



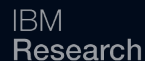
3RD SUMMER SCHOOL ON HIGH-PERFORMANCE AND DISRUPTIVE COMPUTING IN REMOTE SENSING

HDCRS - Working Group of the IEEE GRSS Earth Science Informatics Technical Committee (ESI TC)

PROF. GABRIELE CAVALLARO (FORSCHUNGSZENTRUM JÜLICH | UNIVERSITY OF ICELAND)

PROF. DORA BLANCO HERAS (UNIVERSITY OF SANTIAGO DE COMPOSTELA)

DR. JIN SUN (NANJING UNIVERSITY OF SCIENCE AND TECHNOLOGY)



CITIUS

Centro Singular de Investigación
en Tecnoloxías Intelixentes

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA



TEAM +140 PEOPLE

35 SENIOR RESEARCHERS

citius.gal

COORDINATED EUROPEAN PROJECTS



NANOVR

Nanoscale Design using Virtual Reality



Marie Skłodowska-Curie Actions



NL4XAI

Interactive Natural Language for Explainable AI (Coordinator)



MENLAOS NT

Multimodal Fusion of Sensor Information

HYBRIDS

Hybrid intelligence to monitor, promote and analyse transformations in good democracy practices

SPIN-OFFS



inverbis.



R&D RESULTS 2019-21

ARTICLES **72%** in Q1
36% in D1

22% in top 10% most cited

9,2M €
16% income from european projects

OPPORTUNITIES FOR POSTDOCS

- o Juan de la Cierva contracts
- o Ramón y Cajal contracts
- o Marie Skłodowska-Curie Postdoctoral Fellowships (MSCA-PF)
- o ERC-Starting Grant
- o Postdoc CiTIUS

Welcome package with special benefits for each call. Contact us!



Based in Santiago de Compostela, UNESCO World Heritage City. One of the 100 best places in the world by TIME Magazine 2021.



[citius.gal](https://www.citius.gal)

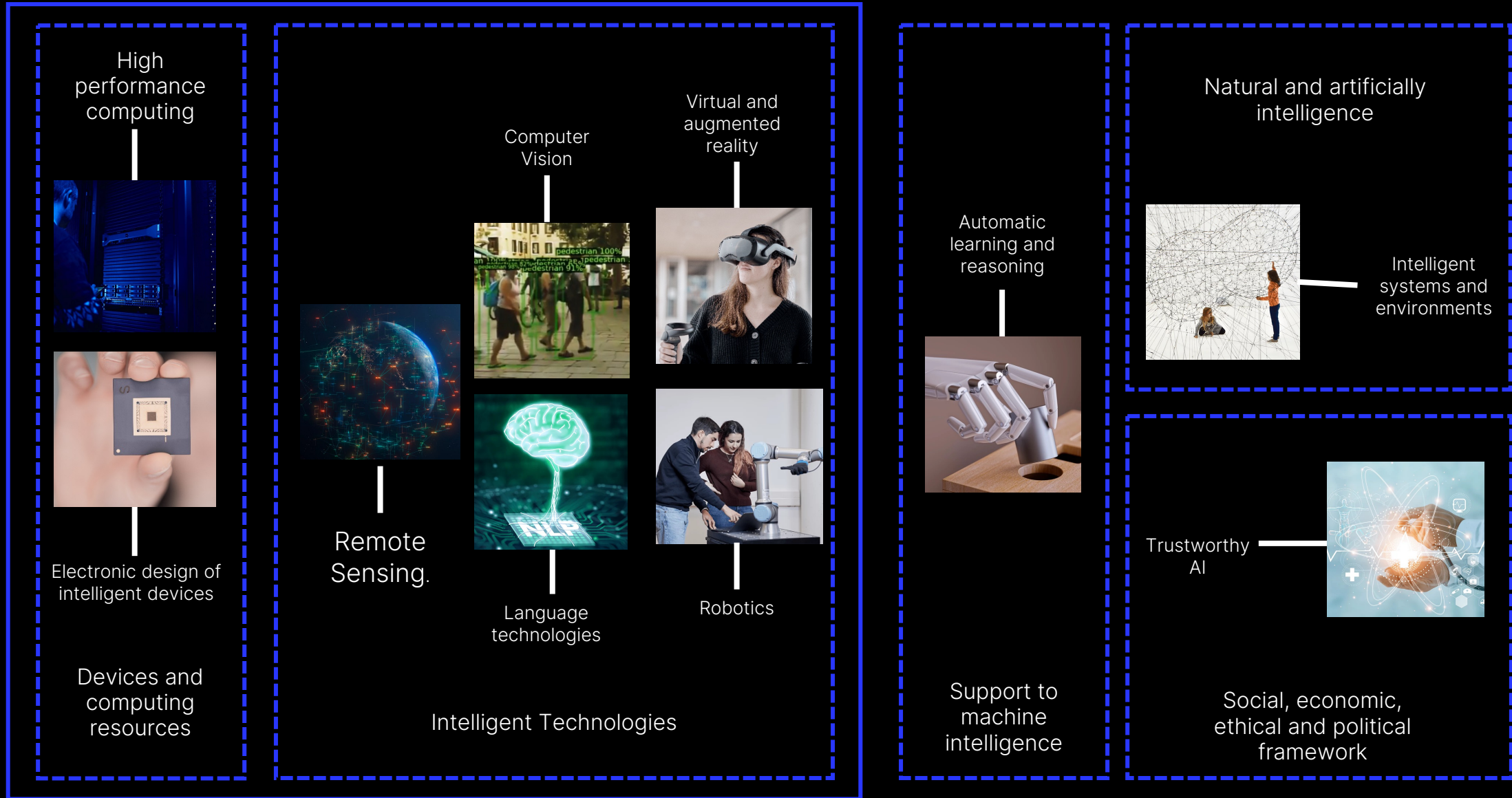


citius.kmt@usc.es



HR EXCELLENCE IN RESEARCH

CiTIUS: Scientific Areas



The HPC and remote sensing group



- 20 years in HPC (15 years in Remote Sensing)
- 7 PhD in Remote Sensing and HPC finished
- 2 seniors and 5 researchers
- 5 PhD students and 10 master students

WHAT IS REMOTE SENSING?

Observing an object from a distance without physical contact



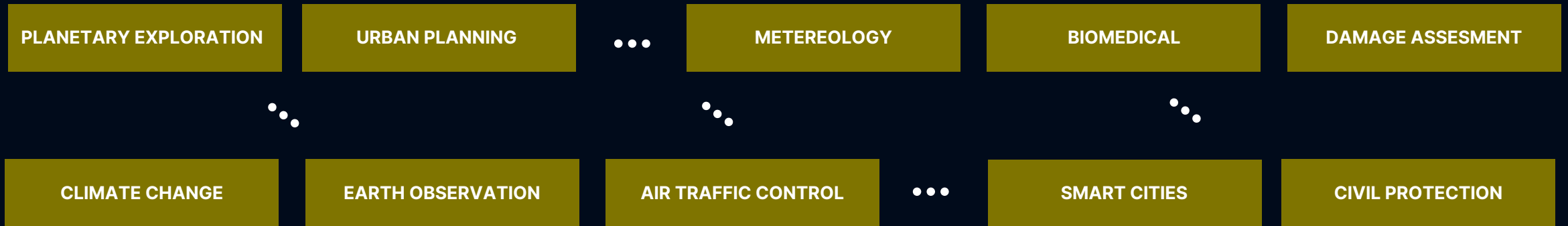
Satellite-based



close-range

APPLICATIONS OF REMOTE SENSING

Observing objects and phenomena from a distance without physical contact allows for numerous applications



