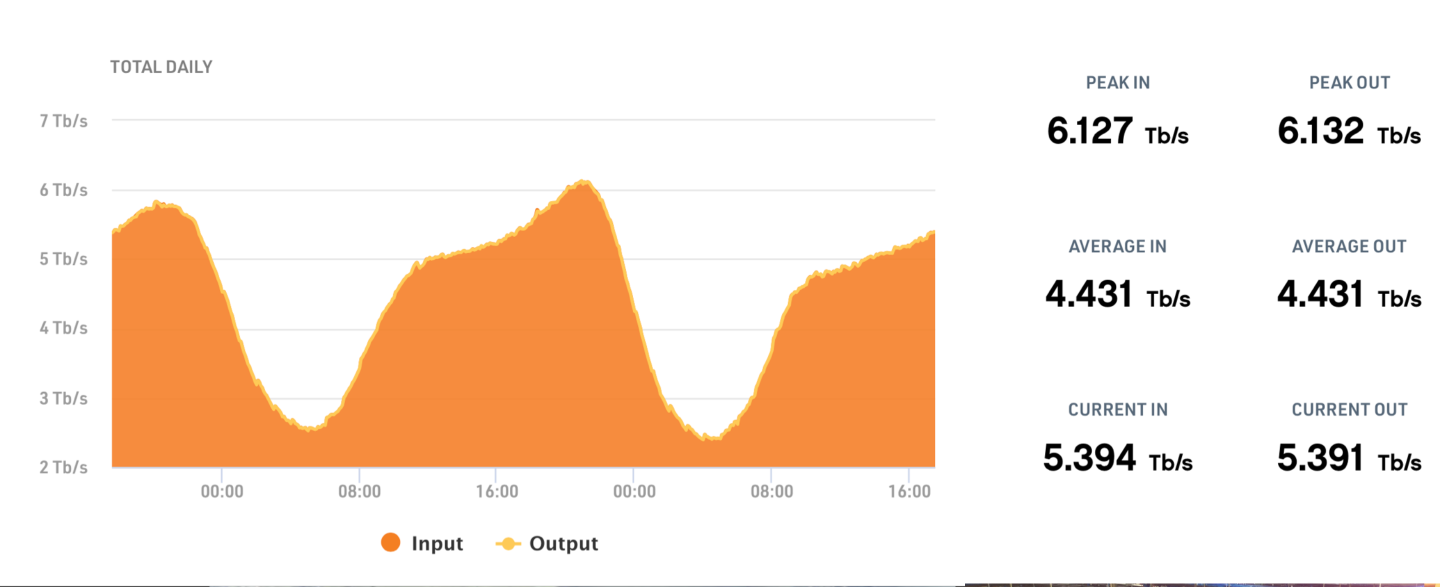
- Internet
  - From end to end to client server bubbles
- Computing
- Data
- Science

