HPC4AI e le istanze k8s multi-tenant come modello di provisioning del cloud GARR

Eventi GARR

WORKSHOP GARR 2019
8-10 ottobre 2019
Università Roma TRE

Marco Aldinucci
Computer Science Department, University of Torino

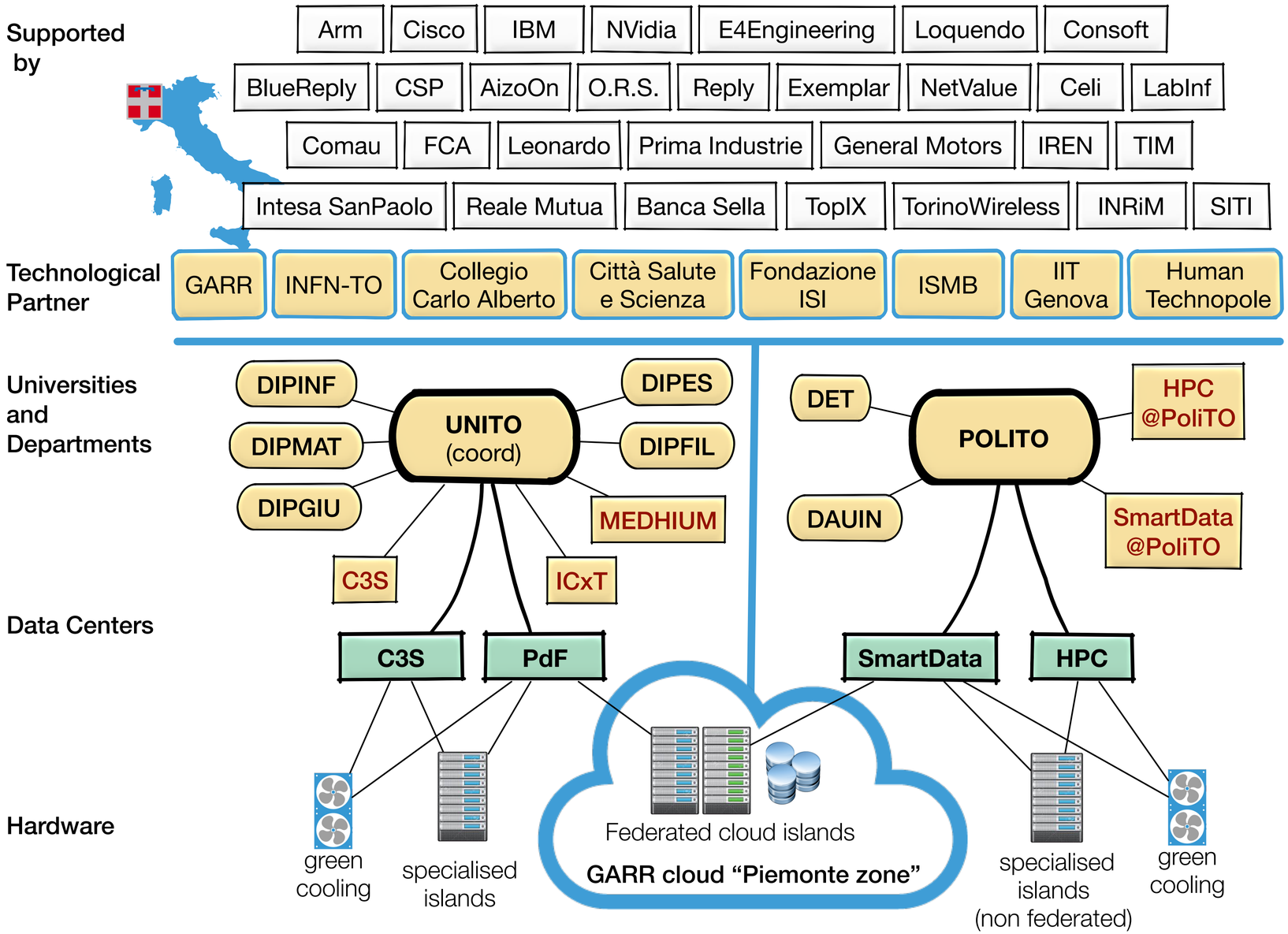
Head of **Parallel Computing group**, coordinator of HPC4AI,
national delegate at EuroHPC JU

Roma, October 10, 2019

UNIVERSITÀ DEGLI STUDI DI TORINO

HPC4AI

EuroHPC
http://EuroHPC.eu



HPC4AI

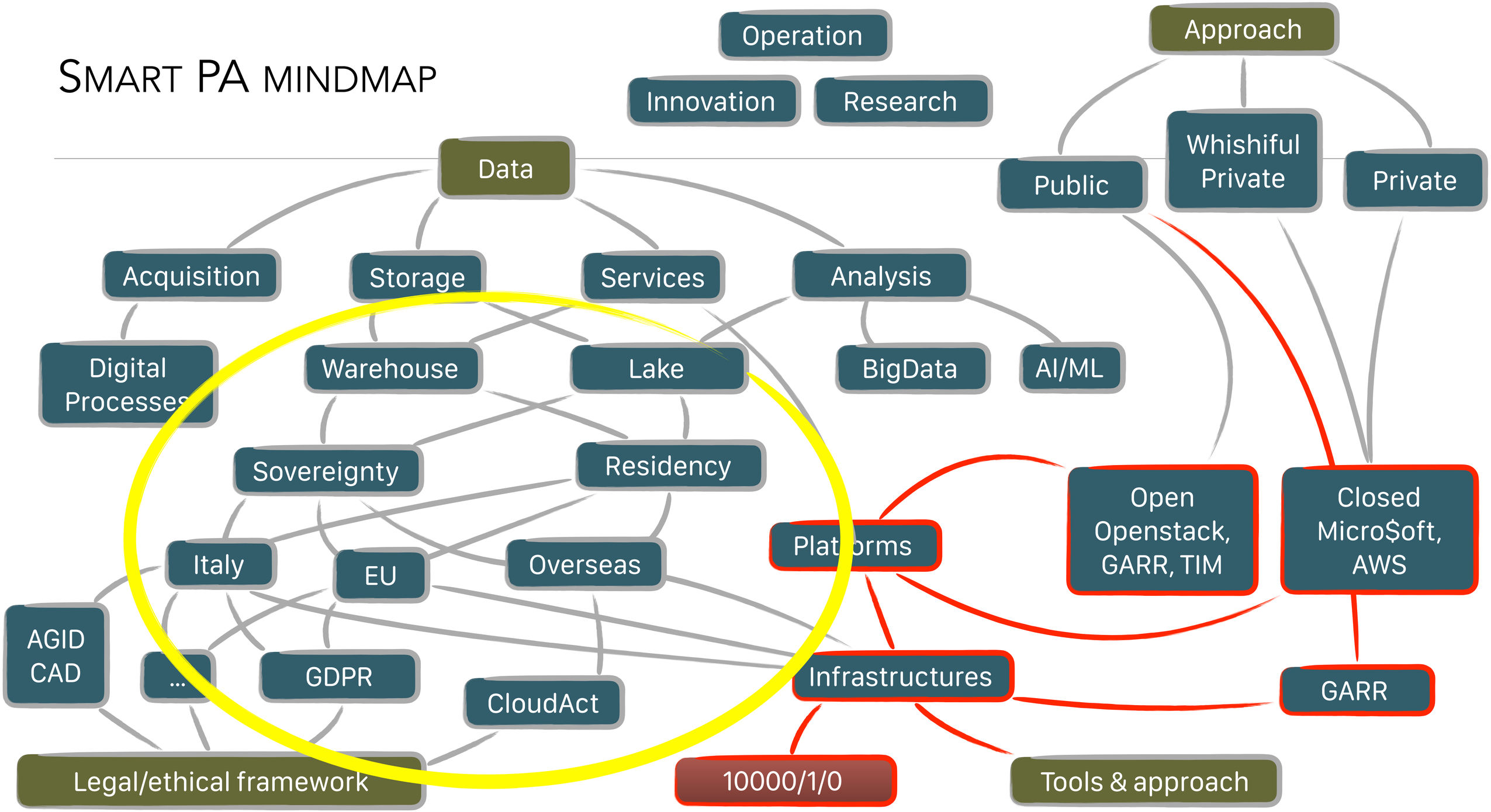
AI-in-demand platform

Facts

- INFRA-P call Nov. 2017
- Ranked 1st on ~30 submitted projects
- Kick-off mid apr 2018
- 4.5M€ funding
- 2 partners
- 8 associated partners
- Coord. M. Aldinucci
- Many industrial stakeholders