- Non invasive method in contrast to in situ or on-site observation
- Efficient and continuous observation of the Earth and its changes
- Satellite platforms provide repetitive and consistent view

EARTH OBSERVATION AT DIFFERENT SCOPES

Solutions developed at three levels of increasing complexity (time response, amount of data, computational resources):

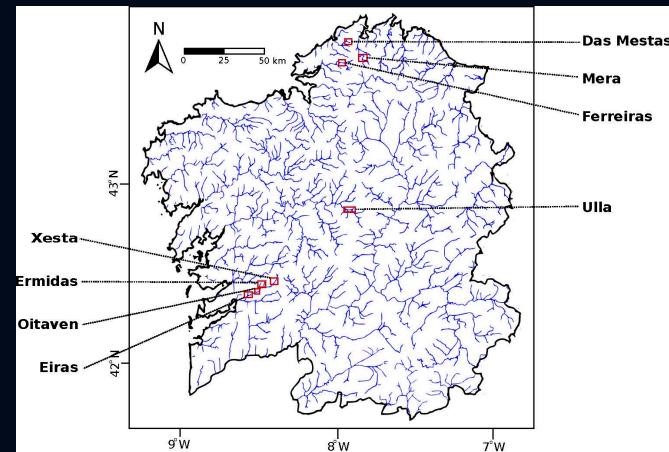
Local : a company, a group of users,...

Time: short



Regional or national: public institutions, companies,...

Time: medium



Global: governments, institutions,...

Time: long

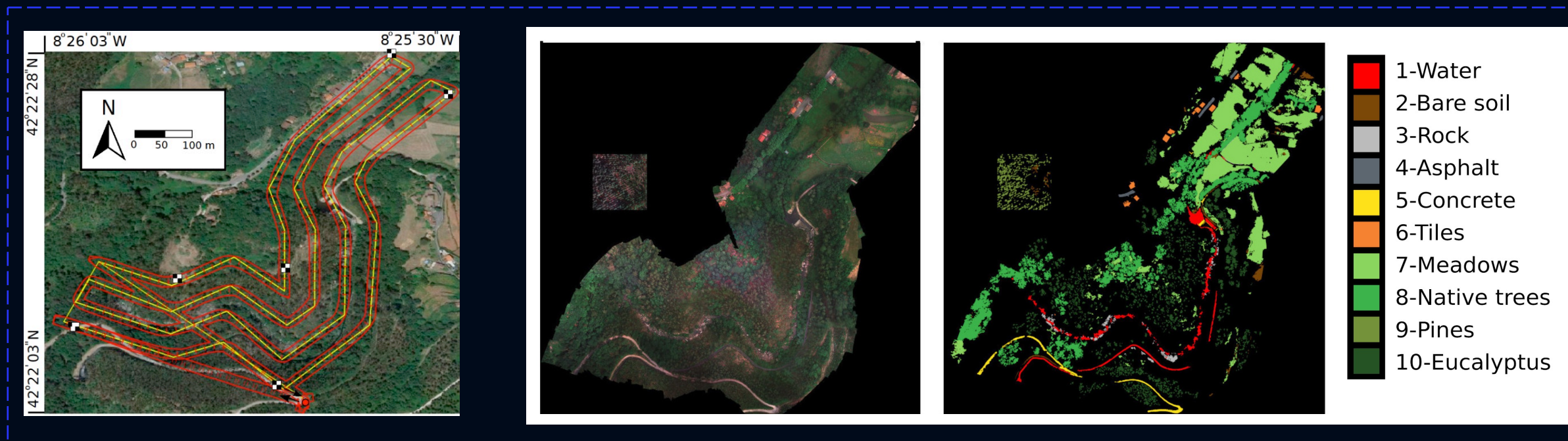


CITIUS GROUP ON RS AND HPC

Solutions at the first and second level for extraction of information aimed at helping the final users in making decisions.

PRE DL ERA

- Mainly from multi and hyperspectral sensor imaging. Integrating information from different sources.
- With strong focus on monitoring agriculture and natural ecosystems.
- Increasing the amount and quality of the information and providing it in a timely manner.



Watershed monitoring at regional level in Spain

CITIUS GROUP ON RS AND HPC

Solutions that require a variety of techniques and approaches

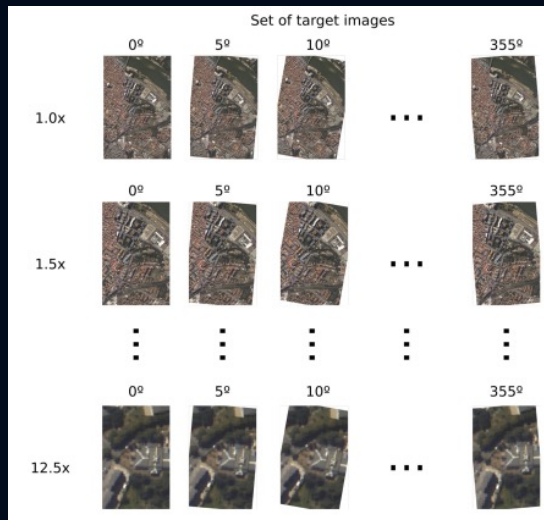
PRE DL ERA

Analyzing, capturing

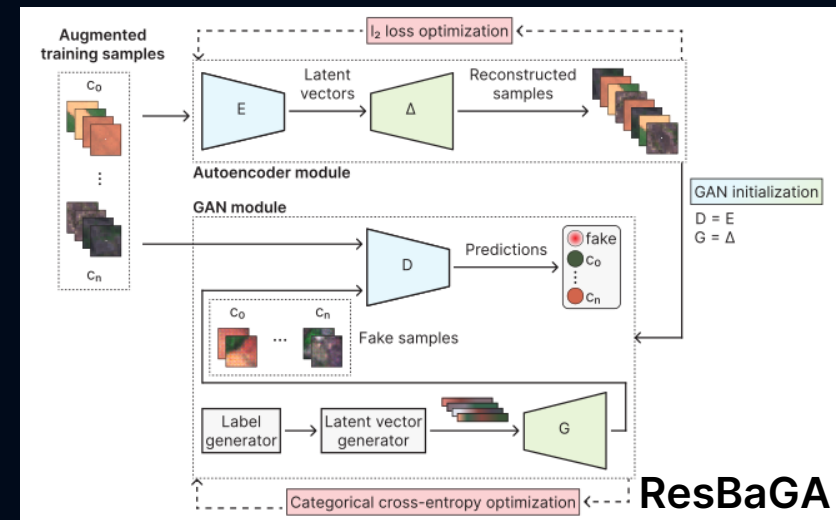
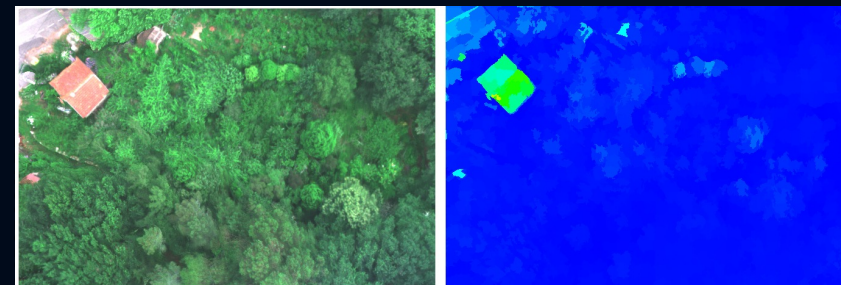
Threats



Preprocessing, storing



Solving: classification, change detection, anomaly detection,...