Fading Trust in Internet





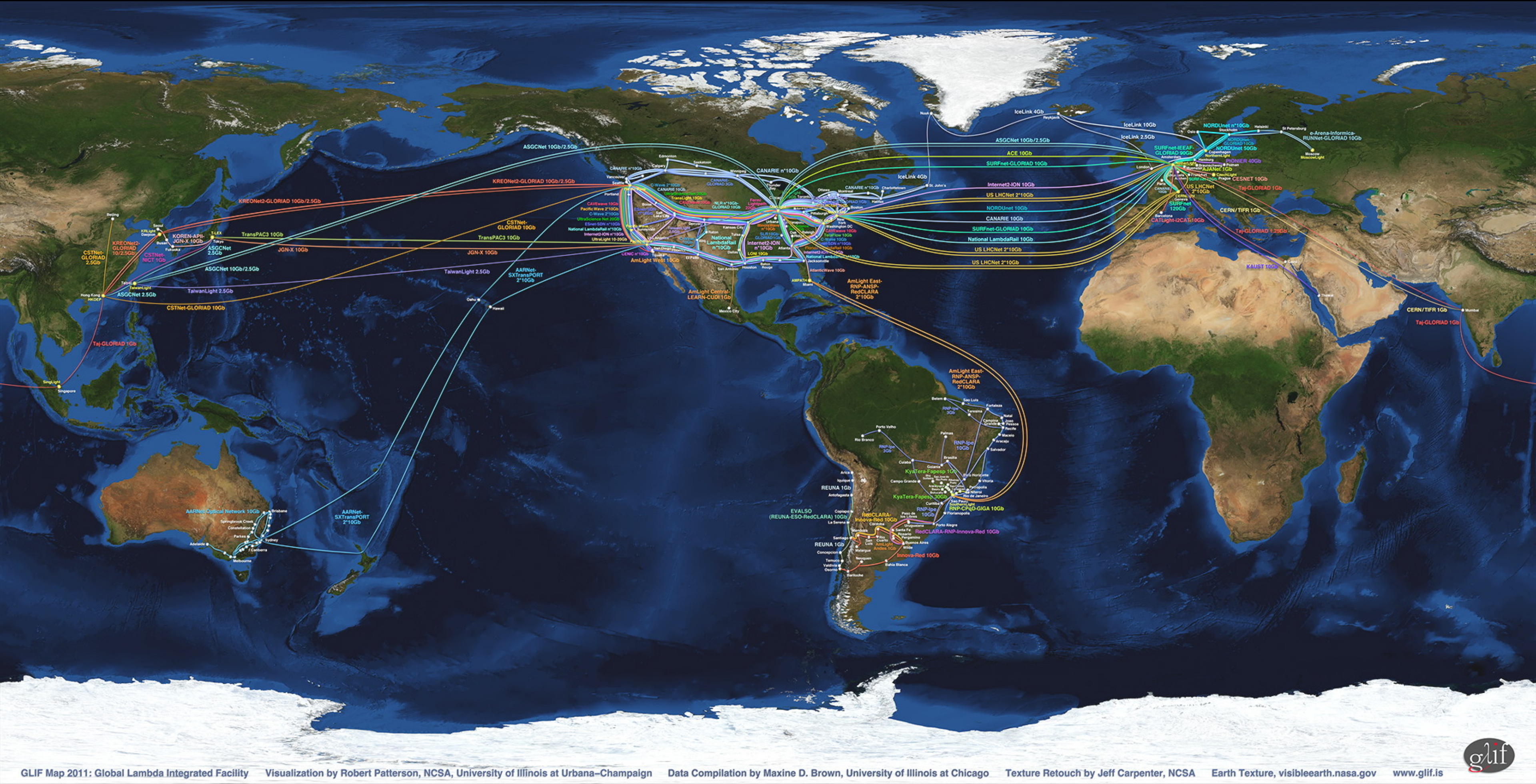
Internet Backbone

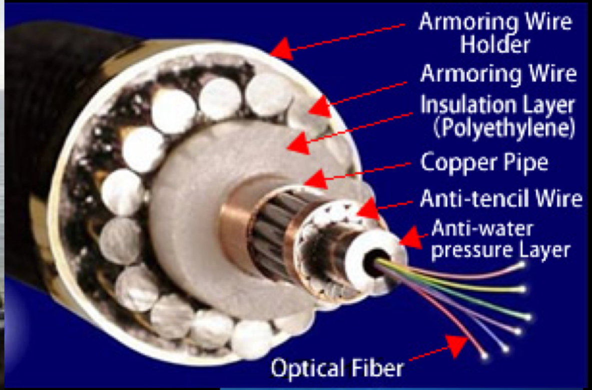


2019

amsix amsterdam internet exchang

The GLIF – LightPaths around the World

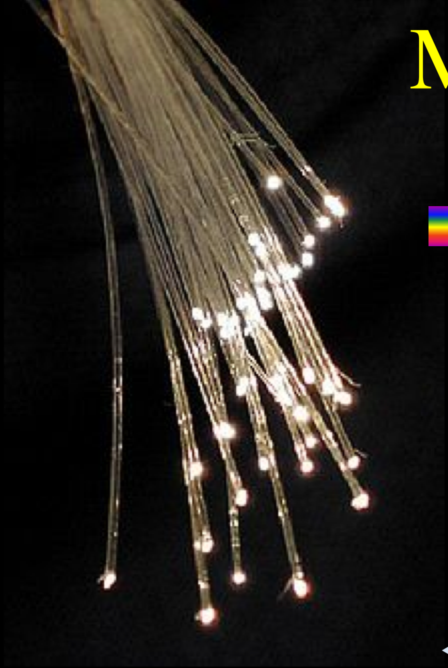


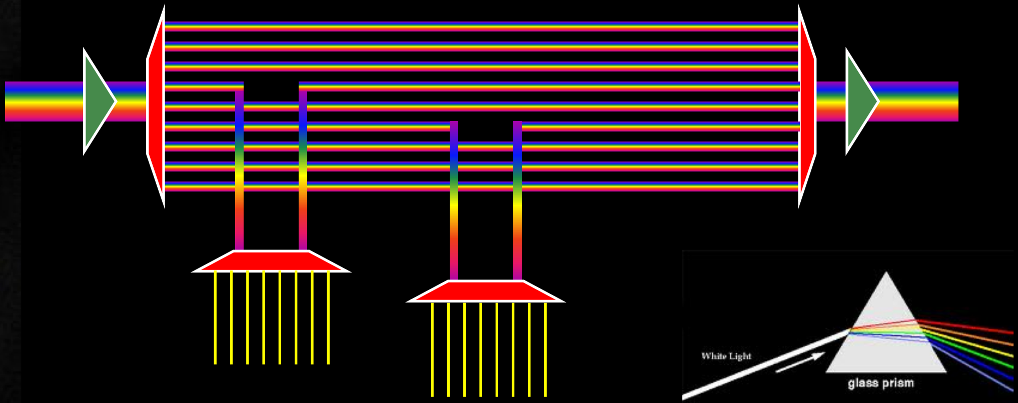


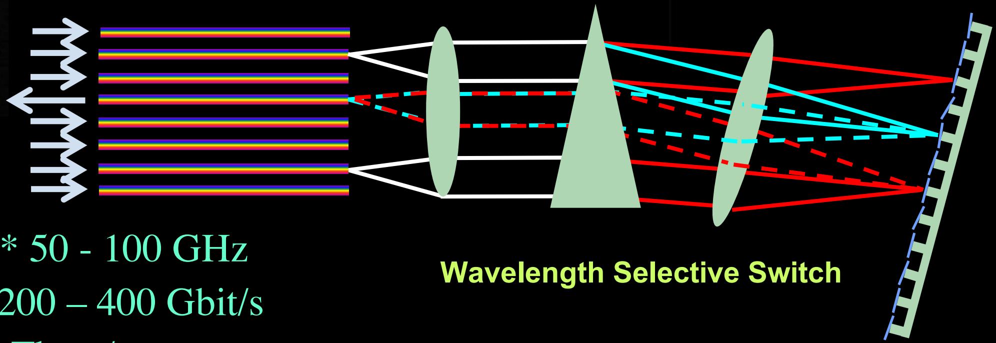
A **cable landing station** may or may not be required, depending on whether, for example, the submarine cable requires power to power submarine repeaters or amplifiers. The voltages applied to the cables can be high **3,000 to 4,000 volts** for a typical trans-Atlantic telecommunications cable system,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

Undersea Cable HV

Multiple colors / Fiber







Per fiber: ~ 80-100 colors * 50 - 100 GHz

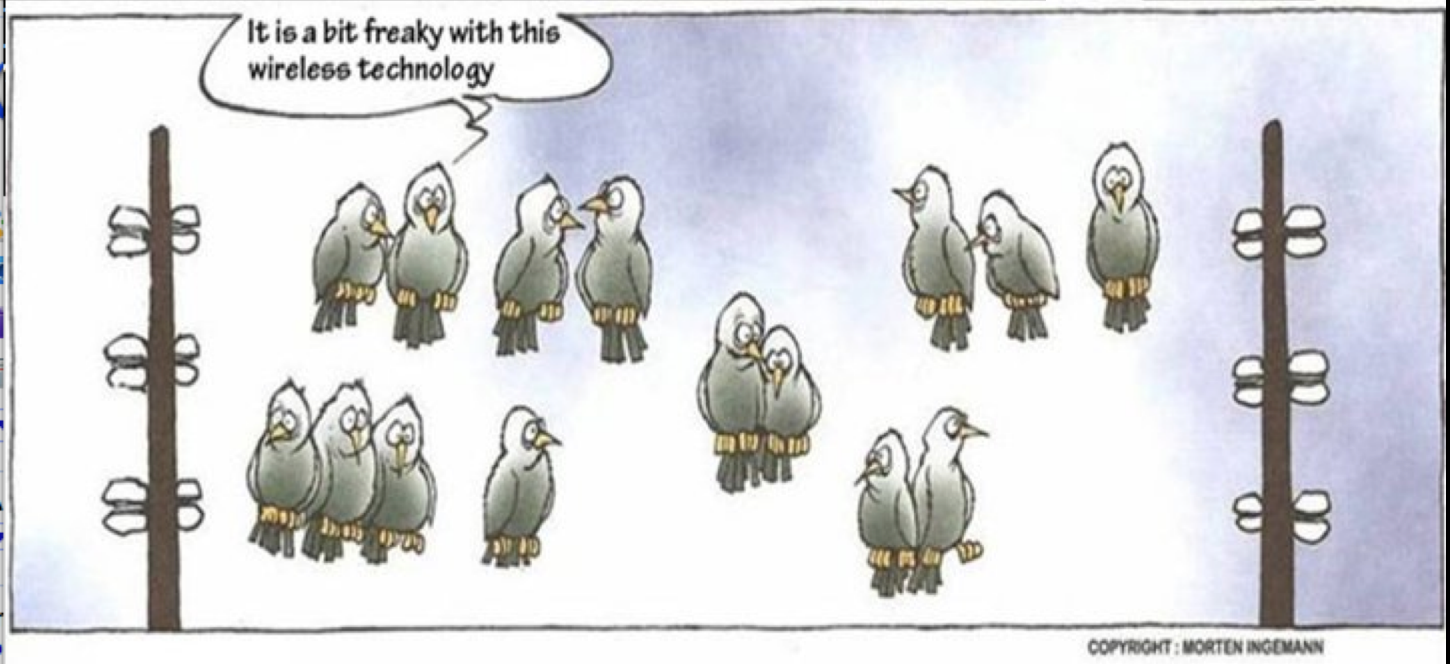
Per color: 10 – 40 – 100 - 200 – 400 Gbit/s

Max total: ~20 Tbit/s = ~2 Tbyte/s

New: Hollow Fiber!

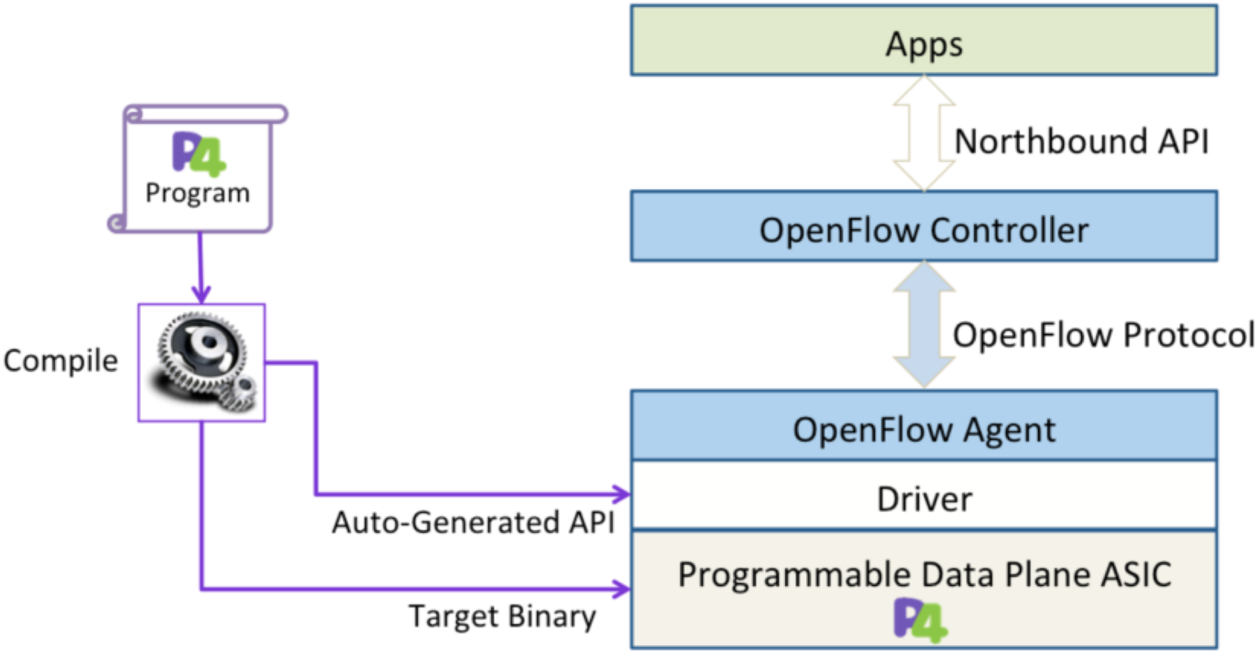
➔ less RTT!