Users	Kind of service	Services	Artifacts
Domain experts with no skills on ML and BDA. Training set not required. Off-the-shelf algorithms/networks.	**Service-as-a-Service (SaaS)**	SaaS for ML and BDA designed within HPC4AI partners	Market place for ML and BDA services: Dashboards, trained models in several domains (NLP, Vision, …)
Domain experts skilled on ML and BDA. **Not expert in parallel computing.** New networks or pipelines; training set required.	**Platform-as-a-Service (PaaS)**	PaaS solutions for ML and BDA directly designed within HPC4AI or companion projects	Market place of VMs and Platforms realising software stacks for ML and BDA. Solutions for data ingestion, data lake, etc.
Researchers, cloud engineering, ML and BDA framework designers, cloud engineers, stack and automation designers.	**1) Infrastructure-as-a-Service (IaaS)** **2) Metal-as-a-Service (IaaS)**	1) GARR/other cloud able to support federation 2) Job scheduler for HPC resources	1) Openstack, docker, VM, object storage, file storage, kubernetes, etc. 2) Alternative cloud, job queue, Big Data Stack (Spark, …).
Researchers, run-time designers.	**Hardware**	Bare Metal	Multicore, GPU, storage, network, switch, UPS, cooling, etc.

SMART PA MINDMAP



Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction

Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools

What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item

CLOUD Act

From Wikipedia, the free encyclopedia

The **Clarifying Lawful Overseas Use of Data Act** or **CLOUD Act** (H.R. 4943) is a United States federal law enacted in 2018 by the passing of the Consolidated Appropriations Act, 2018, PL 115-141, section 105 executive agreements on access to data by foreign governments. Primarily the CLOUD Act amends the Stored Communications Act (SCA) of 1986 to allow federal law enforcement to compel U.S.-based technology companies via warrant or subpoena to provide requested data stored on servers regardless of whether the data are stored in the U.S. or on foreign soil.

Contents [hide]

Background [edit]

The CLOUD Act was introduced following difficulties that the Federal Bureau of Investigations (FBI) had with obtaining remote data through service providers through SCA warrants, as the SCA was written before cloud computing was a viable technology. The situation was highlighted from a 2013 drug trafficking

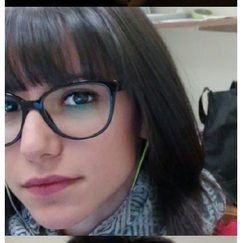
Clarifying Lawful Overseas Use of Data Act

Acronyms (colloquial)	CLOUD Act
Enacted by	the 115th United States Congress
Effective	March 23, 2018
Citations	
Public law	Pub.L. 115–141
Codification	
Acts amended	Stored Communications Act, Electronic Communications Privacy Act
Titles amended	18
U.S.C. sections amended	2523

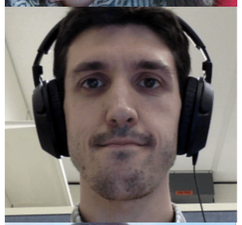
XXVI - G. Peretti Pezzi, Head of app support group, ETH Zurich, Switzerland

XXVII - F. Tordini, Research scientist Diatech Pharmacogene

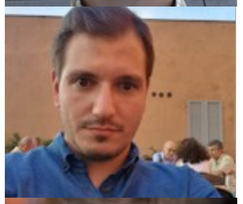
My past
PhD Students



XXVIII - C. Misale, Research scientist, IBM research, TJ Watson, NY, USA



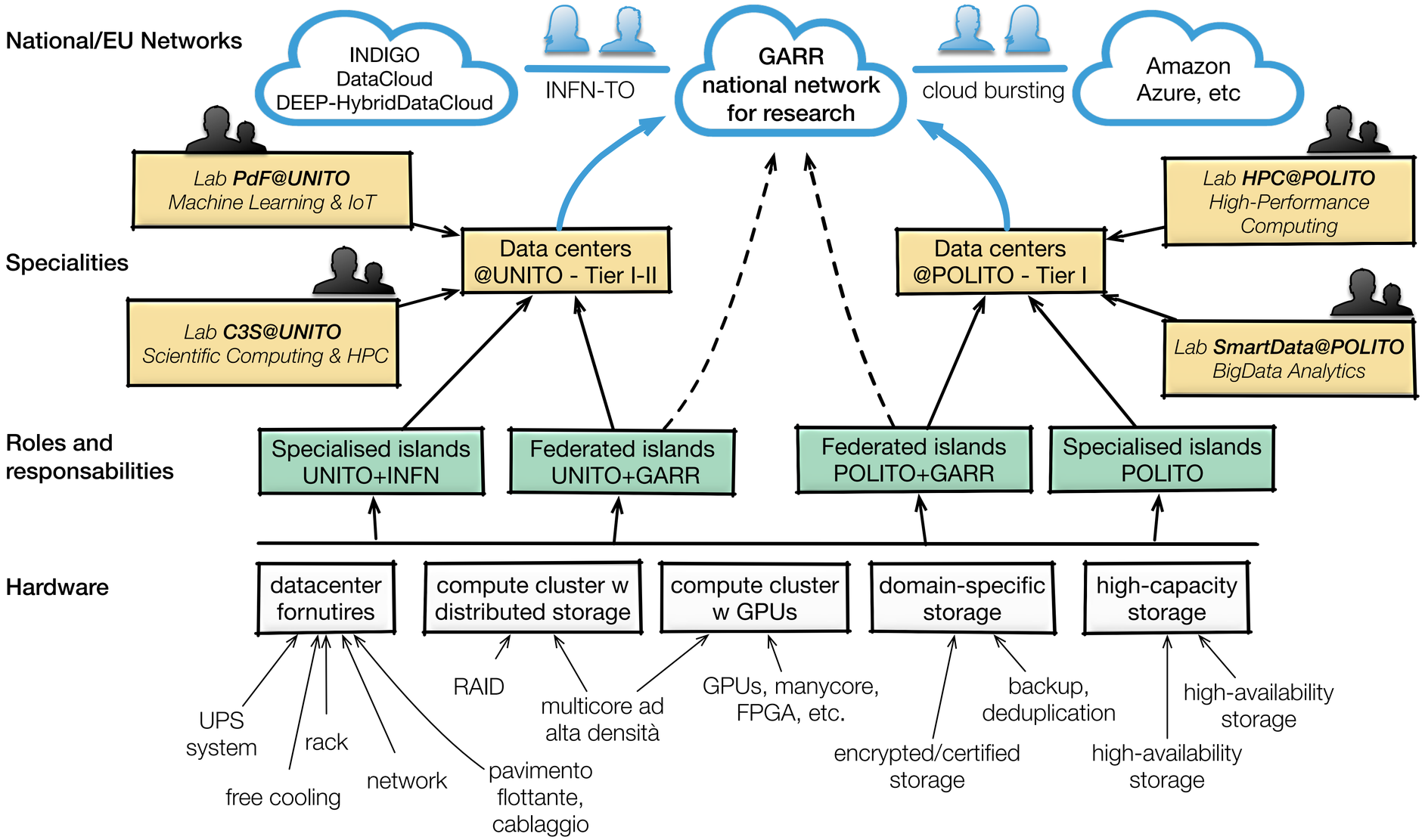
XXIX - M. Drocco, Research scientist, IBM research, TJ Watson, NY, USA



XXX - P. Viviani, Research scientist, Noesis Solution, Leuven, Belgium

XXXI - F. Spiga, Research scientist, Arm research, Cambridge, UK

FEDERATED ARCHITECTURE BASED ON THE GARR CLOUD (I.E. OPENSTACK VERSION MAINTAINED BY GARR+CANONICAL)



ailure

(

inistration of virtualized clusters
and provides an efficient, iSCSI

ICS 2007

A fully functional linux distribution that installs as a cloud

Much simpler to install w.r.t. openstack

Acquired by Eurotech then died

Marco Aldinucci
University of Pisa, Italy
aldinuc@di.unipi.it

Pierfrancesco Zuccato
Eurotech HPC, Italy
p.zuccato@exadron.com

INTERNATIONAL SUPERCOMPUTING CONFERENCE
26-29 June 2007, Dresden, Germany

VirtuaLinux

Virtual Clustering with no single point of failure

http://sourceforge.net/projects/virtualinux

Marco Aldinucci
University of Pisa, Italy
aldinuc@di.unipi.it

Pierfrancesco Zuccato
Eurotech HPC, Italy
p.zuccato@exadron.com

VirtuaLinux is a Linux meta-distribution that allows the creation, deployment and administration of virtualized clusters with no single point of failure. VirtuaLinux architecture supports diskless configurations and provides an efficient, iSCSI based abstraction of the SAN. Clusters running VirtuaLinux exhibit no master node, thus boosting resilience and flexibility.