SATELLITES ORBITING THE EARTH IN 2022

8261 individual satellites orbiting the Earth at the end of January; an increase of 11.84% compared to April 2021



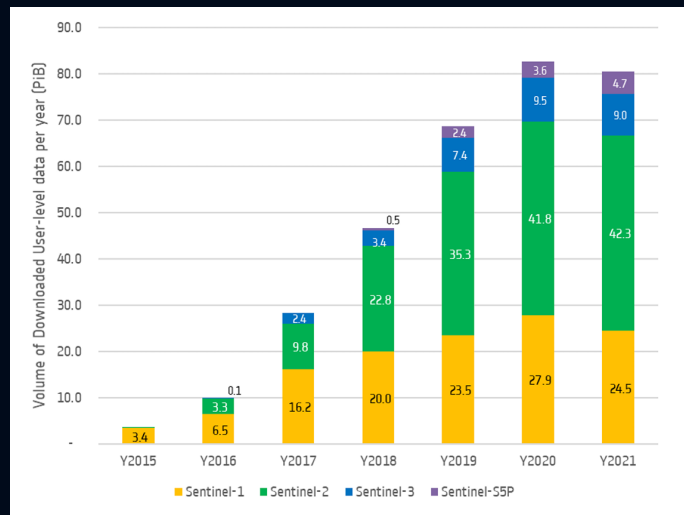
<https://www.pixalytics.com/satellites-in-2022/>

<https://physicsworld.com/a/space-debris-threat-to-geosynchronous-satellites-has-been-dramatically-underestimated/>

- To the end of January 2022, there have been 12293 objects launched into space
- Communications (3135), **Earth observation (1030)**, Technology development/demonstration (385)
- Navigation/positioning (154), Space science/observation (108), Earth science (22), Other purposes (18)

COPERNICUS SENTINEL DATA

Sentinel data available for retrieval in 2021 was 41.86 PB, with a total download volume of 80.5 PB



<https://scihub.copernicus.eu/twiki/do/view/SciHubWebPortal/AnnualReport2021>

- The Copernicus Open Access Hub provides complete, free and open access to Sentinel data
- Platforms: Sentinel-1, Sentinel-2, Sentinel-3 and Sentinel-5P user products
- 490,000 registered users with an average daily download volume of 203 TiB

EARTH OBSERVATION DATA

Information about the physical, chemical, and biological systems of the planet Earth



A market in constant expansion

Upstream

Space
Infrastructure

Manufacturing of satellites,
ground systems, launchers,
ground operations

Midstream

Data Sales

Data acquisition Data
processing Data
archiving Data
distribution and delivery

Downstream

Value-adding
services Data

Processing, adaptation to
users' need, integration with
data from other sources

Digital Transformation Monitor, Big Data in Earth Observation, EASME/COSME/2017:
<https://ati.ec.europa.eu/sites/default/files/2020-06/Big%20Data%20in%20Earth%20Observation%20%28v1%29.pdf>
EUSPA EO and GNSS Market Report 2022:
https://www.euspa.europa.eu/sites/default/files/uploads/euspa_market_report_2022.pdf

EO IS BEYOND BIG DATA

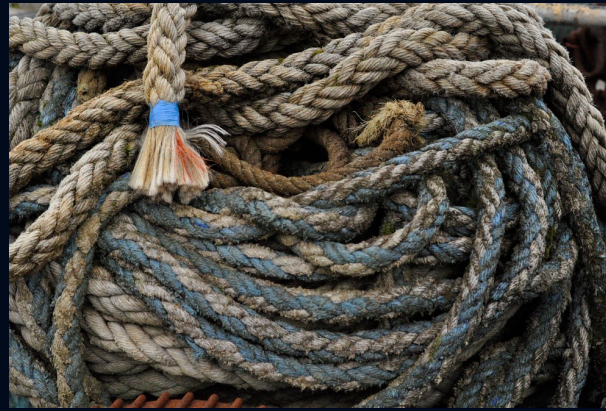
Volume, Variety, Veracity, ... Velocity and ...

High Dimensional



Data reflect complex relationships between natural and social phenomena

High Complexity



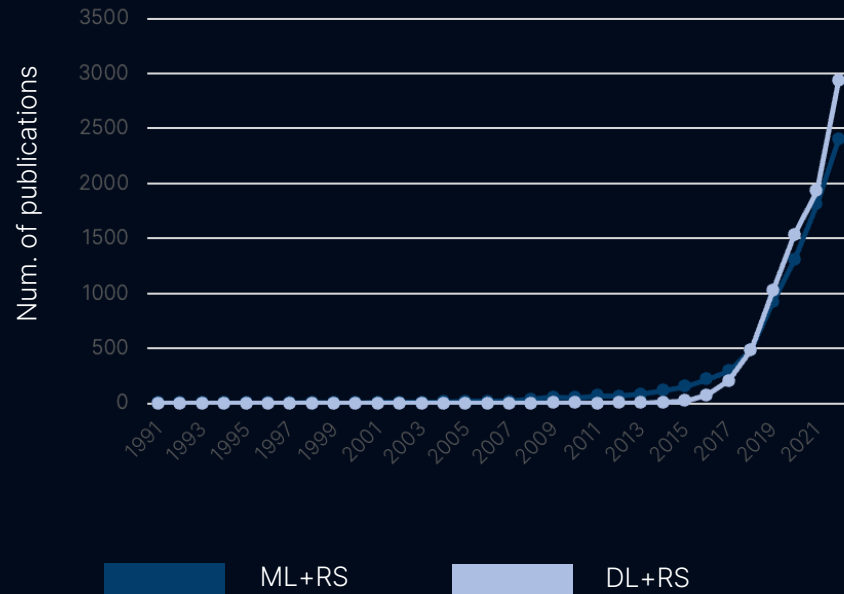
Data is incorporated in highly complex models

High Uncertainty

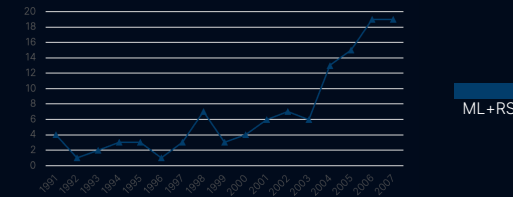


Data errors or incompleteness of data is unavoidable

MACHINE LEARNING AND DEEP LEARNING IN REMOTE SENSING

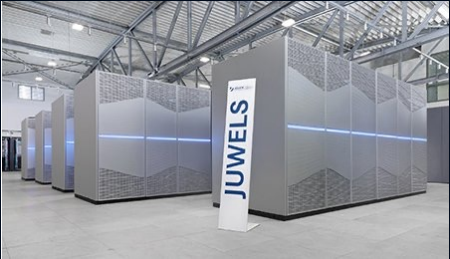


- Classical ML such as Support Vector Machine (SVM) and Random Forest (RF) since the '90s



- DL unleashed advances in the last decade
- Figures may differ depending on the source, but the overall trend remains consistent