Draadloze netwerken



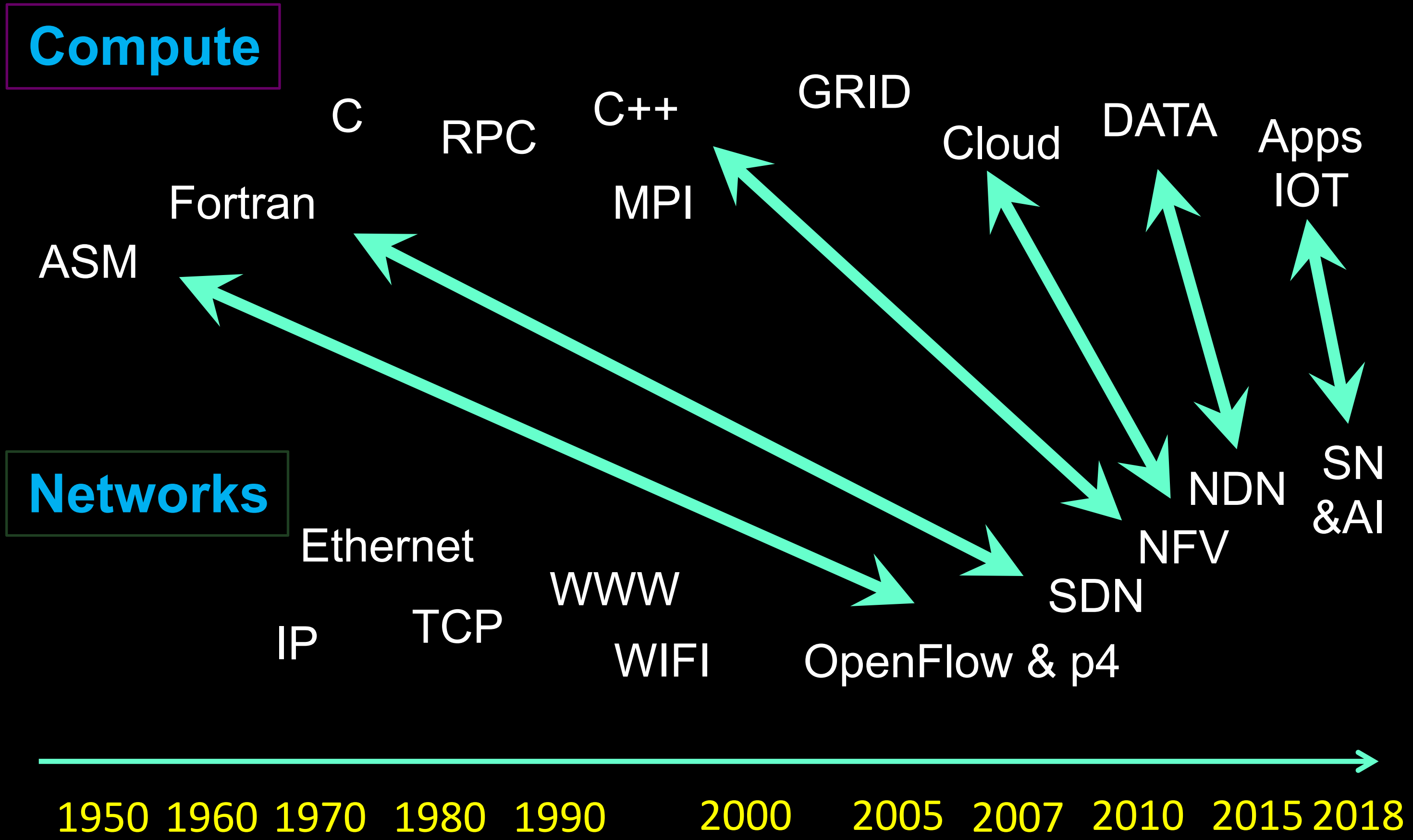
protocol LAN due to the easy comparison and convenience in the digital home. While consumer PC products has just started to migrate to a much higher bandwidth of 802.11n wireless LAN now working on next-generation standard definition is already in progress.

P4 & OpenFlow



Apps
Northbound API
OpenFlow Controller
OpenFlow Protocol
OpenFlow Agent
Driver
Programmable Data Plane ASIC
Program
Compile
Auto-Generated API
Target Binary

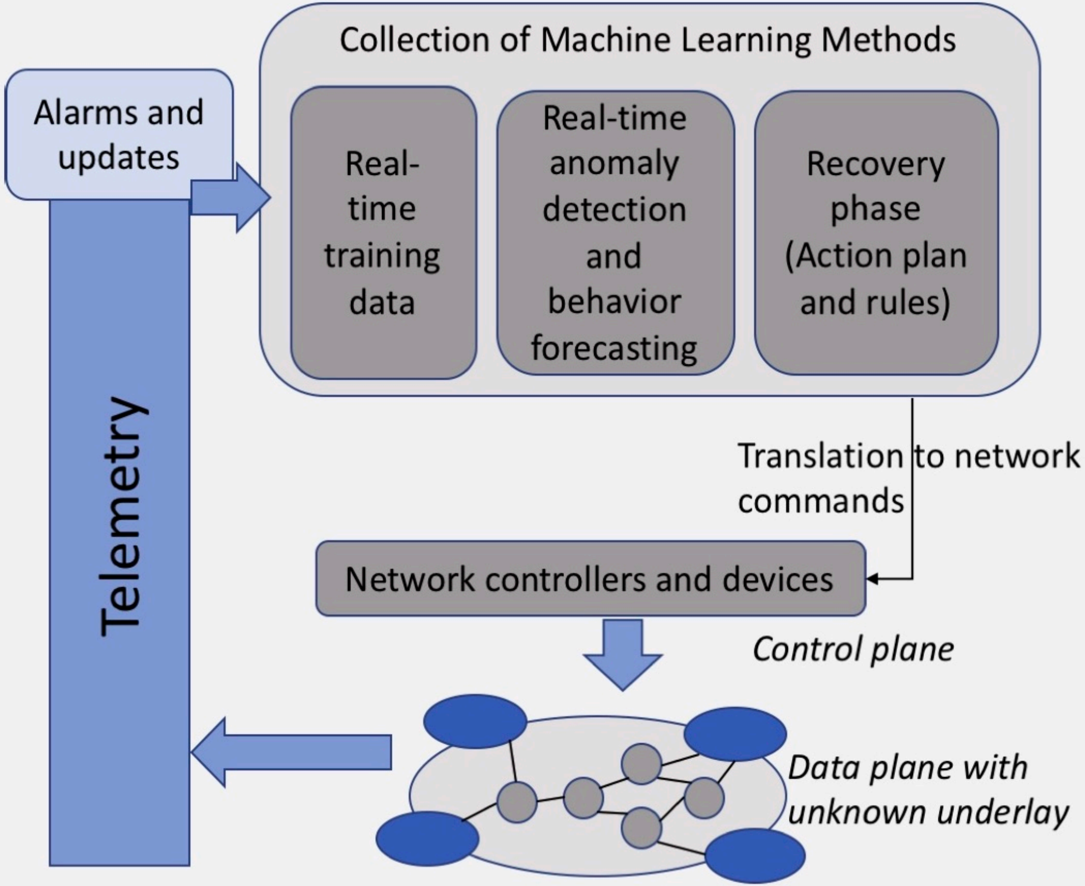
TimeLine



Compute
C
RPC
C++
GRID
Cloud
DATA
Apps
IOT
Fortran
MPI
ASM
Networks
Ethernet
WWW
IP
TCP
WIFI
OpenFlow & p4
SDN
NFV
NDN
SN
&AI
1950
1960
1970
1980
1990
2000
2005
2007
2010
2015
2018