Diskless Cluster + External SAN

- Using EVMS and iSCSI, the architecture provides a flexible, high-level description of the underlying hardware that frees the administrator from the traditional, rigid allocation of resources.
- Performance and scalability are achieved by means of the direct access of cluster blades to the external SAN via a switched Infiniband network (10-20Gb/s).
- Storage reliability is implemented within the external SAN, which can exploit a redundant array of disks (disks that can be mounted on the blades are usually quite slow and fragile).

Masterless Cluster Configuration

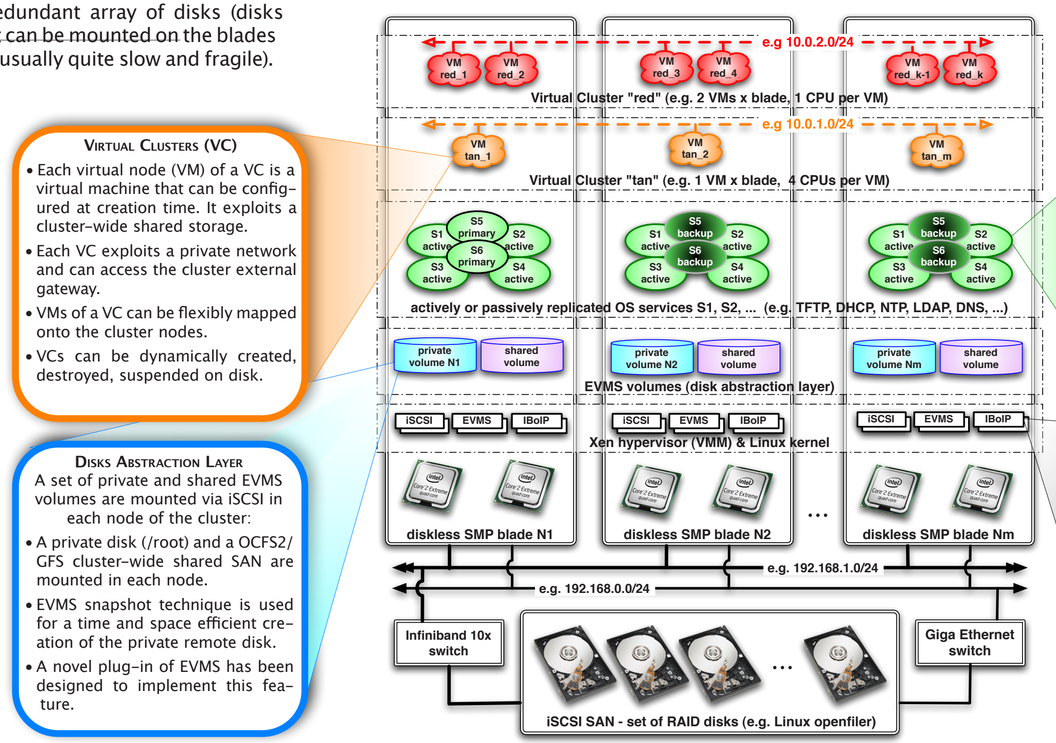
- VirtuaLinux has no master, all nodes have a symmetric configuration.
- Critical OS services are categorized and made redundant by either active or passive replication in such a way they are, at each point in time, cooperatively implemented by the running nodes.
- Any blade of the cluster can be hot-swapped with no impact on cluster operation.

Cluster Virtualization

VirtuaLinux separates three environments, targeted to different classes of administrators:

- The physical cluster, including the physical devices, which are insulated and made transparent (hardware technician).
- The privileged cluster, i.e. the environment that provides services and interfaces to the virtual clusters (skilled OS administrator).
- The virtual clusters that are sets of virtual machines. They can run any OS and configuration in such a way Grid and Beowulf style virtual clusters can coexist simultaneously (standard OS administrator).

VirtuaLinux Architecture



Virtual Clusters (VC)

- Each virtual node (VM) of a VC is a virtual machine that can be configured at creation time. It exploits a cluster-wide shared storage.
- Each VC exploits a private network and can access the cluster external gateway.
- VMs of a VC can be flexibly mapped onto the cluster nodes.
- VCs can be dynamically created, destroyed, suspended on disk.

Disks Abstraction Layer

A set of private and shared EVMS volumes are mounted via iSCSI in each node of the cluster:

- A private disk (/root) and a OCFS2/GFS cluster-wide shared SAN are mounted in each node.
- EVMS snapshot technique is used for a time and space efficient creation of the private remote disk.
- A novel plug-in of EVMS has been designed to implement this feature.

Linux OS services

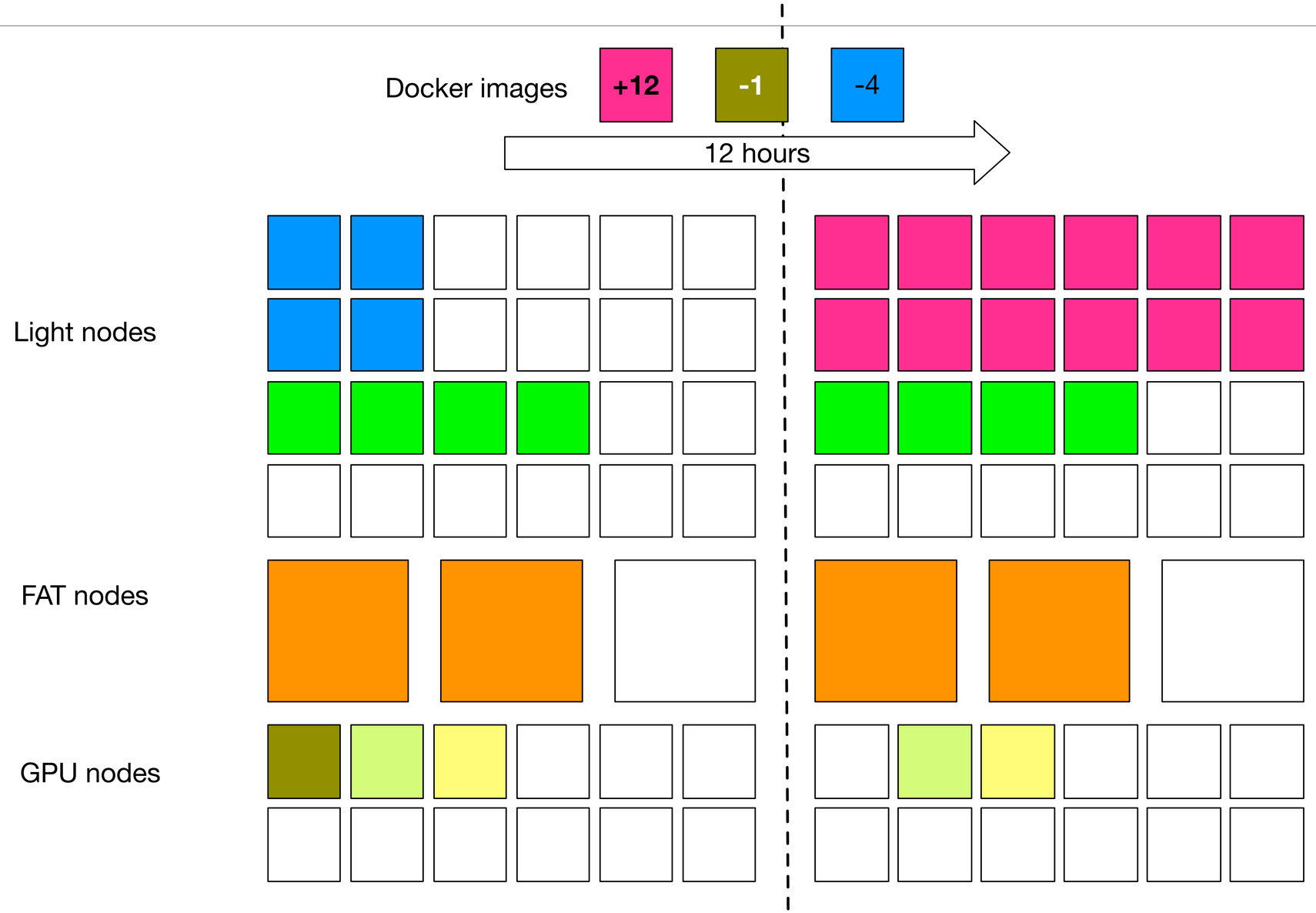
All standard Linux services are made fault-tolerant via either active or passive replication:

- Active: Services are started in all nodes; a suitable configuration enforces load balance on client requests. E.g. NTP, DNS, TFTP, DHCP.
- Passive (primary-backup): Linux HA with heartbeat is used as fault detector. E.g. LDAP, IP gateway.

Kernel Basic Features

- All standard Linux modules.
- Xen hypervisor, supporting Linux paravirtualization, and Microsoft Windows via QEMU binary translation (experimental).
- Network connectivity, including Infiniband userspace verbs and IP over Infiniband.
- iSCSI remote storage access.
- OCFS2 and GFS shared file systems.

Examples of a Specialised Island: GITLAB-based Docker Orchestrator for HPC



MOTIVATIONS

America Has a Monopoly Problem—and It's Huge

The Nobel Prize winner argues that an economy dominated by large corporations has failed the many and enriched the few.

By Joseph E. Stiglitz

OCTOBER 23, 2017

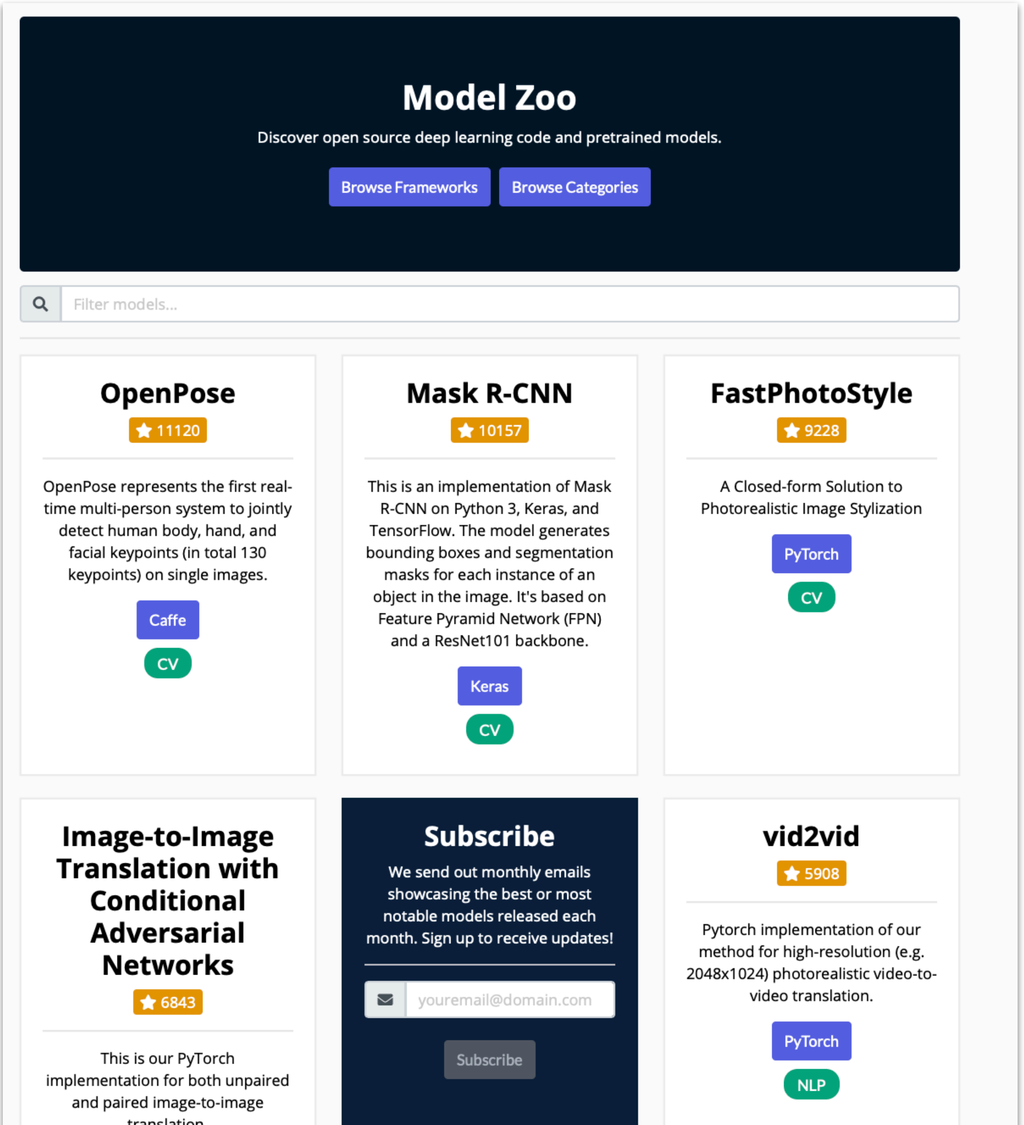
Motivations: AI/BDA services are typically operated by over-the-top