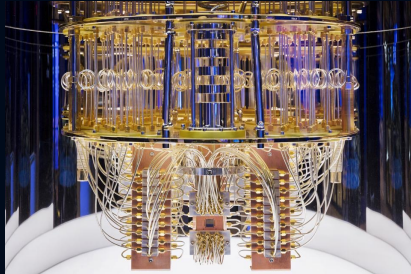
EMERGING COMPUTING PARADIGMS



Supercomputing



Specialized Hardware Computing



Quantum Computing



Edge Computing



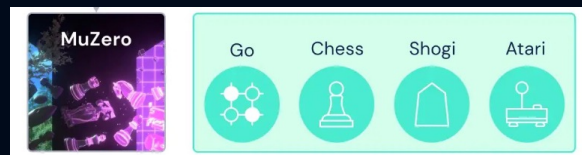
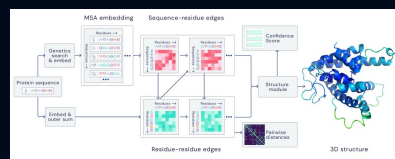
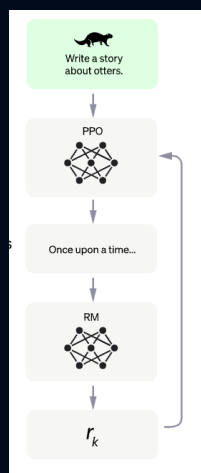
Cloud computing

...



Blockchain

MOST BREAKTHROUGHS REQUIRE HEAVY COMPUTE POWER, USING MANY ACCELERATORS SIMULTANEOUSLY



- GPT-3: natural language generation, language understanding
- CLIP, DALL-E 2, Stable Diffusion: image understanding and image generation
- AlphaFold 2: protein structure prediction
- AlphaZero, MuZero: learning control in highly dimensional state-action spaces

<https://openai.com/blog/chatgpt>

AlphaFold: a solution to a 50-year-old grand challenge in biology, <https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology>

MuZero: Mastering Go, chess, shogi and Atari without rules, <https://www.deepmind.com/blog/muzero-mastering-go-chess-shogi-and-atari-without-rules>

Jenia Jitsev, Towards Scalable Deep Learning, Scalable Learning & Multi-Purpose AI Lab, Helmholtz AI, LAION @ JSC

CURRENT POPULARITY OF SUPERCOMPUTERS

TPU v4: An Optically Reconfigurable Supercomputer for Machine Learning with Hardware Support for Embeddings

Industrial Product*

Norman P. Jouppi, George Kurian, Sheng Li, Peter Ma, Rahul Nagarajan, Lifeng Nai, Nishant Patil, Suvinay Subramanian, Andy Swing, Brian Towles, Cliff Young, Xiang Zhou, Zongwei Zhou, and David Patterson
Google, Mountain View, CA

FORBES > INNOVATION > SUSTAINABILITY

Tesla's Biggest News At AI Day Was The Dojo Supercomputer, Not The Optimus Robot

James Morris Contributor
I write about the rapidly growing world of electric vehicles

Follow

0

Oct 6, 2022, 07:23am EDT

Tech > Science

BABY STEPS Google artificial intelligence supercomputer creates its own 'AI child' that can outperform its human-made rivals

The NASNet system was created by a neural network called AutoML earlier this year

Mark Hodge

Published: 15:22, 5 Dec 2017 | Updated: 11:27, 6 Dec 2017

RESEARCH

Introducing the AI Research SuperCluster — Meta's cutting-edge AI supercomputer for AI research

January 24, 2022

- Big tech companies are announcing their AI supercomputers
- Supercomputing now goes far beyond traditional scientific computing, which was driven by large governments
- The field is currently propelled by major industries building highly specialized supercomputers

N. P. Jouppi, G. Kurian, et al., "TPU v4: An Optically Reconfigurable Supercomputer for Machine Learning with Hardware Support for Embeddings", 2023, <https://doi.org/10.48550/arXiv.2304.01433>

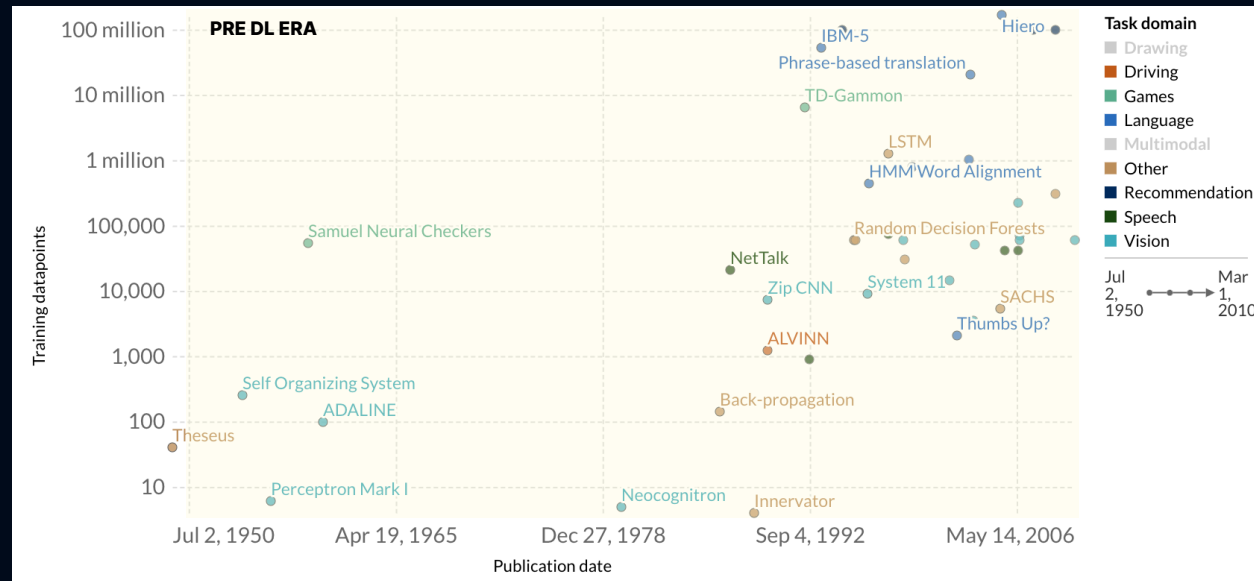
<https://www.forbes.com/sites/jamesmorris/2022/10/06/teslas-biggest-news-at-ai-day-was-the-dojo-supercomputer-not-the-optimus-robot/?sh=22ba4ab780bd>

<https://ai.facebook.com/blog/ai-rsc/>

<https://www.thesun.co.uk/tech/5072741/google-nasnet-ai-child-reinforcement-learning/>

Torsten Hoefler, "Efficient AI: From supercomputers to smartphones", Scalable Parallel Computing Lab @ ETH Zurich, <https://youtu.be/xxwT45IjG4o>

PRE DEEP LEARNING ERA (1950-2010)