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

Bundles

&

Fragmentation

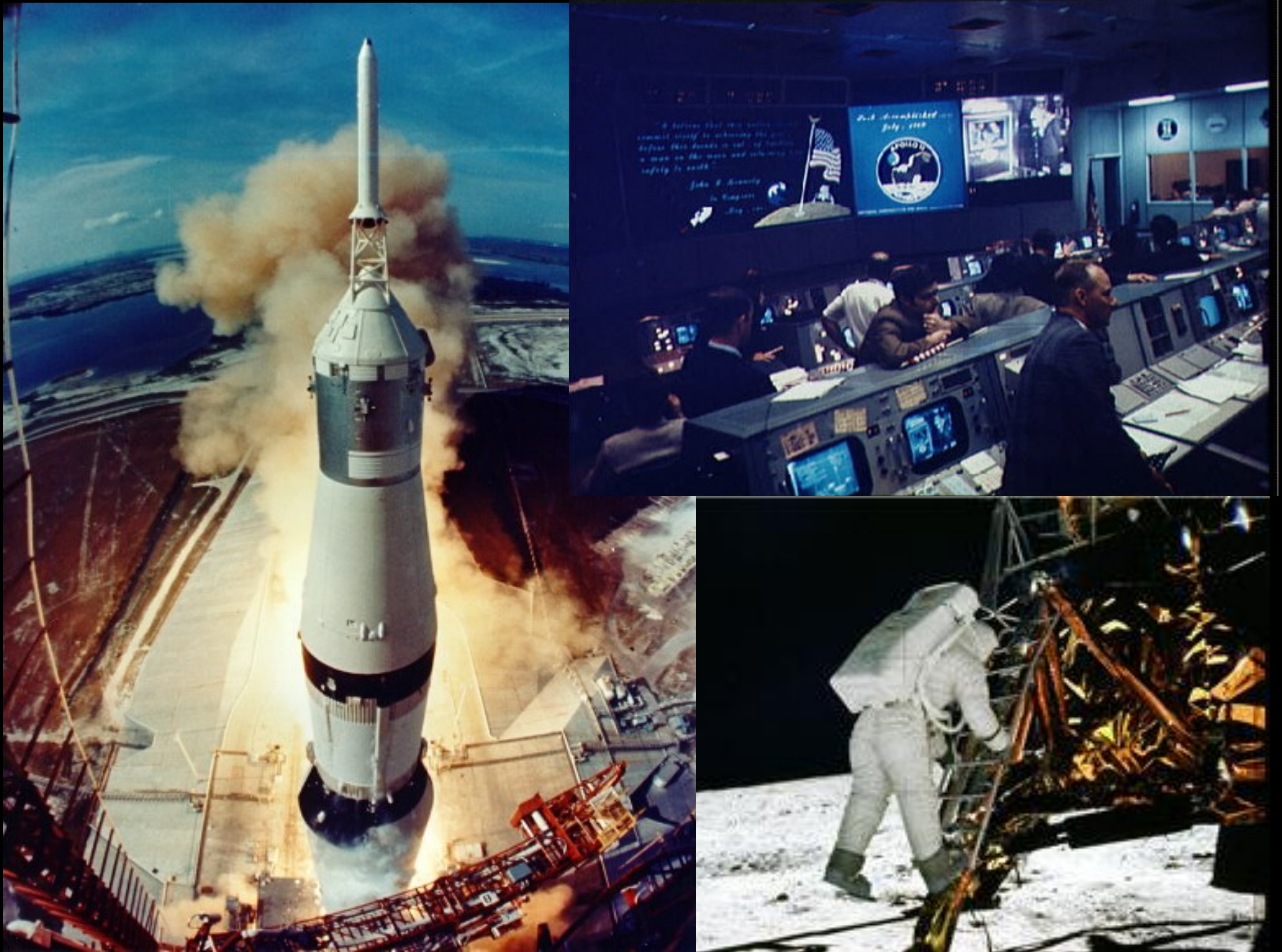
The Trend

- 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




IBM


TEMP
GIMBAL LOCK
PROG
RESTART
TRACKER
VERB
NOUN

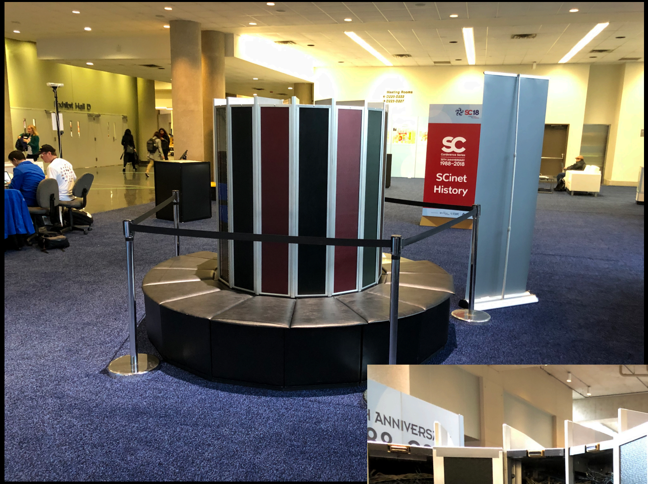

UPLINK ACTY
TEMP
AUTO
GIMBAL LOCK
PROG
RESTART
NO ATT
TRACKER
STBY
ALT
KEY REL
OPR ERR
COMP ACTY
PROG
VERB
NOUN
VERB
NOUN
+
7
8
9
CLR
ENTR
−
4
5
6
STBY
RSET
1
2
3
KEY REL


AT&T
3G
9:42 AM
Tuesday
9
SMS
Text
Calendar
Photos
Camera
YouTube
Stocks
280
Maps
73°
Weather
Clock
Calculator
Notes
Settings
iTunes
App Store
Phone
Mail
Safari
iPod
UPLINK ACTY
TEMP
GIMBAL LOCK
PROG
RESTART
TRACKER
NO ATT
STBY
KEY REL
OPR ERR
VERB
NOUN
CLR
ENTR
RSET

Some progress

2018

= ~7 ×



1976

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

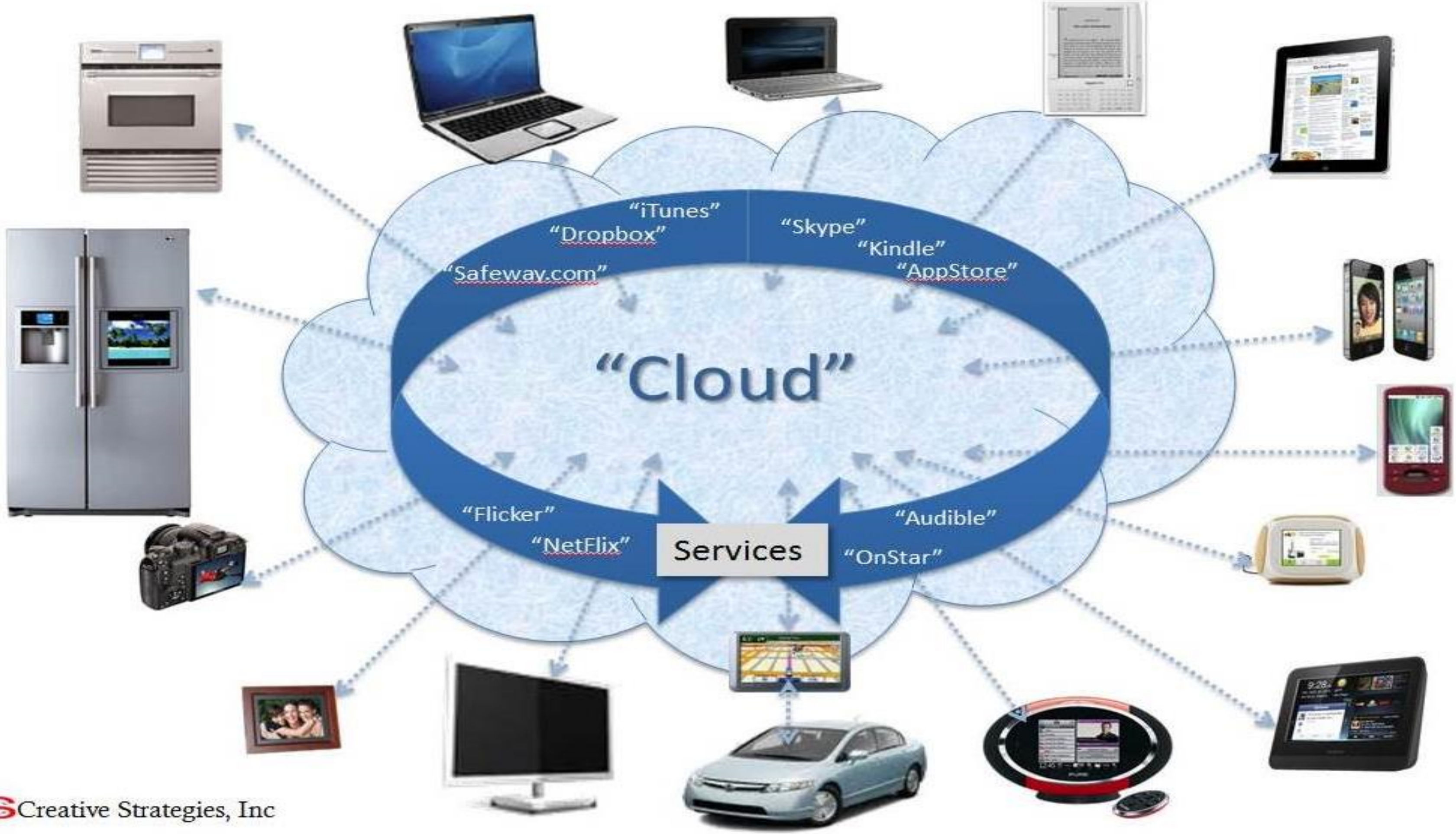
80 MHz
160 MFlops
8 MByte memory
300 MByte disks
120 kW


14

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

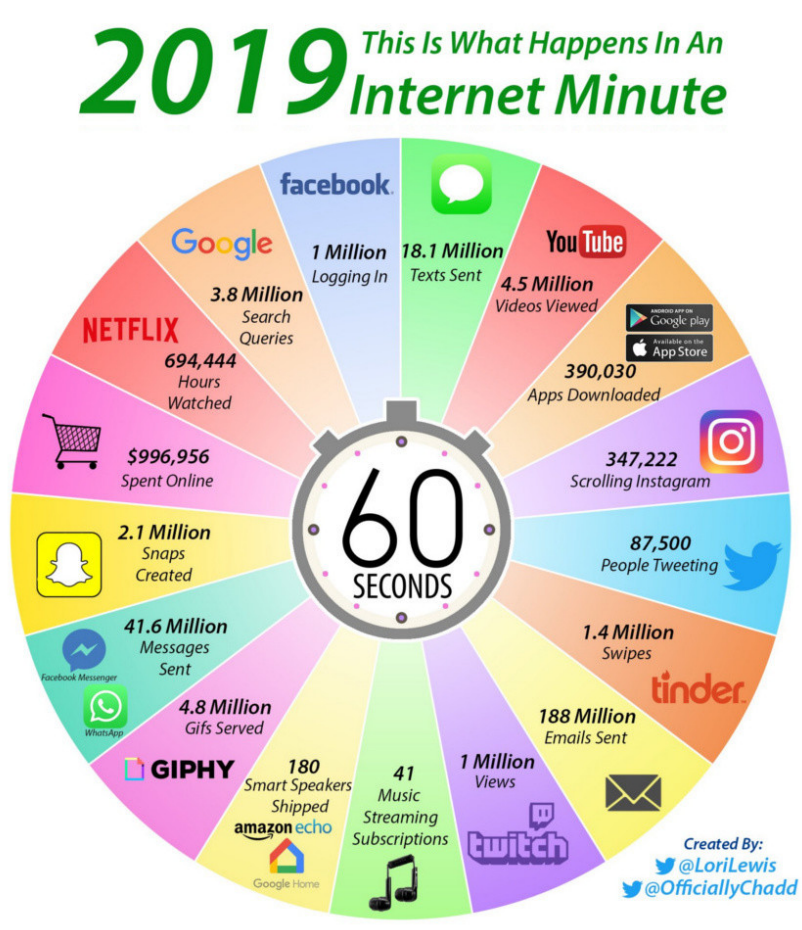


Creative Strategies, Inc

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?