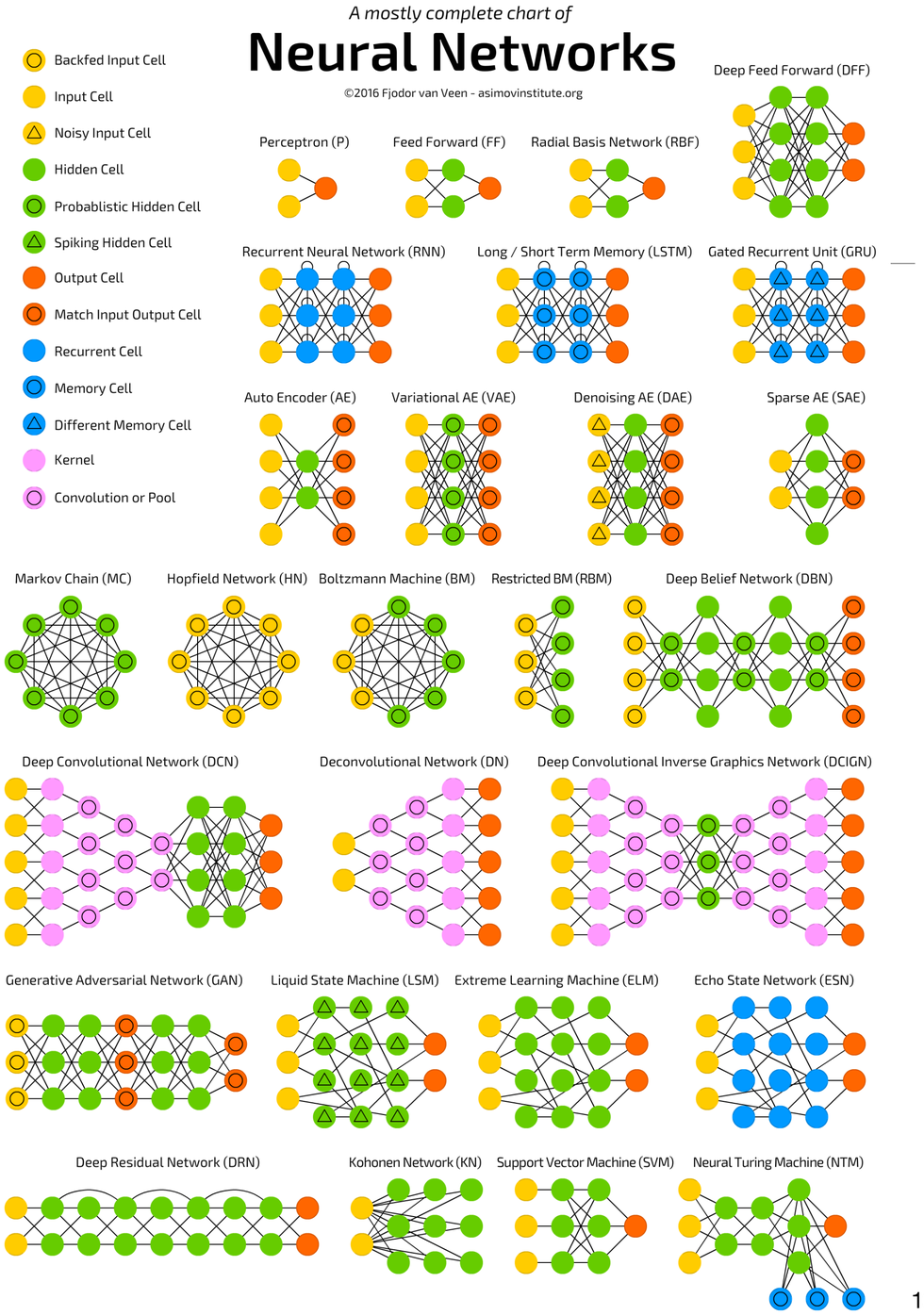
- Proprietary solutions produce technological lock-in. And continue to exact a rent
  - *The long-tail of this rental economy model increases inequalities rather than controlling them [J.E. Stiglitz]*
- AI require GPUs/TPUs. Their cost is very high in commercial public clouds
  - *Buying-vs-renting break-even point occurs early*
- AI and BDA need annotated datasets that require human effort for curation and maintenance
  - *Local availability of these dataset is an enabling feature for training/analysis*
- Datasets contain sensitive information, they should be stored appropriately
  - *Compliant with European regulations (including privacy, e.g. GDPR) if within EU borders [Argonne AoT]*

HPC4AI AIMS

- Facilitate scientific research and engineering in the areas of Artificial Intelligence and Big Data Analytics
  - Support large scale experimentation of applications
  - Engage regional industry in joint research projects, also boosting their R&D capabilities
  - Gather and store dataset with specific local/EU value (medical, business, code, …)
- Focusing on methods for the on-demand provisioning of AI and BDA cloud services

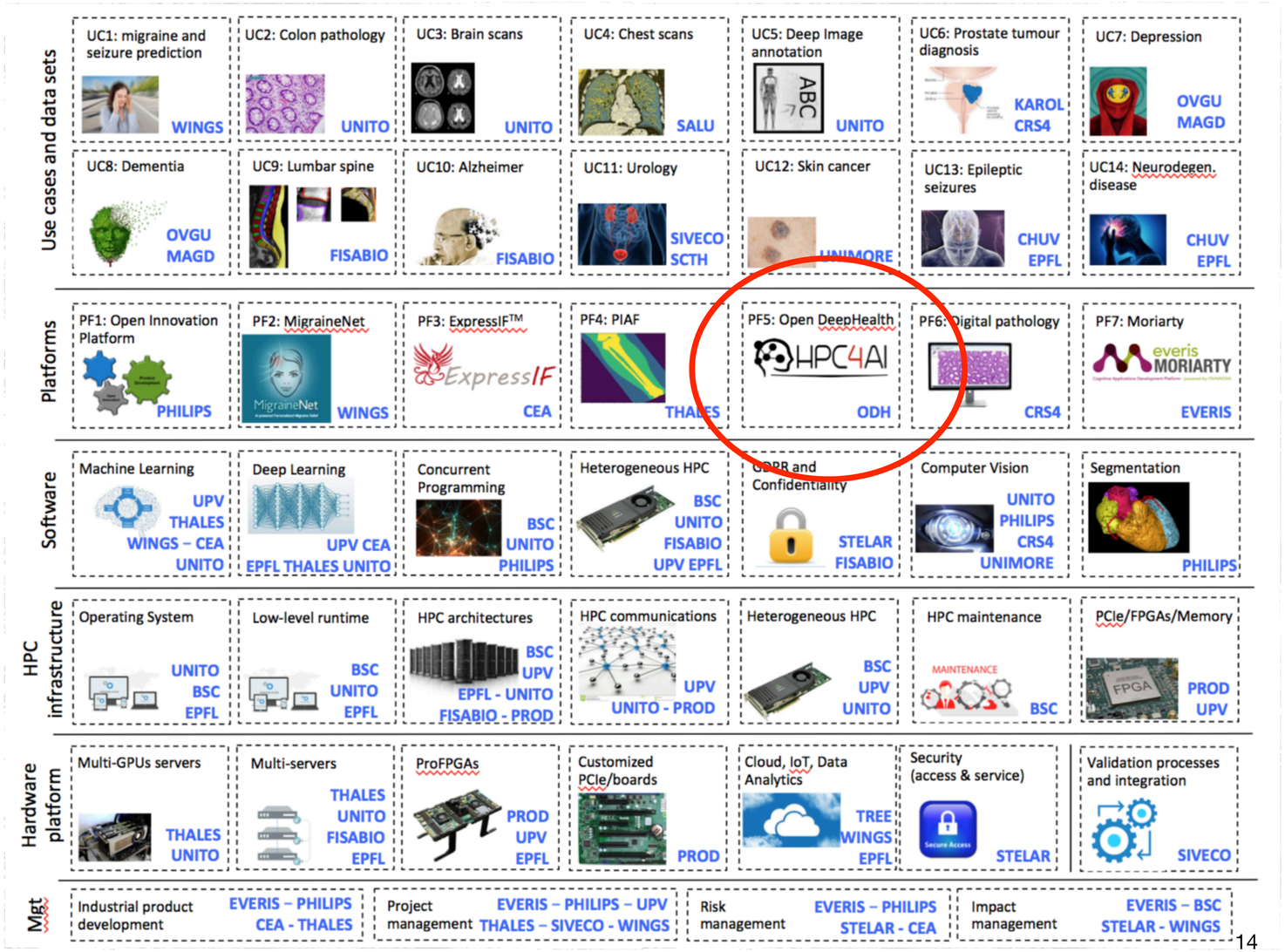
WHAT'S NEXT?