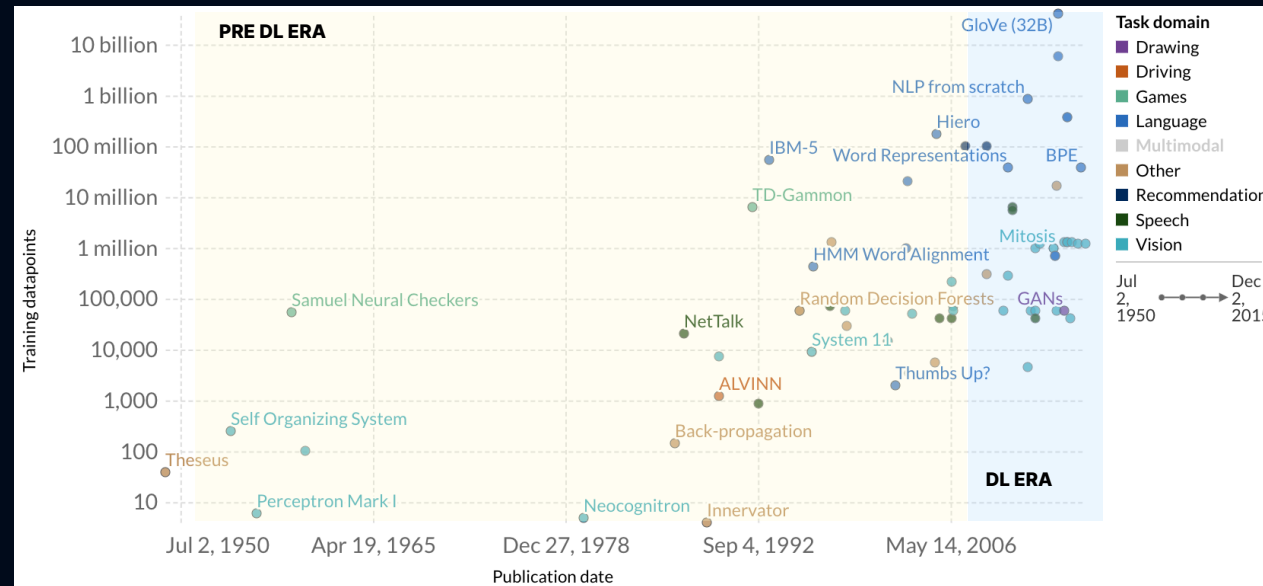


C. Giattino, et al., Artificial Intelligence, <https://ourworldindata.org/artificial-intelligence>

J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn and P. Villalobos, "Compute Trends Across Three Eras of Machine Learning," 2022 International Joint Conference on Neural Networks (IJCNN), pp. 1-8, 2022, <https://doi.org/10.1109/IJCNN55064.2022.9891914>

- Training compute doubled every 17 to 29 months
- This increase is roughly in line with Moore's Law

DEEP LEARNING ERA AROUND 2010

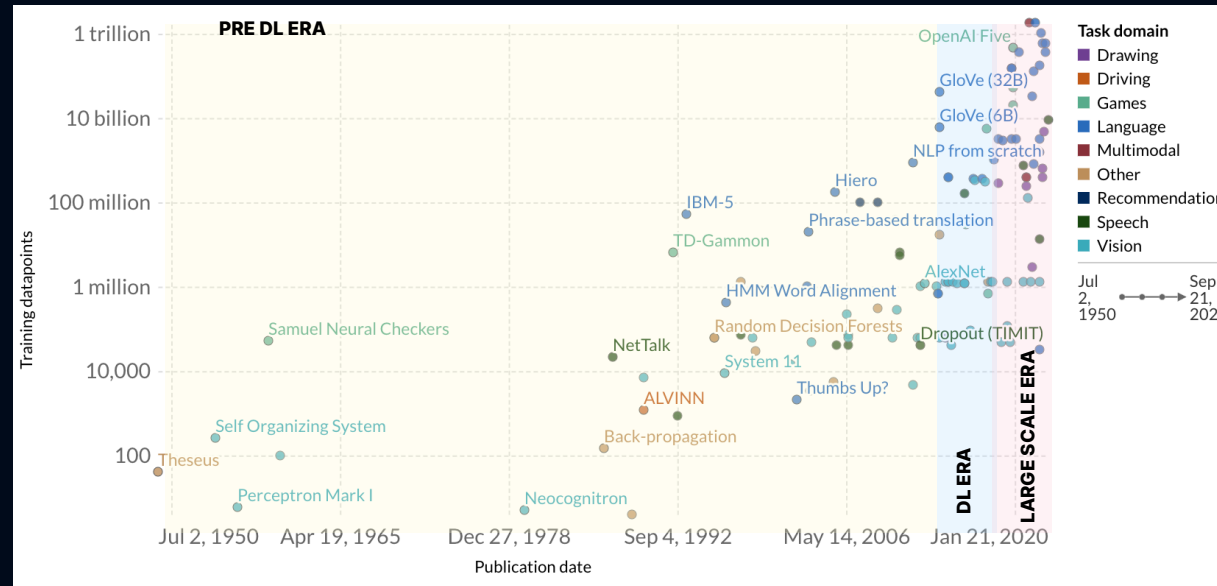


C. Giattino, et al., Artificial Intelligence, <https://ourworldindata.org/artificial-intelligence>

J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn and P. Villalobos, "Compute Trends Across Three Eras of Machine Learning," 2022 International Joint Conference on Neural Networks (IJCNN), pp. 1-8, 2022, <https://doi.org/10.1109/IJCNN55064.2022.9891914>

- The trend accelerated along the with the popularity of deep learning
- From around 2010, compute doubles every 4 to 9 months

LARGE-SCALE ERA AROUND 2015



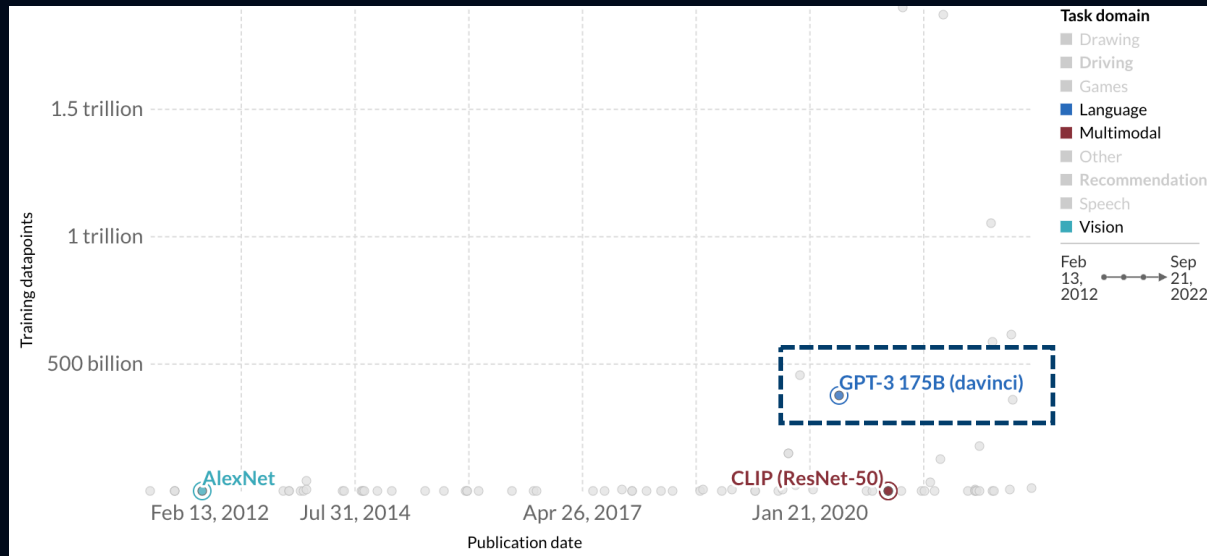
C. Giattino, et al., *Artificial Intelligence*, <https://ourworldindata.org/artificial-intelligence>

J. Sevilla, L. Heim, A. Ho, T. Besiroglu, M. Hobbhahn and P. Villalobos, "Compute Trends Across Three Eras of Machine Learning," 2022 International Joint Conference on Neural Networks (IJCNN), pp. 1-8, 2022, <https://doi.org/10.1109/IJCNN55064.2022.9891914>