- 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 PUBLISHING

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

➔

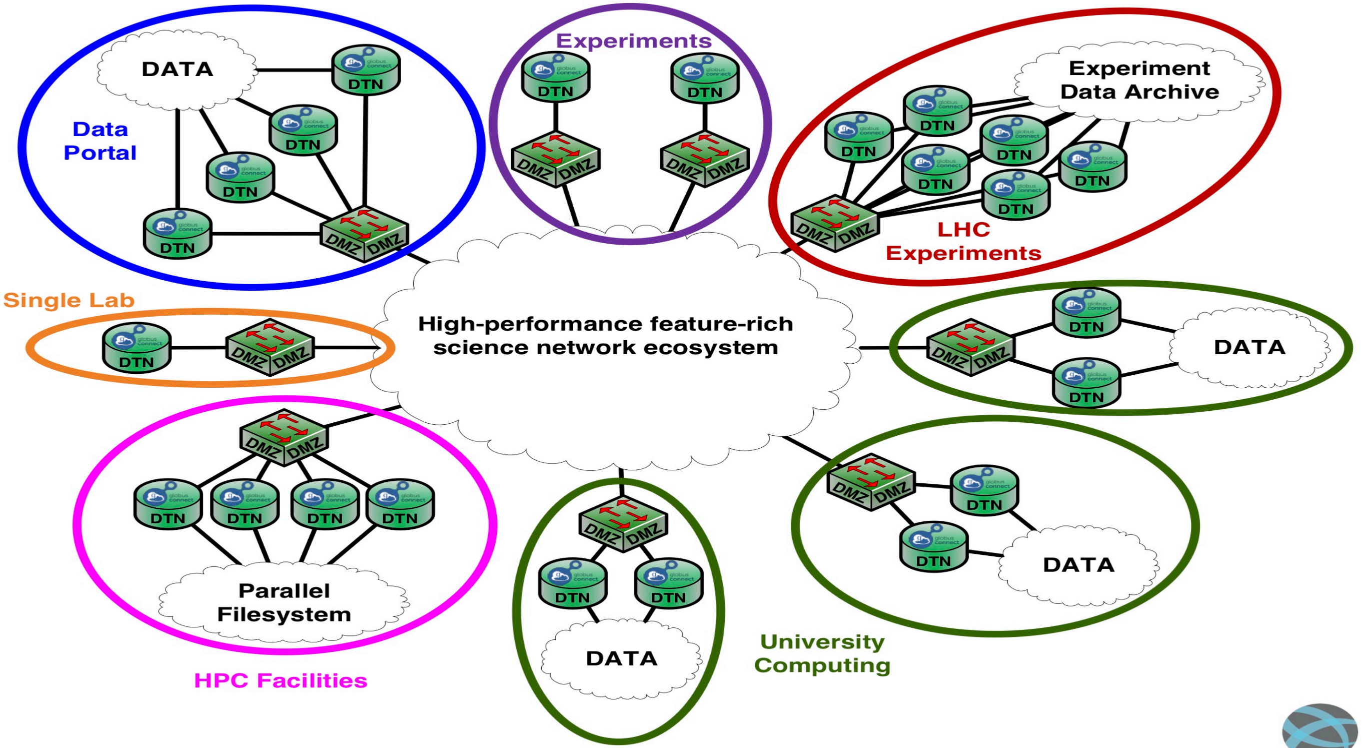
Create an equal playing field

➔

Sound Market principles

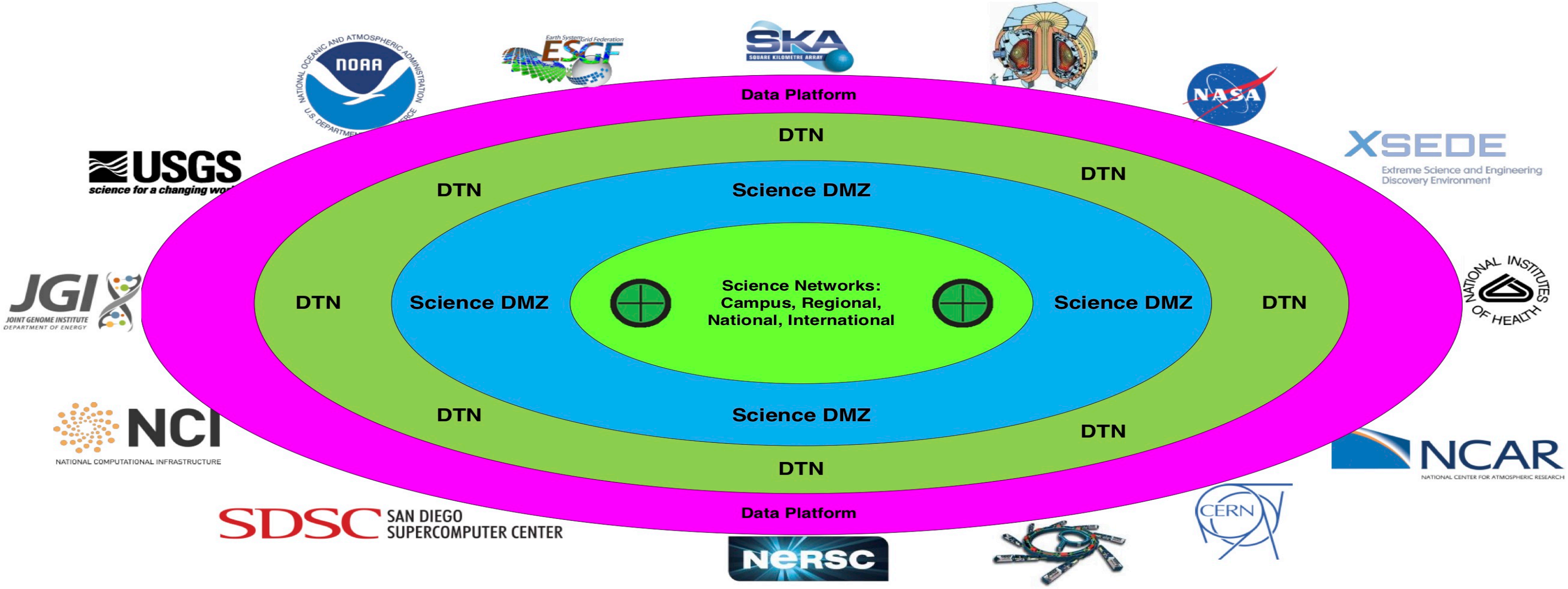
https://hbr.org/2017/09/managing-our-hub-economy