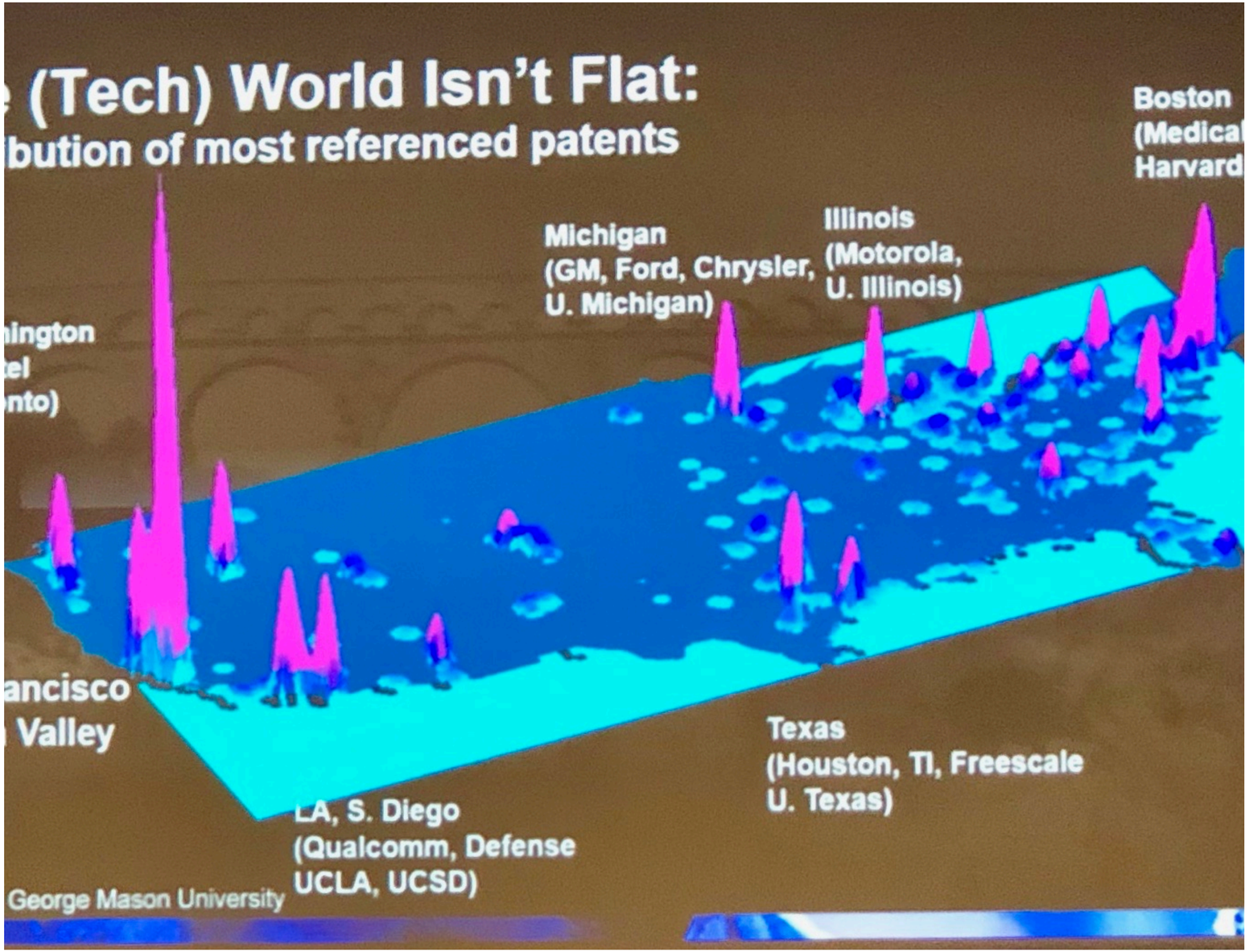

Model Zoo
Discover open source deep learning code and pretrained models.
Browse Frameworks
Browse Categories
Filter models...
OpenPose
11120
OpenPose represents the first real-time multi-person system to jointly detect human body, hand, and facial keypoints (in total 130 keypoints) on single images.
Caffe
CV
Mask R-CNN
10157
This is an implementation of Mask R-CNN on Python 3, Keras, and TensorFlow. The model generates bounding boxes and segmentation masks for each instance of an object in the image. It's based on Feature Pyramid Network (FPN) and a ResNet101 backbone.
Keras
CV
FastPhotoStyle
9228
A Closed-form Solution to Photorealistic Image Stylization
PyTorch
CV
Image-to-Image Translation with Conditional Adversarial Networks
6843
This is our PyTorch implementation for both unpaired and paired image-to-image
Subscribe
We send out monthly emails showcasing the best or most notable models released each month. Sign up to receive updates!
youremail@domain.com
Subscribe
vid2vid
5908
Pytorch implementation of our method for high-resolution (e.g. 2048x1024) photorealistic video-to-video translation.
PyTorch
NLP

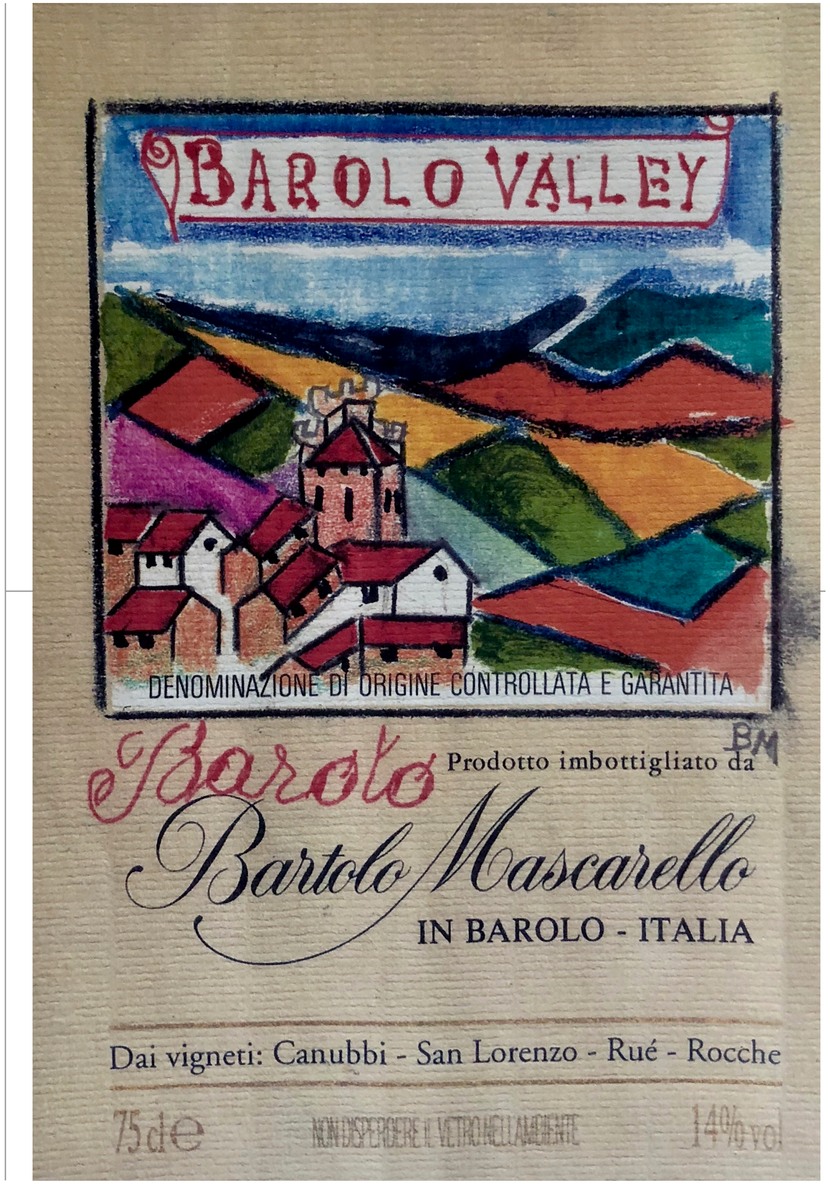

A mostly complete chart of
Neural Networks
©2016 Fjodor van Veen - asimovinstitute.org
Backfed Input Cell
Input Cell
Noisy Input Cell
Hidden Cell
Probablistic Hidden Cell
Spiking Hidden Cell
Output Cell
Match Input Output Cell
Recurrent Cell
Memory Cell
Different Memory Cell
Kernel
Convolution or Pool
Perceptron (P)
Feed Forward (FF)
Radial Basis Network (RBF)
Deep Feed Forward (DFF)
Recurrent Neural Network (RNN)
Long / Short Term Memory (LSTM)
Gated Recurrent Unit (GRU)
Auto Encoder (AE)
Variational AE (VAE)
Denoising AE (DAE)
Sparse AE (SAE)
Markov Chain (MC)
Hopfield Network (HN)
Boltzmann Machine (BM)
Restricted BM (RBM)
Deep Belief Network (DBN)
Deep Convolutional Network (DCN)
Deconvolutional Network (DN)
Deep Convolutional Inverse Graphics Network (DCIGN)
Generative Adversarial Network (GAN)
Liquid State Machine (LSM)
Extreme Learning Machine (ELM)
Echo State Network (ESN)
Deep Residual Network (DRN)
Kohonen Network (KN)
Support Vector Machine (SVM)
Neural Turing Machine (NTM)

Follow-up project (2018_ICT-11-a EU IA, 13M€)
DeepHealth: Deep-Learning and HPC to Boost Biomedical Applications for Health

- 18 Parterns: Everis, Siveco, Wings, Philiphs, SIVECO, IBM, Thales, CEA, Treelogic, EPFL, UPV, UNITO, UNIMORE, …
- Design and develop:
  - European Library for Distributed Deep Learning for health
  - AI-on-demand cloud platforms (HPC4AI, …)







FUTURE HORIZONS

PIEDMONT AI ECOSYSTEM

A high-level regional system, from research to industry

Artificial Intelligence and Intelligent Systems — CINI National Lab

- Top region in Italy for R&D investment (2,03% of GDP above national average,1.6% private).

TRL3

- Three prestigious universities, three industrial research institutes.

TRL5

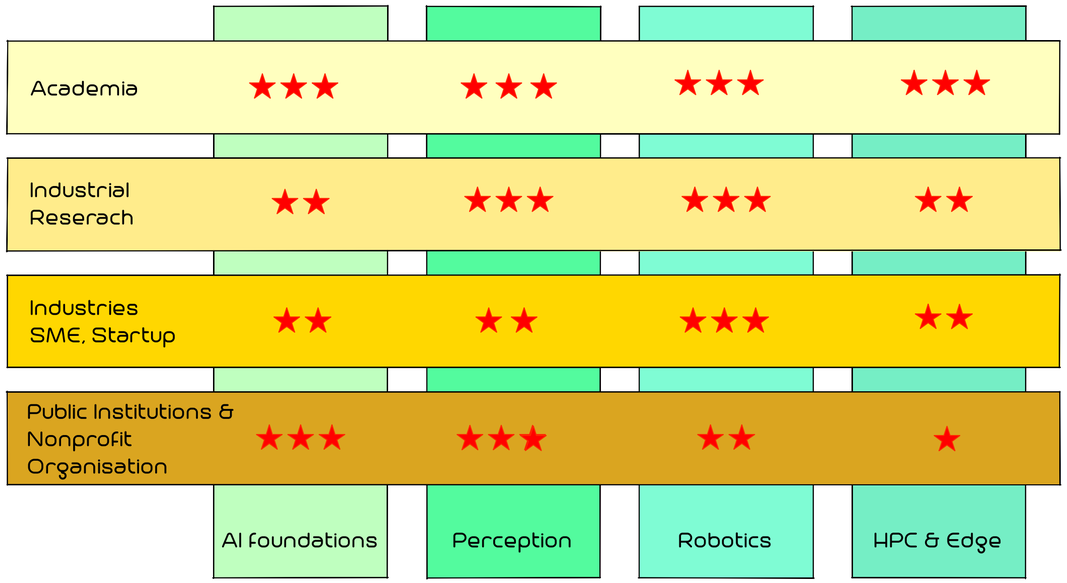
- International NGO, Government Office, innovation clusters, competence center.

TRL7

TRL9

- Vibrant deep tech Industry (Large enterprise, SME, startup, incubator) and research center of multinationals across main areas: health, robotics, space, automotive, telco, finance, agri-food.

We propose a single and coordinated Center-of-Excellence (CoE) of tech players ready to cooperate with pan-European AI laboratories



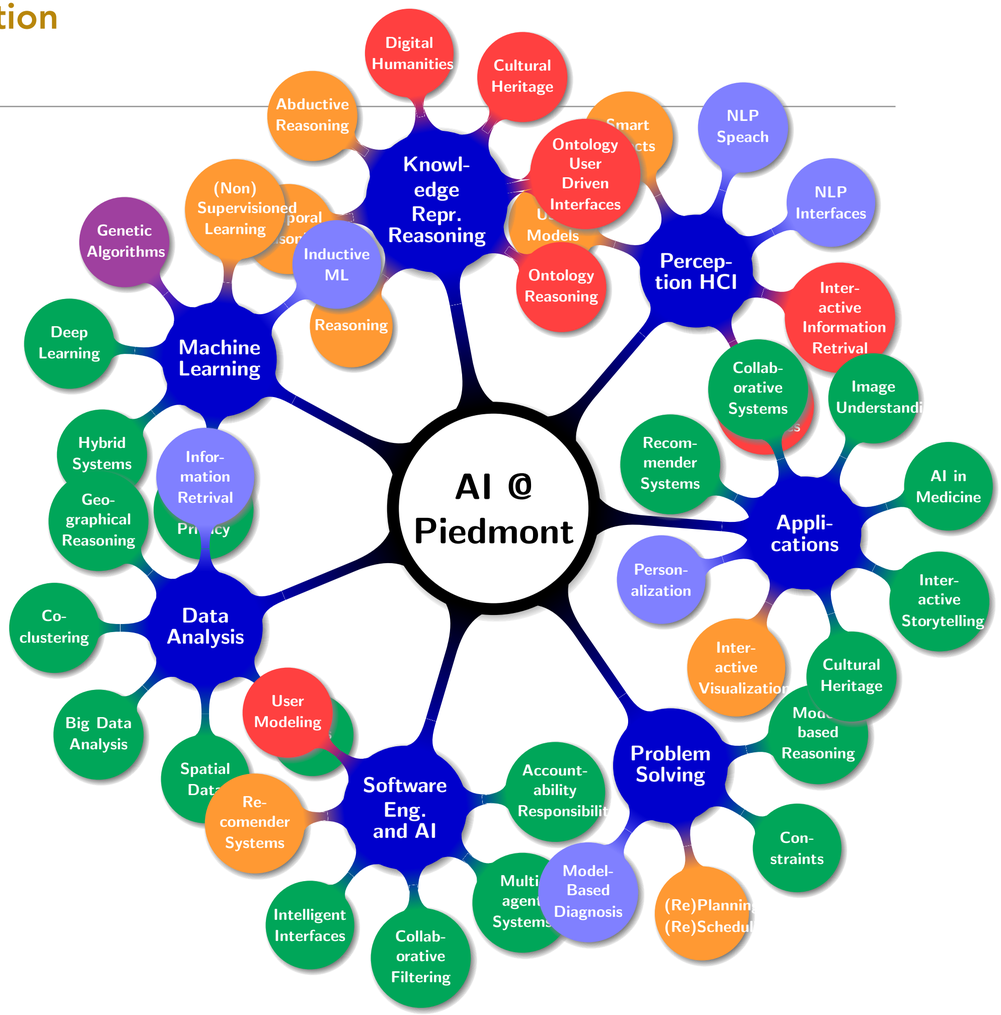
Contact: Marco Aldinucci

Tech transfer knowledge Track record 2018

First region in Italy for projects financed by European Territorial Cooperation

- **DeepHealth** (EU ICT-11-b 15M€)
  - Deep-Learning and HPC to Boost Biomedical Applications for Health
- **Lexis** (EU ICT-11-b 14M€)
  - Large-scale EXecution for Industry & Society
- **HPC4AI** (EU FESR 2018 - 4.5M€)
  - High-Performance Computing for Artificial Intelligence (AI-on-Demand platform)
- **CANP**: (EU FESR 2018 -11M€)
  - AI for Elderly Care and Telemedicine
- **Industry 4.0 competence center** (IT MISE 2018 - 20M€)
  - POLITO+UNITO+24 industries (FCA, GM, Thales-Alenia Space, Leonardo, STMicroelectronics, Prima Industria, TIM, …)

Competencies



nei pazienti reclutati in 5 studi conservati in TCGA.