D. Silver et al., "Mastering the game of Go without human knowledge," *Nature*, vol. 550, pp. 354–359, 2017, <https://doi.org/10.1038/nature24270>

- In late 2015, a new trend of large-scale models emerged
- Computational capacity significantly higher than that of other models published in the same year (e.g., release of AlphaGo)
- This growth trend is slower than the overall DL trend, with a doubling time of roughly every 8 to 17 months

GPT-3 MODEL (175 BILLION WEIGHT PARAMETERS)



- AlexNet, winner ILSRVC 2012 – 60M
- ResNet-50, winner ILSRVC 2015 – 25M
- CLIP ViT L/14, multi-modal learning, 2021 - 600M

TIME REQUIRED FOR A FULL TRAINING OF GPT-3 MODEL



Ranking in November 2020
(TOP500 (7 World, 1 Europe), Green500 (1 in TOP100) TOP10 AI (4))

- ≈ 355 years with one Nvidia Volta 100 GPU
- ≈ 90 years with one Nvidia Ampere 100 GPU
- ≈ 16 days if scaled well with 2,000 A100 GPUs on JUWELS Booster

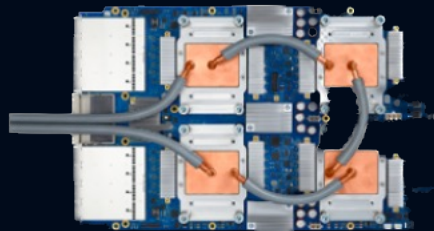
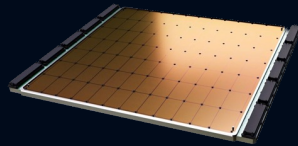
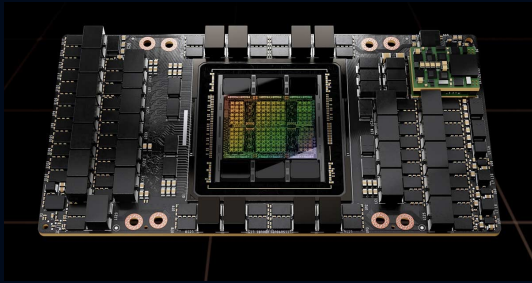


Julich Supercomputing Centre, "JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Julich Supercomputing Centre," *Journal of large-scale research facilities*, vol. 7, no. A138, 2021.

S. Kesselheim, A. Herten, K. Krajsek, J. Ebert, J. Jitsev, M. Cherti, M. Langguth, B. Gong, S. Stadler, A. Mozaffari, G. Cavallaro, R. Sedona, A. Schug, A. Strube, R. Kamath, M. G. Schultz, M. Riedel, and T. Lippert, "JUWELS Booster – A Supercomputer for Large-Scale AI Research," in *High Performance Computing* (H. Jagode, H. Anzt, H. Ltaief, and P. Luszczek, eds.), (Cham), pp. 453–468, Springer International Publishing, 2021.

Jenia Jitsev, Towards Scalable Deep Learning, Scalable Learning & Multi-Purpose AI Lab, Helmholtz AI, LAION @ JSC

TRAINING OF MODELS REQUIRES ACCELERATORS



- GPUs (currently NVIDIA dominant), TPUs (Google)
- GPUs: generic deep learning hardware (parallelizing matrix/tensor operations via vectorization)
- Specialized hardware, eg. in-memory computing chips, Graphcore IPU: Colossus MK2, Cerebras Wafer Scale Engine 2 (850k cores)

<https://www.nvidia.com/en-us/data-center/technologies/hopper-architecture/>

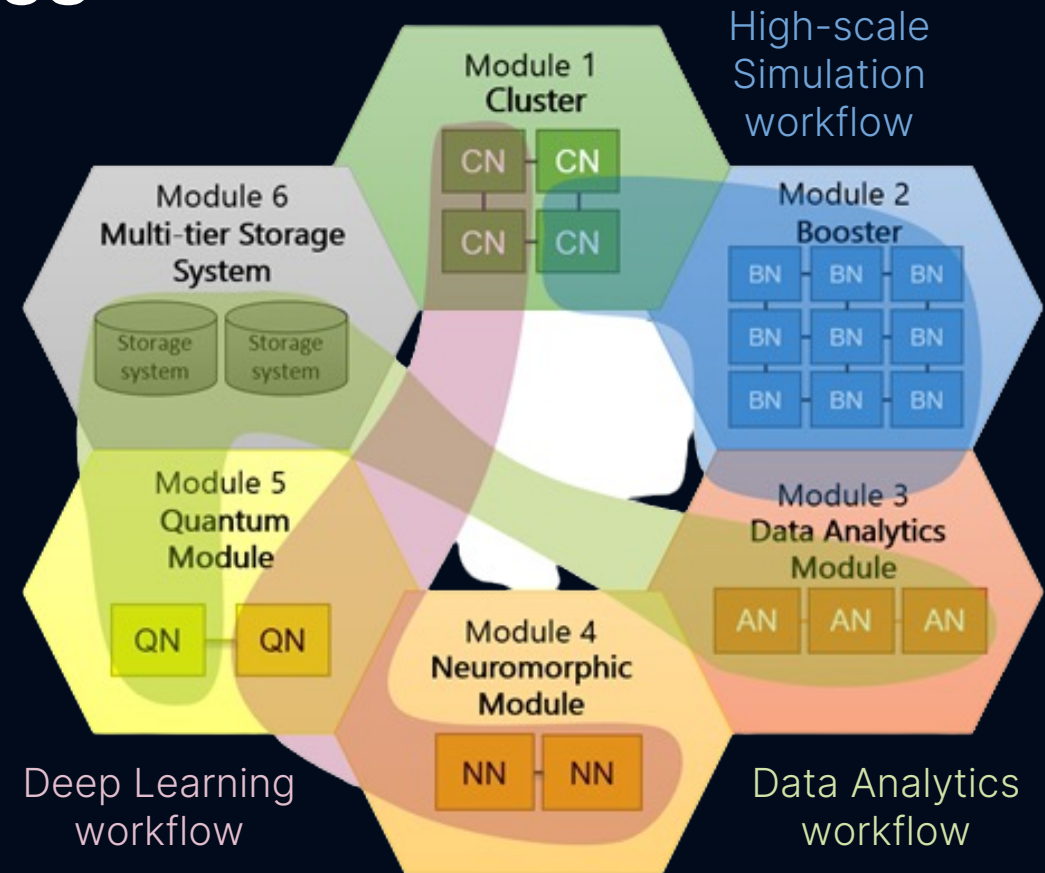
<https://cloud.google.com/tpu>

<https://www.graphcore.ai/products/ipu>

<https://www.cerebras.net/product-chip/>

MODERN HPC SYSTEMS ARE NOT ONLY MASSIVELY PARALLEL BUT ALSO HETEROGENEOUS

- Cost-effective scaling
- Effective resource-sharing
- Fit application diversity
 - Large-scale simulations
 - Data analytics
 - Machine- and Deep Learning
- Composability of heterogeneous resources



Modular Supercomputing Architecture (MSA)

JU PIONEER FOR INNOVATIVE AND TRANSFORMATIVE EXASCALE RESEARCH (JUPITER)



JÜLICH
SUPERCOMPUTING
CENTRE



Mars
(JUWELS)

Jupiter



Mars: d= 6,800 Km
Jupiter: d= 143,00 Km

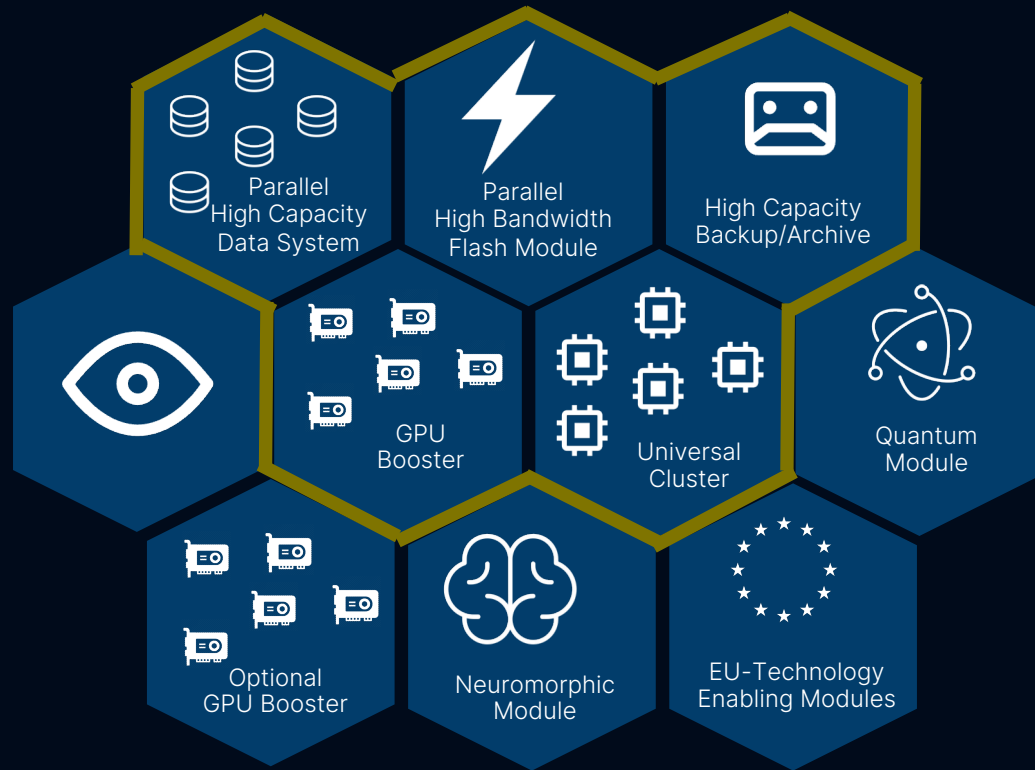
JUWELS: 85 PF/s peak
JUPITER: 1,300 PF/s peak

Thomas Lippert, "Jupiter Ascending, Driving the Forefront of Europe's HPC Effort", EuroHPC Summit 2023, <https://www.eurohpcsummit.eu/>

JUPITER | The Arrival of Exascale in Europe, <https://www.fz-juelich.de/en/ias/jsc/jupiter>

- The Forschungszentrum Jülich in Germany will house Europe's first exascale computer, named JUPITER
- This machine is set to surpass the threshold of one quintillion (a "1" followed by 18 zeros) calculations per second
- JUWELS 20 = JUPITER (when using the diameters of planets as a measurement scale)

JUPITER WILL BE A MODULAR SUPERCOMPUTER



- Its modules will be able to interactively and dynamically work together (it is not just a set of modules/machines)
- Target >20x application performance compared to JUWELS Booster
- Performance: >1 Exabyte storage (with different storage technologies), > 1 EF (GPU), 7,5 PB/s (x86 or ARM)



EuroHPC
Joint Undertaking



JÜLICH
Forschungszentrum
JÜLICH SUPERCOMPUTING
CENTRE

Thomas Lippert, "Jupiter Ascending, Driving the Forefront of Europe's HPC Effort", EuroHPC Summit 2023, <https://www.eurohpcsummit.eu/>

JUPITER | The Arrival of Exascale in Europe, <https://www.fz-juelich.de/en/ias/jsc/jupiter>

DESTINATION EARTH



- HPC is used intensively to run highly complex simulations of the Earth system at the very high resolution of the digital twins
- The flexibility of modular systems is suitable for DestinE
- It's important to quantify the uncertainty of the forecasts and to perform complex data fusion, data handling, and access operations.

Destination Earth (DestinE), <https://destination-earth.eu/>

Destination Earth, a European Commission flagship initiative for a sustainable future, <https://digital-strategy.ec.europa.eu/en/policies/destination-earth>

Destination Earth, https://www.esa.int/Applications/Observing_the_Earth/Destination_Earth

GRSS TECHNICAL COMMITTEES

Earth Science Informatics



Frequency allocations
in Remote Sensing



Geoscience Spaceborne
Imaging Spectroscopy



Image Analysis and
Data Fusion



Instrumentation and Future
Technologies



Modeling in Remote Sensing



Remote sensing Environment,
Analysis and Climate
Technologies



GRSS Standards for
Earth Observations



EARTH SCIENCE INFORMATICS (ESI) TECHNICAL COMMITTEE

Objectives

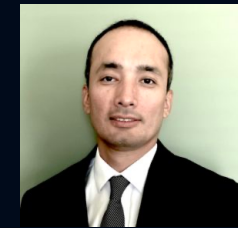
Advance application of informatics to
geoscience and remote sensing

<https://www.grss-ieee.org/technical-committees/earth-science-informatics/>

Chairs



Peter Baumann



Manil Maskey

TWO WORKING GROUPS

High-performance and Disruptive Computing in Remote Sensing (HDCRS)



Gabriele Cavallaro



Dora Blanco Heras



Jin Sun

Databases in Remote Sensing (DBRS)



Dai-Hai Ton That



Kesheng (John) Wu



Khalid Belhajjame

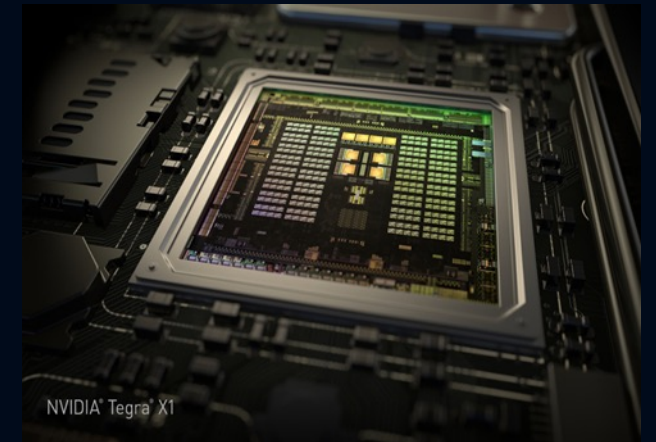
High Performance and Disruptive Computing in Remote Sensing Working Group

Main Objective:

Connect and support the community of interdisciplinary researchers in remote sensing who are specialized in emerging computing paradigms

The idea:

Innovative computing technologies applied to efficient computation of remote sensing problems



<https://www.grss-ieee.org/community/groups/initiatives/high-performance-and-disruptive-computing-in-remote-sensing-hdcrs/>

COME ABOARD

About Earth Science Informatics



MISSION

The Earth Science Informatics Technical Committee (ESI TC) provides a venue for informatics professionals to exchange ideas and share knowledge. It aims at advancing application of informatics to geosciences and remote sensing, assessing technology to support data stewardship and management, and promoting best practices and lessons learned.

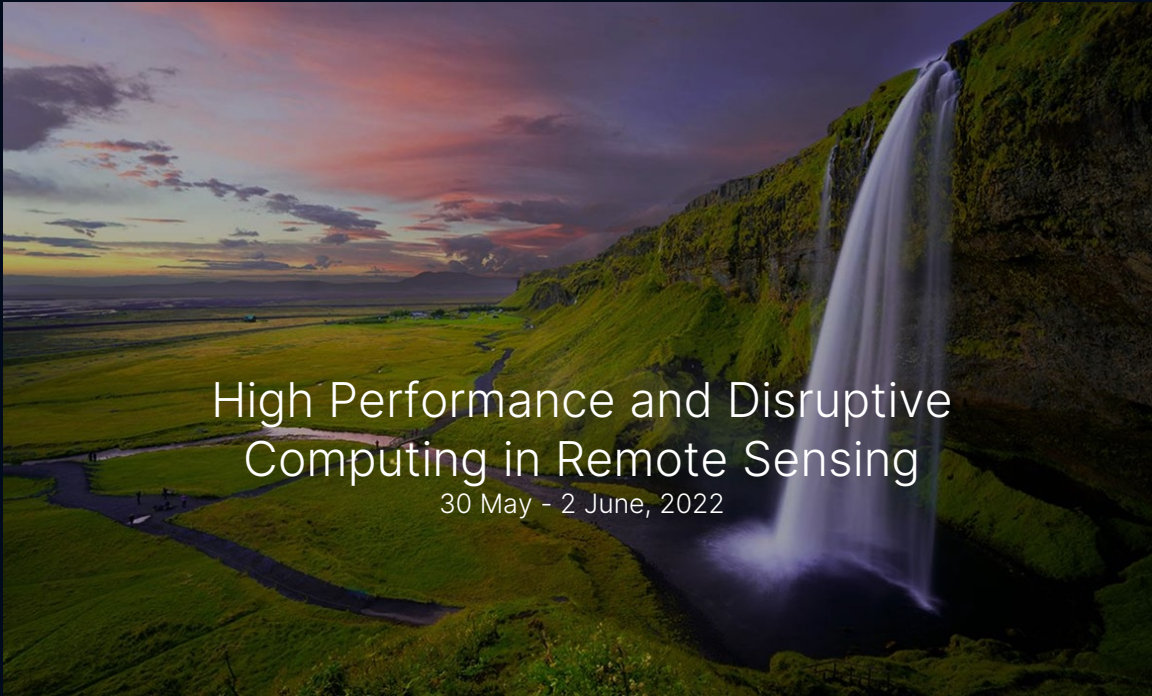
The mission of the ESI TC is to bring together informatics experts and practitioners to share ideas and information to support open science and maximize the use of science data for research and applications.



<https://www.grss-ieee.org/technical-committees/earth-science-informatics/>

HDCRS Activities

HDCRS SUMMER SCHOOL - 2022



Teaching material and videos available at
<https://www.grss-ieee.org/community/groups-initiatives/high-performance-and-disruptive-computing-in-remote-sensing-hdcrs/hdcrs-summer-school-2022/>



UPCOMING ACTIVITIES AT IGARSS 2023

Tutorial on End-to-End Machine Learning with Supercomputing and in the Cloud



Hands-on experience in Supercomputer Systems and Cloud Computing Services for Remote Sensing Applications

Community Contributed Sessions



- (1) Scalable Parallel Computing for Remote Sensing
- (2) Quantum computing next generation HPC
- (3) Quantum Machine Learning algorithms for EO



<https://2023.ieeeigarss.org/>

UPCOMING ACTIVITIES AT M2GARSS 2024



Tutorial on “Quantum Computing for Earth Observation”

<https://2024.m2garss.org/tutorials.php>

ABOUT THIS SCHOOL

Network with students and young professionals, as well as senior researcher and professors who work on interdisciplinary research in remote sensing

Supercomputing

Distributed computing

Specialized hardware computing

Edge computing

Blockchain

Parallel Programming Models

Quantum Computing

Neuromorphic computing



Will be organized every year and cover different research topics

AGENDA – MONDAY MAY 29TH

09:30 - 09:45 (GMT)	Welcome at the University of Iceland and Opening of the Summer School	Jón Atli Benediktsson
09:45 - 10:30	Work and Activities of the HDCRS Working Group	Dora Blanco Heras Gabriele Cavallaro
10:30 - 11:00	Coffee break	-
11:00 - 12:00	Lowering the Barrier for Modern Cloud-based Geospatial Big Data Analysis by Combined Use of Innovative and Traditional Infrastructure	Serkan Girgin
12:00 – 13:30	Lunch break	-
13:30 - 15:00	Hyperspectral technology: inspiring ideas, challenges and opportunities	José Francisco López Roberto Sarmiento
15:00 - 15:30	Coffee break	-
15:30 - 16:30	An Overview of the European HPC Strategy and Highlights from the Icelandic HPC Communities	Hemanadhan Myneni Morris Riedel
20:00 - 22:00	Social Dinner	-

*Greenwich Mean Time (GMT)

SOCIAL DINNER

20:00 – 22:00



- Höfnin Restaurant, Geirsgata 7, 101 Reykjavík

AGENDA – TUESDAY MAY 30TH

Data Science at Scale: Harmonized Landsat Sentinel (HLS) Case Study

09:00 - 10:30 (GMT)	Data Science Overview, Large Scale Data Harmonization
10:30 - 11:00	Coffee break
11:00 - 12:00	Analysis and Exploration: Analytics Platform, Science Analysis
12:00 - 13:30	Lunch break
13:30 - 15:30	Application: Geospatial Foundation Model
15:30 - 16:00	Coffee break
16:00 - 17:00	Inference with Fine-Tuned Model and Conclusion

Speakers: Manil Maskey, Sean Harkins, Brian Freitag, Muthukumaran Ramasubramanian, Iksha Gurung, Linsong Chu, Paolo Fraccaro, Johannes Jakubik, Blair Edwards

*Greenwich Mean Time (GMT)

AGENDA – WEDNESDAY MAY 31TH

Remote Sensing Deployable Analysis Environment

09:00 - 10:30 (GMT)	HPC, RSDAT, and the EO Software Ecosystem
10:30 - 11:00	Coffee break
11:00 - 12:00	Deployment with RS-DAT and Data Retrieval
12:00 – 13:30	Lunch break
13:30 - 15:00	Scaling EO workflows with HPC
15:00 - 15:30	Coffee break
15:30 - 17:00	Hands-on session

Speakers: Francesco Nattino, Meiert Willem Grootes, Pranav Chandramouli

*Greenwich Mean Time (GMT)

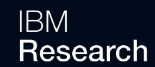
AGENDA – THURSDAY JUNE 1TH

Quantum Computing for Earth Observation

09:00 - 10:30 (GMT)	Introduction to Quantum Computing and its Ecosystem	Riccardo Mengoni
10:30 - 11:00	Coffee break	-
11:00 - 12:00	Introduction to Quantum Computing and its Ecosystem	Riccardo Mengoni
12:00 - 13:30	Lunch break	-
13:30 - 14:30	ESA Quantum Computing for Earth Observation (QC4EO): Current Activities and Perspectives	Bertrand Le Saux
14:30 - 15:30	Is Space ready for the Quantum Leap? A Thales Alenia Space Perspective on Quantum Technologies for Earth Observation	Mattia Verducci Tommaso Catuogno
15:30 - 16:00	Closing	-

*Greenwich Mean Time (GMT)

ACKNOWLEDGMENT



Thank you for your attention