Science DMZs for Science Applications



Courtesy Eli Dart, ESnet

Data Ecosystem – Concentric View

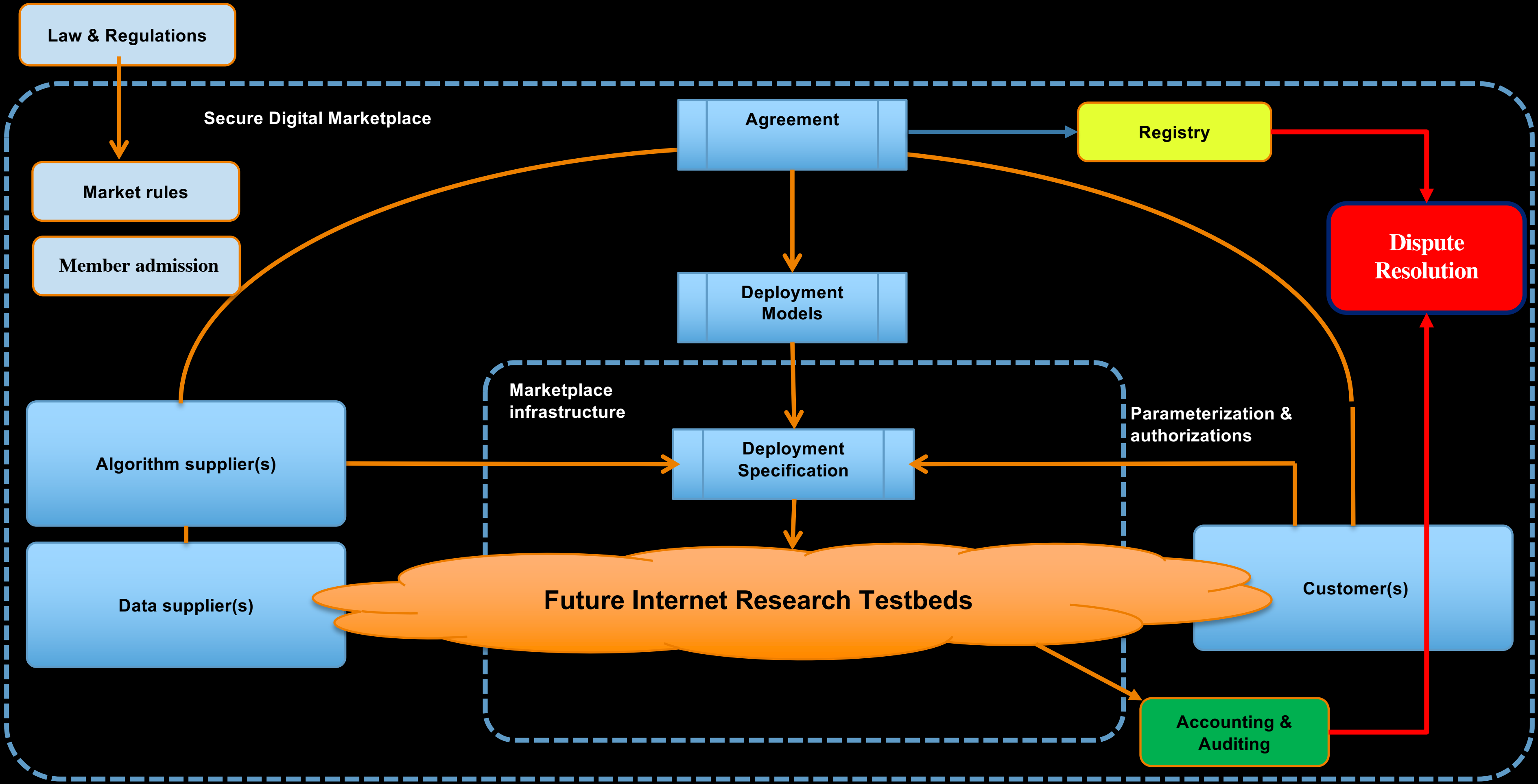


Courtesy Eli Dart, ESnet

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

National Scale

City / regional Scale

Campus / Enterprise Scale

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

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

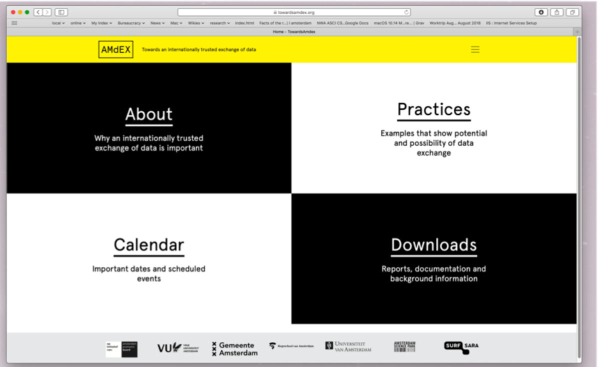
NLIP iShare project

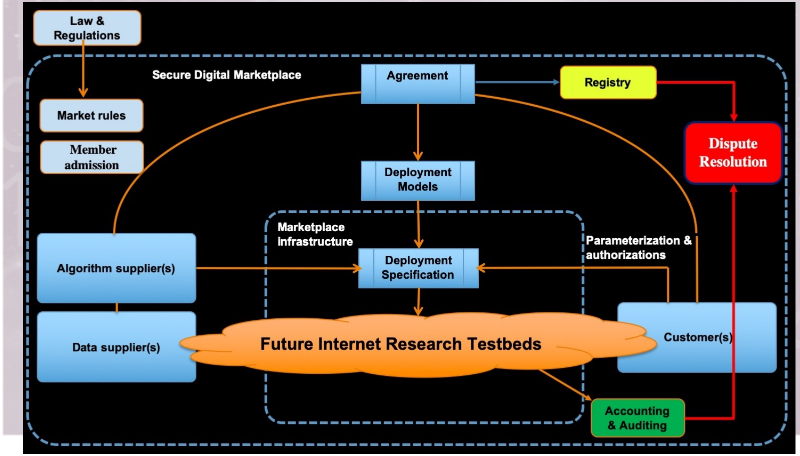
iSHARE
powered by NLIP

Cybersecurity Big Data
NWO COMMIT/
SARNET project
3.5 FTE

AMdEX.eu

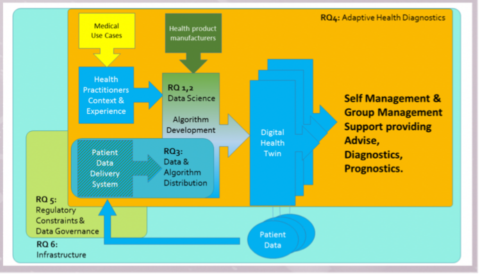
- Competing organisations, share data for common benefit
- Trust, Risk, data ownership & control
  - Industry: AF-KLM, Health, etc
  - Science: European Open Science Cloud
  - Society: Amsterdam Economic Board

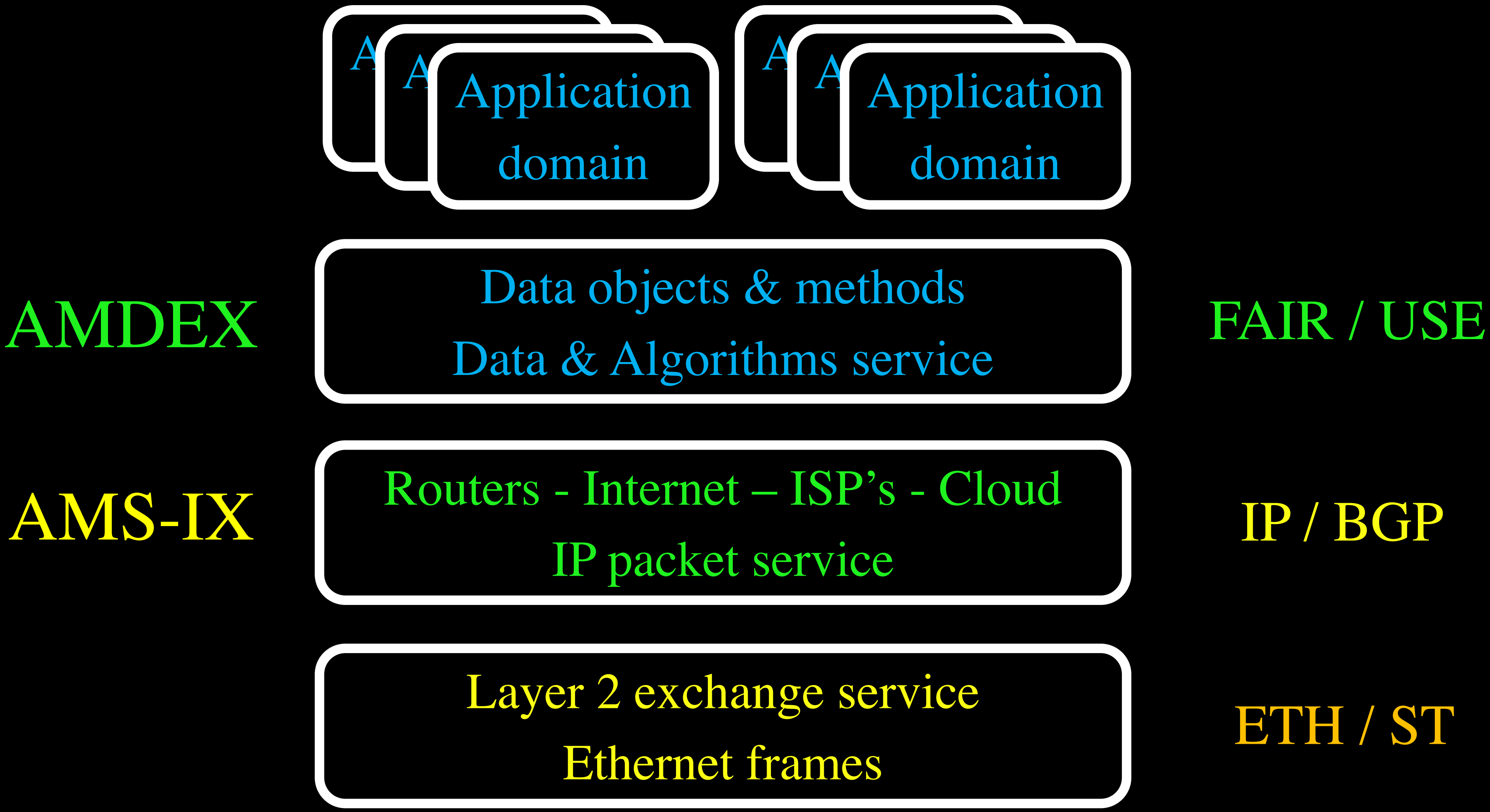




Aircraft Maintenance AF-KLM

Health: Enabling Personal Interventions

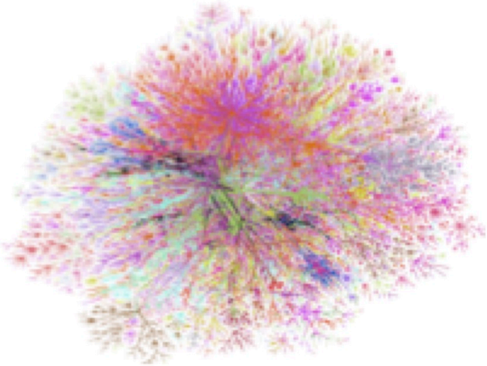



Application domain
Application domain
AMDEX
Data objects & methods
Data & Algorithms service
FAIR / USE
AMS-IX
Routers - Internet – ISP’s - Cloud
IP packet service
IP / BGP
Layer 2 exchange service
Ethernet frames
ETH / ST