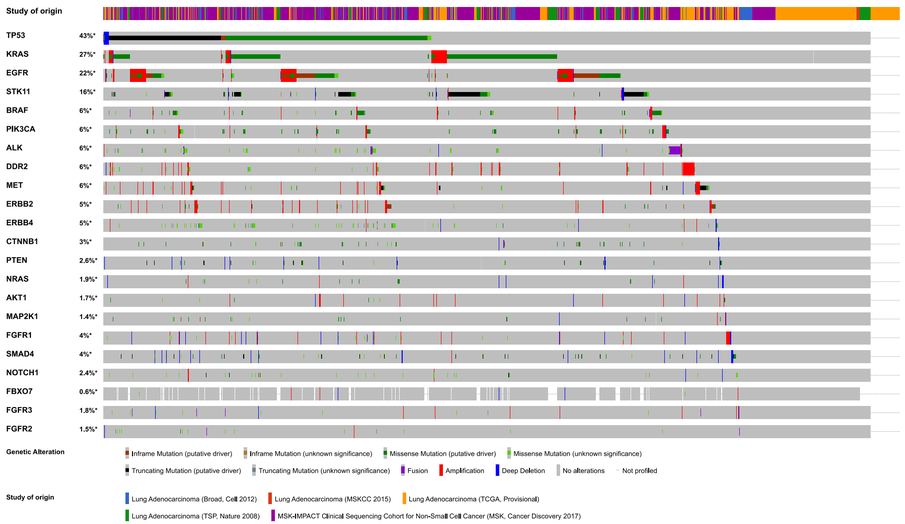


Figura 2. Profili mutazionali dei 22 geni coinvolti nel pannello clinico mutazionale da utilizzare nello studio in oggetto

Disegno sperimentale

Tipo si studio: osservazionale retrospettivo monocentrico

Dipartimenti e S.C.D.U. coinvolti:

A) S.C.D.U. Radiodiagnostica (Direttore Prof. Veltri), S.C.D.U. Anatomia Patologica Oncologia (Direttore Prof. Volante), S.C.D.U. Oncologia (Direttore Prof. Scagliotti) - **Dipartimento di Oncologia - Università degli Studi di Torino**

B) **Dipartimento di Informatica - Università degli Studi di Torino** (Prof. Marco Aldinucci, Dott.ssa Francesca Cordero)

SISTEMA SANITARIO REGIONALE DEL PIEMONTE
A.O.U. SAN LUIGI GONZAGA
Regione Gonzole, 10 – 10043 Orbassano (TO)

COMITATO ETICO INTERAZIENDALE
A.O.U. SAN LUIGI GONZAGA DI ORBASSANO
AA.SS.LL. TO3 – TO4 – TO5

Azienda Ospedaliero-Universitaria San Luigi Gonzaga di Orbassano

Al **prof. Andrea VELTRI**
S.C.D.U. Radiodiagnostica
A.O.U. San Luigi Gonzaga – Orbassano

Al **DIRETTORE GENERALE**
A.O.U. San Luigi Gonzaga – Orbassano

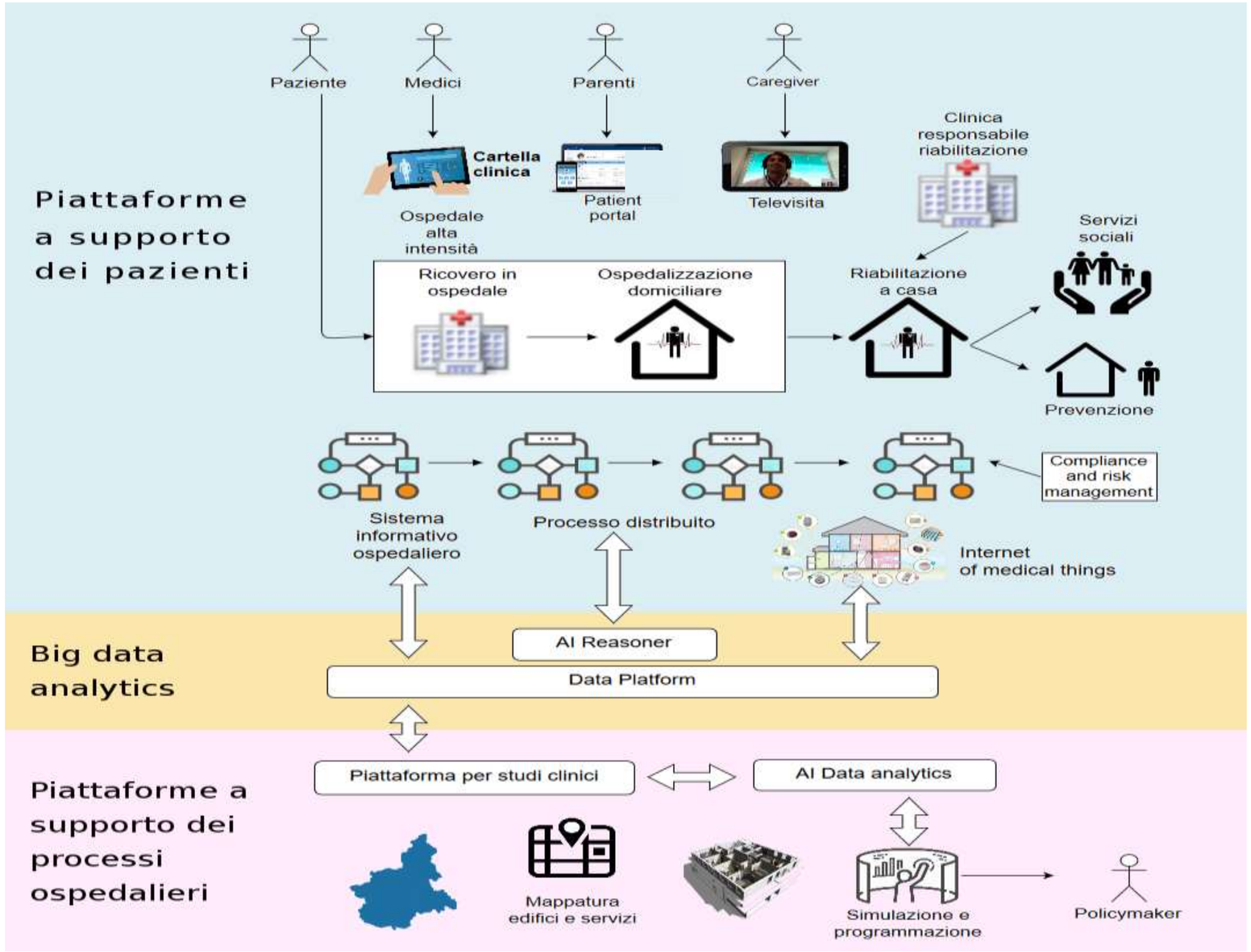
A.O.U. S. LUIGI GONZAGA
11 MAR. '19
Prot. n° 4005
Tit. 02 Cat. 06
Classe ... Sottoclasse ...

PRATICA N. 32/2019 **SEDUTA DEL 27 FEBBRAIO 2019**

Protocollo: CT-guided biopsy of NSCLC: radiomic analysis of the CT images and correlation with NGS findings
PROMOTORE:S.C.D.U. Radiodiagnostica – A.O.U. San Luigi Gonzaga
SEDE: S.C.D.U. Radiodiagnostica – A.O.U. San Luigi Gonzaga
SPERIMENTATORE RESPONSABILE: **prof. Andrea VELTRI**

Il Comitato Etico Interaziendale San Luigi Gonzaga, è istituto con deliberazione del Commissario dell'A.O.U. San Luigi Gonzaga di Orbassano n. 609 del 20 dicembre 2016, ai sensi del Decreto Legislativo 24/06/2003 n. 211 e D.M. 12/05/2006, del decreto-legge 13 settembre 2012, n. 158, convertito, con modificazione, dalla legge 8.11.2012, n. 189 e del D.M. Salute 8 febbraio 2013, secondo quanto previsto dalla D.G.R. del Piemonte 25-6008 del 25/06/2013, opera secondo i principi bioetici fondamentali contenuti nella Costituzione Italiana, nel Codice di Norimberga (1946), nella Dichiarazione Internazionale dei Diritti dell'Uomo (1948), nella revisione corrente della Dichiarazione di Helsinki, nella Convenzione sui diritti umani e la Biomedicina del Consiglio d'Europa (Oviedo, 1997), nelle Good Clinical Practice – Linee Guida di Buona Pratica Clinica.

Nella seduta del 27 febbraio 2019 il Comitato Etico San Luigi Gonzaga, dopo lunga e approfondita disamina della documentazione presentata per il protocollo in oggetto, ritiene che lo studio possa esser approvato purché vengano apportate le seguenti modifiche/ integrazioni:



CANP – la CAsa Nel Parco

Settembre 2018 SAL

CANP

LA CASA NEL PARCO (REGIONE PIEMONTE FESR2014-18 11M€)
2018 - COORD. GUIDO BOELLA, INFORMATICA@UNITO

HPC4AI

METHODOLOGY