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

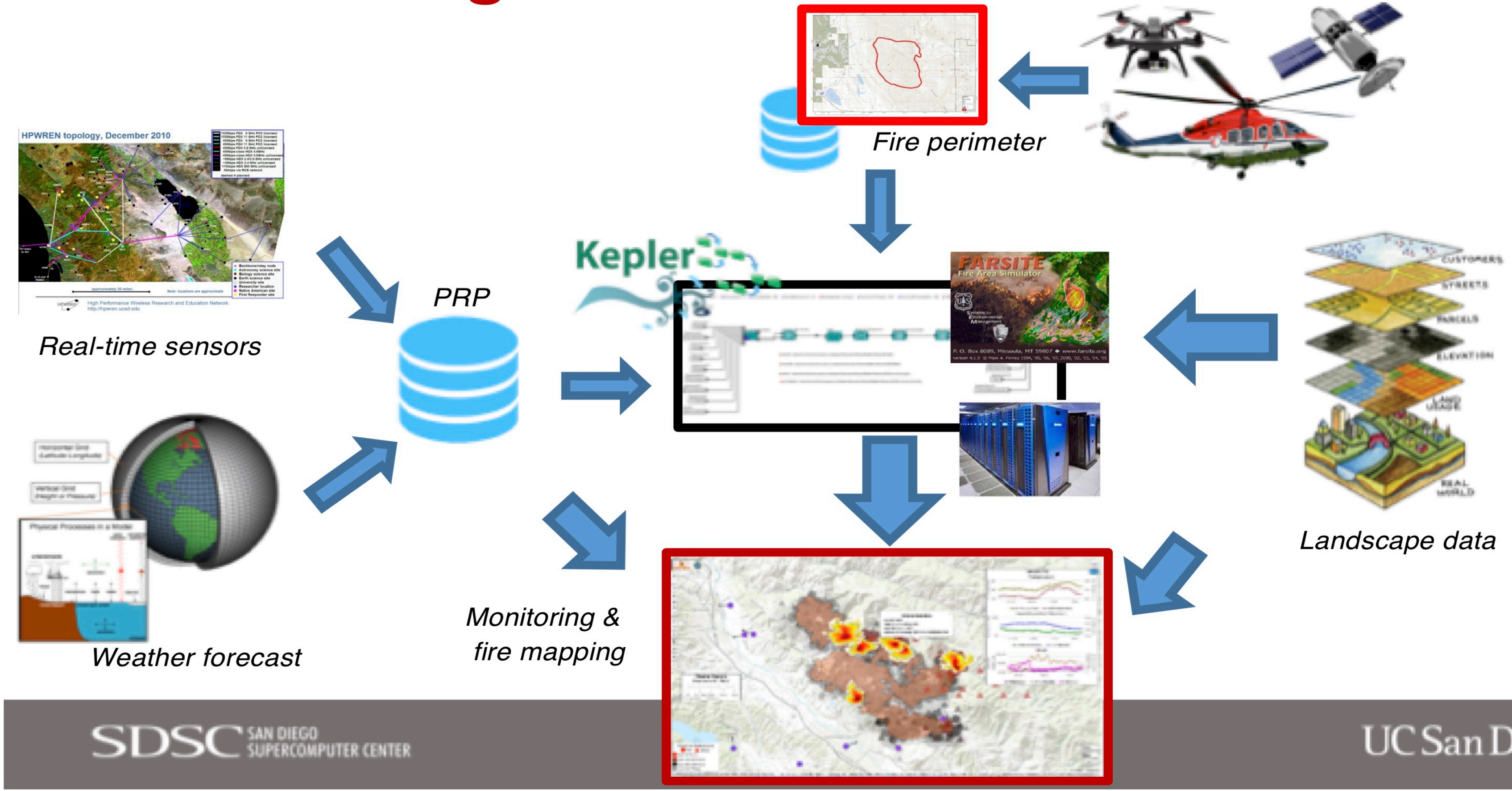
Enables dynamic data-driven applications

Computer-Aided Drug Discovery *Smart Cities* *Disaster Resilience and Response*

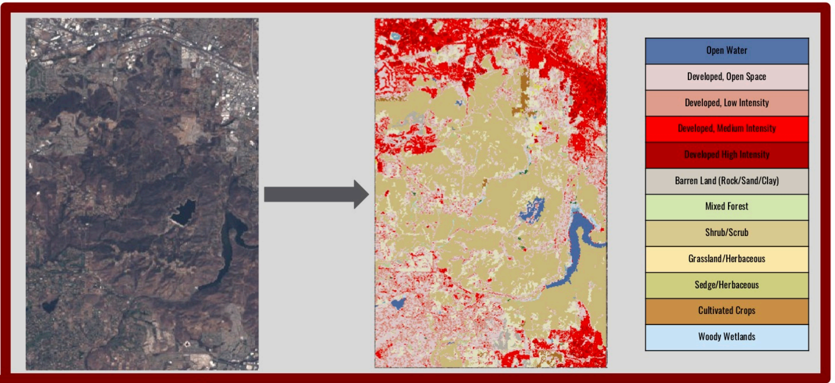
Smart Manufacturing *Personalized Precision Medicine* *Smart Grid and Energy Management*

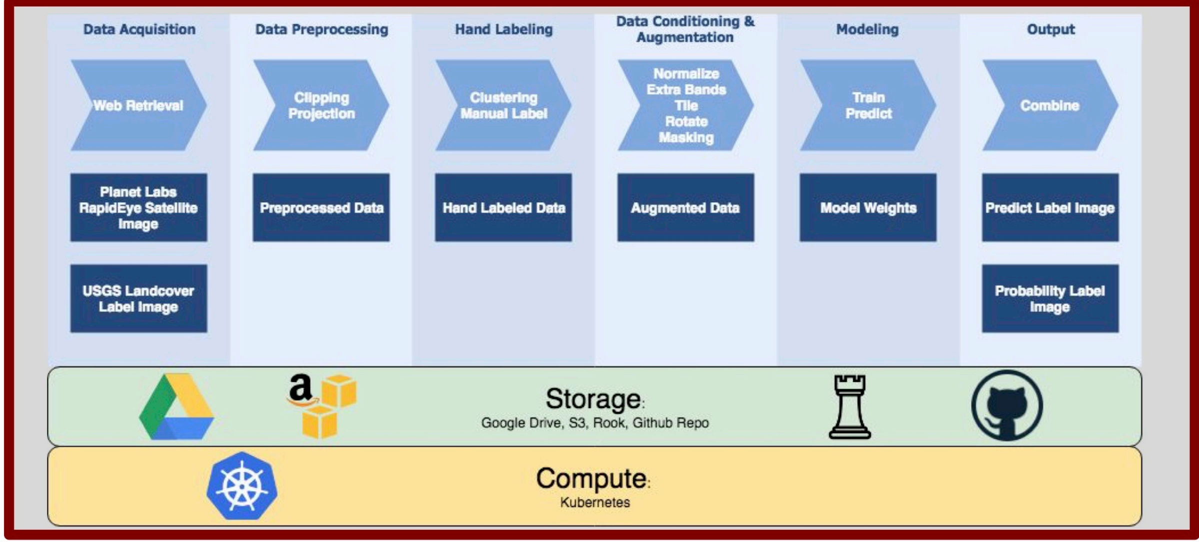
İlkay ALTINTAŞ, Ph.D.
ialtintas@ucsd.edu

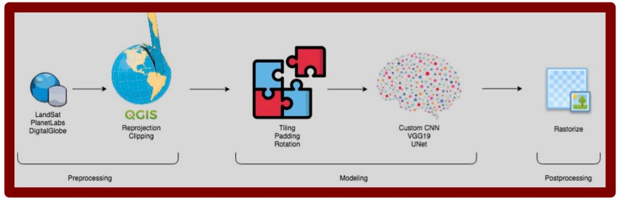
Fire Modeling Workflows in WIFIRE

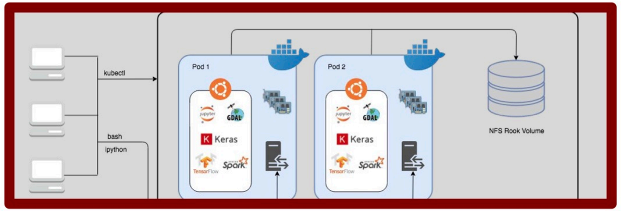


One Piece of the Puzzle: Vegetation Classification using Satellite Imagery





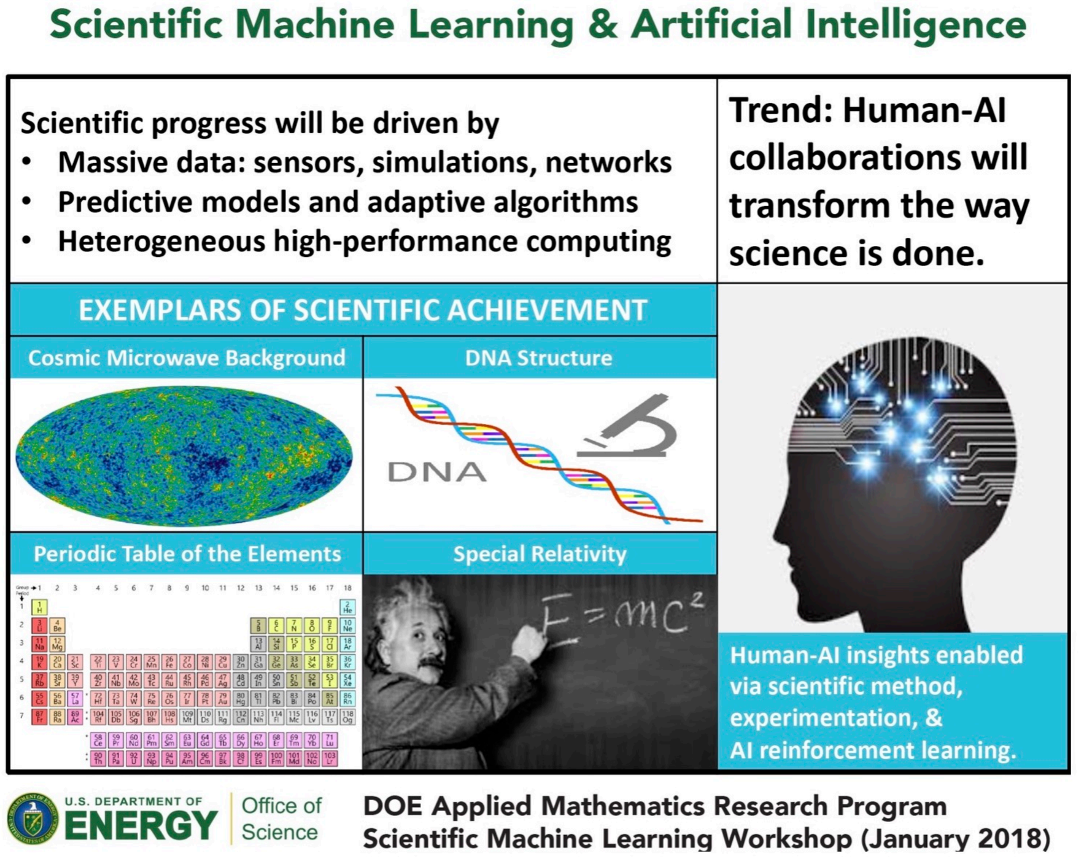




SDSC SAN DIEGO SUPERCOMPUTER CENTER

İlkay ALTINTAŞ, Ph.D.
ialtintas@ucsd.edu

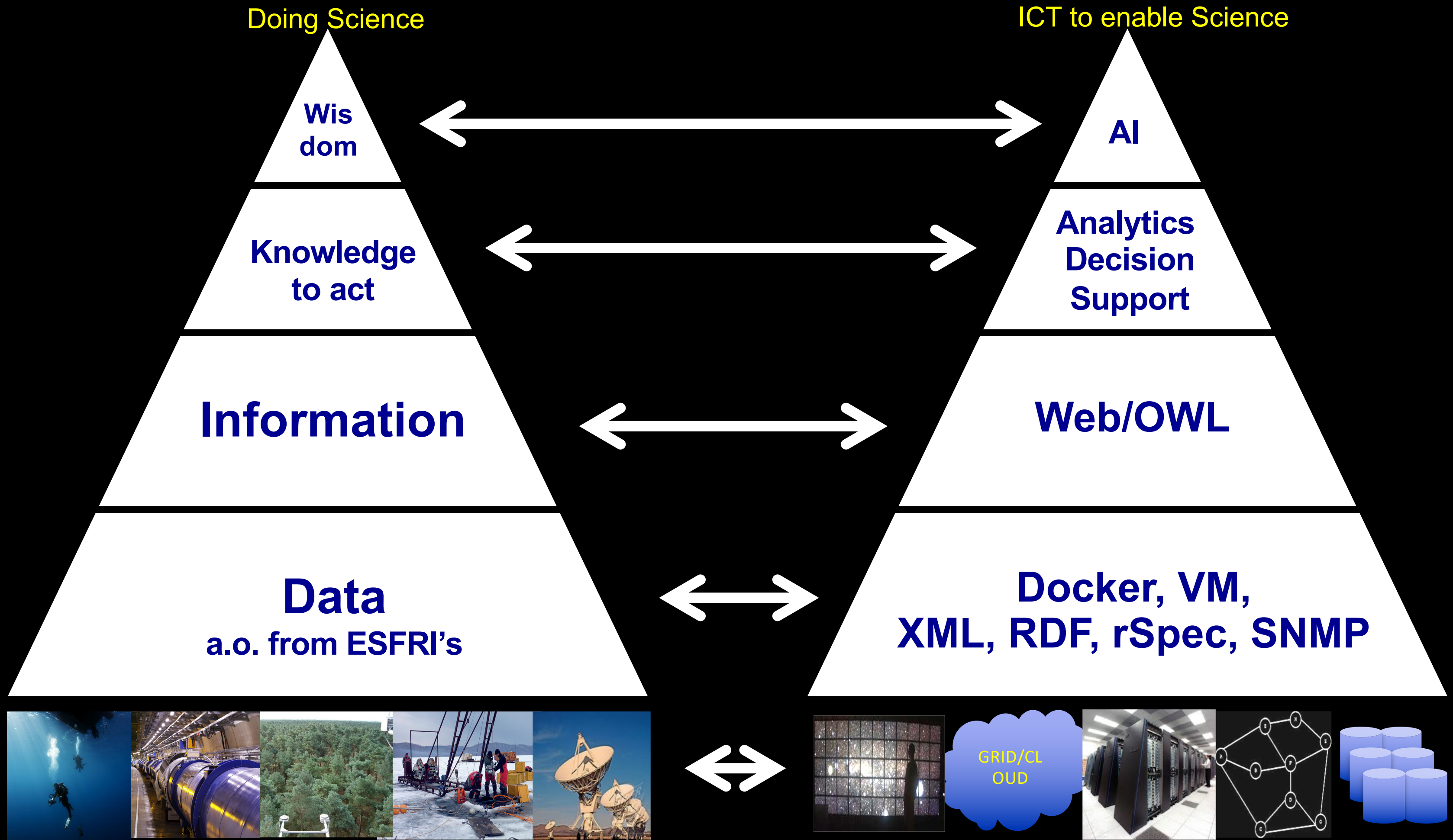
UC San Diego



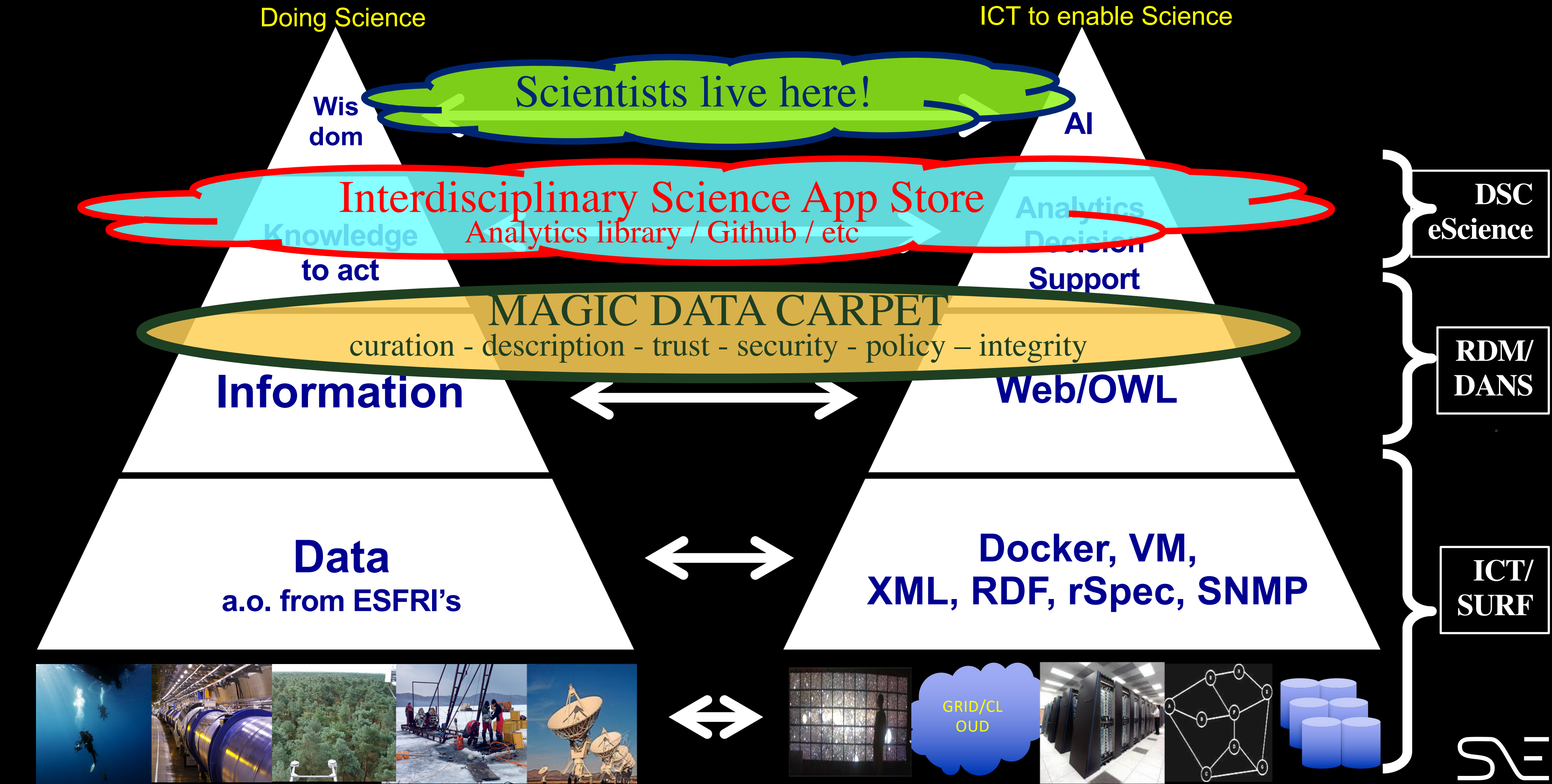
Workshop report:

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

The Big Data Challenge



The Big Data Challenge



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

Supported by: