

# CINECA

## 2020 - 2021

### HPC annual report



LEONARDO



ISBN 9788886037426

Cineca HPC Report 2020  
website: [www.hpc.cineca.it](http://www.hpc.cineca.it)  
mail: [info-hpc@cinca.it](mailto:info-hpc@cinca.it)

Contents are © CINECA 2020- 2021 are released under  
the Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)  
Creative Commons license, version 3.0.

Editorial Office  
Cineca Consorzio Interuniversitario  
Via Magnanelli 6/3  
40033 - Casalecchio di Reno (BO) - Italy

Editorial Board  
Sanzio Bassini  
Gabriella Scipione  
Antonella Guidazzoli  
Maria Chiara Liguori

Design and Layout project  
Simona Caraceni

Thanks to:  
Fabio Affinito; Paola Alberigo; Claudio Arlandini; Sanzio Bassini; Carlo Cacciamani; Simona Caraceni; Cinzia Caroli; Mirko Cestari; Patrizia Coluccia; Maurizio Cremonesi; Mattia D'Antonio; Chiara Della Casa; Francesca Delli Ponti; Andrew Emerson; Francesco Falciano; Luca Ferraro; Federico Ficarelli; Franz Fischnaller; Daniela Galetti; Estibaliz Gascon; Silvia Gioiosa; Luca Govoni; Massimiliano Guarrasi; Antonella Guidazzoli; Tim Hewson; Silvano Imboden; Giannandrea Inchingolo; Maria Chiara Liguori; Alessandro Marani; Fabio Massaccesi; Paola Mercovigliano; Andrea Montani; Maurizio Ortali; Daniele Ottaviani; Eric Pascolo; Marco Rorro; Elda Rossi; Elisa Rossi; Monia Santini; Gabriella Scipione; Donatella Sforzini; Deanna Shemek; Giuseppe Trotta; Roberta Turra; Augustin Vintzileos; Francesca Zanella; Cinzia Zannoni.

# CINECA 2020 - 2021

## HPC annual report

Dear colleagues,

the publication of the HPC annual report, summarizing the activities and projects carried out during the period 2020 - early 2021, comes when the pandemic crisis of the SARS Covid 19 in Western countries seems, if not completely overcome, at least under control.

This is due to the response that scientific research has been able to give, initially to contain the effects of the spread of the infection, subsequently making available in an extraordinarily short time effective vaccines capable of immunizing the population.

CINECA has participated in numerous projects relating to this research front to combat the pandemic, and this is a source of pride and a sense of responsibility for the increased awareness of how much HPC is an essential component of the scientific research methodologies of our time.

Time that has appropriately proposed the centrality of scientific research to address the big challenges that affect life on our planet, from climate change to the protection of biodiversity. Problems so complex that only by combining the numerical simulation of phenomena with data assimilation into machine learning and artificial intelligence models can they be faced and solved. And this requires the extreme computing powers at exaflop scale for which as CINECA we are committed to a mature, concrete and sustainable roadmap.

Thinking about the complexity of physical phenomena, one cannot fail to recall another moment of pride, certainly for the Italian research system, but more generally for the international scientific community, relating to the awarding of the Nobel Prize to Prof. Giorgio Parisi.

A year dense of important events that raise great global problems, but also many elements of hopes and positive opportunities.

From this point of view, the next period 2021 - 2022 will be a period of great intensity. First of all for the next production start of the pre-exascale Leonardo system, and then for the planning that we will be called to face due to the many strategic deadlines that are looming on the horizon already in the short term.

First of all, the planning related to the Horizon Europe research and development framework program funded by the European Commission, which enters its implementation phase, in the second instance and as a relative consequence, the activity plan relating to the second period of agreement 2021 - 2030 of EuroHPC, finally, but of extreme importance, the National Recovery and Resilience Plan, financed with funds from Next Generation Europe. These design scenario all see HPC, in all its classic and innovative expressions up to the use of quantum technologies, at the center and as a key enabling reference.

So resilience, relaunch and development in some strategic directions. In the first instance, the consolidation towards the upgrade of the computing infrastructure, both through the planning in the EuroHPC area, and in the EuroFUSION area, due to the fact that calls will be released for the expression of interest for the enhancement of the related supercomputing systems that we host in our datacenter. Secondly, the expansion, through the inclusion of new staff in the HPC high level support team, with particular reference to some areas, but not only, such as meteo climate, in this regard we also greet with exultation the operational start of the National Agency Italia Meteo which will be based in Bologna, physics of materials, health and life science and also with increasing attention to the humanities and social disciplines, finally a relaunch, also by reactivating class in presence, of the training activity, which has not suffered slowdowns even in phase of profound lockdown, but which constitutes a priority of our action, both as regards the continuous and advanced training of the internal staff, and to contribute to the preparation of personnel who will increasingly require research and innovation both public and enterprises and industry.



		The Italy Flash Flood use case in MISTRAL	40
		The new climate datasets in HIGHLANDER	42
		AIDA	44
		Into the (un)known of the Cosmo	46
		<b>BIG DATA AND ARTIFICIAL INTELLIGENCE</b>	
		Big Data in the analysis of TV ratings	50
		Fostering AI innovation	52
		<b>MATERIAL SCIENCES</b>	
		TREX and MaX supporting science to exascale	54
		<b>QUANTUM COMPUTING</b>	
		Cineca quantum computing perspective	57
		<b>HUMANITIES AND SOCIAL SCIENCES, DIGITAL ECOSYSTEMS</b>	
		Isabella d'Este Virtual Studiolo	61
		Memory and archive in motion	63
		TÜBKE Monumental	65
		From I-Media-Cities to DARE Ravenna project	67
		<b>COMPUTING PROVISIONING SERVICES</b>	
		<b>INDUSTRIES</b>	
		Supporting the National Industry through HPC-POC	71
		EuroCC bio	72
		Training for industries	73
		Supporting ENI HPC infrastructure	74
		<b>FUTURE TRENDS</b>	
		What's next: evolution of the computing infrastructure	77
	<b>2020 IN NUMBERS</b>		
2	HPC Infrastructures		
4	Users' statistics		
10	Resource allocation: projects, agreements, collaborations		
11	PRACE		
13	Iscra		
14	Agreements		
15	Training HPC		
18	Staff		
19	Events		
25	HPC H-Index		
	<b>SCIENTIFIC OVERVIEW</b>		
	<b>LIFE SCIENCES</b>		
28	HPC and Life Sciences		
29	Orchestra		
31	NIG		
32	GENOMED4ALL		
33	Exscalate4Cov		
35	LIGATE		
36	<b>EARTH AND ENVIRONMENTAL SCIENCES</b>		
38	The operational forecast in Italy		

# 2020 IN NUMBERS

Cineca HPC is the main Center for scientific computing in Italy. It runs a large computing infrastructure and make it available to Italian and European Researchers, as well as to important Italian Companies within a programme for supporting national industrial competitiveness.

Here we present data about usage and users, projects, events and classes offered in the field of High Performance Computing.

At last of this section we present our staff, the true resource that makes everything work.

# HPC infrastructures

Elda Rossi  
Cineca

In year 2020, our main HPC system (Marconi) went under a major upgrade. The largest partition (the so-called Marconi-A2 aka KNL) was substituted by a new cluster engineered by IBM and based on Power 9 processors and NVIDIA GPU accelerators. This new partition (known as Marconi100) is now part of our Tier-0 system that started its life in mid-2016, with the set-up of the first partition (A1) based on Broadwell chips and a computational power of 2PFlops peak. After several upgrades and inclusions, at the end of 2020 it is configured in two partitions: A3, made of 3.188 SkyLakes nodes; and M100, an accelerated partition made of 980 nodes (IBM Power 9 + Volta100 Nvidia GPUs), with a total peak performance in excess of 40PFlops. Marconi100 was ranked for the TOP500 list in June 2020 and reached position no.

9 with a sustained performance (Rmax) of 21.640 TFlops. A high-performance Lenovo GSS storage subsystem, that integrates the IBM Spectrum Scale™ (GPFS) file system, provides about 10 PByte net data storage capacity to Marconi-A3. A different storage system based on GPFS as well provides about 5 PByte net storage capacity to Marconi100. In addition to that, a common storage of about 20 PByte is available as Long Term Storage and Archiving of user data. The other important HPC system Galileo remained unchanged. It is made of 1022 nodes (Intel Broadwell) connected with an Intel Omnipath, 100Gb/s internal network. Sixty nodes are equipped with nVidia K80 GPUs. This is the Tier-1 system, available for the Italian research community.

## Tier-0

	CPU	Total nodes/ cores	Memory/ node
MARCONI 100 (IBM)	IBM Power AC922 + NVIDIA Volta 100	980 nodes 31.360 cores, 4x980 GPUs	256 GB
MARCONI A3 – SKL (LENOVO)	Intel SkyLake 2x Intel Xeon 8160 @2.1GHz 24 cores each	3.216 nodes 154.368 cores	192 GB

## Tier-1

	CPU	Total nodes/ cores	Memory/ node
GALILEO NextScale OmniPath 100Gb/s	Lenovo Intel Broadwell 2 x Xeon E5-2697 @2.1GHz 36 cores each	1022 nodes 36.792 cores	128 GB

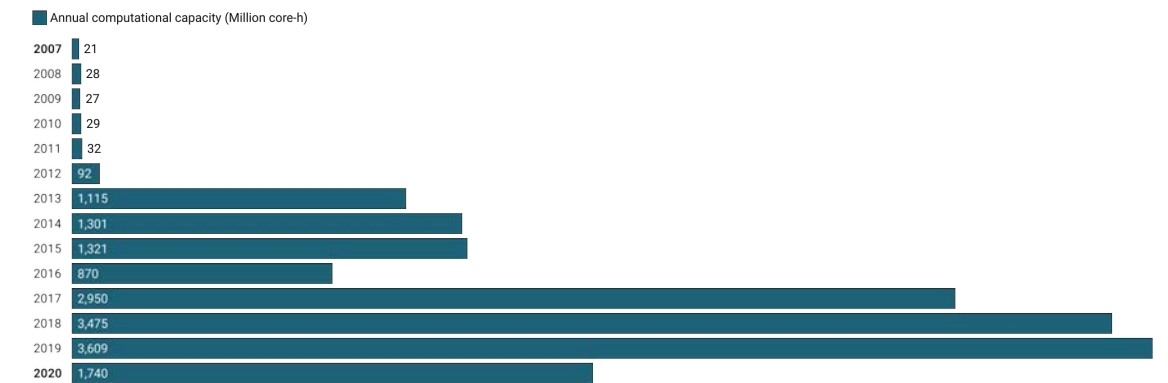
In addition to aforementioned systems, other devices were present in the Cineca HPC infrastructure in 2020. They were devoted to special functions or dedicated to specific users.

**MEUCCI:** Part of this cluster is dedicated to weather/climate service operations, part is running the “HPC Cloud” service. It is made of 288 nodes (Broadwell Intel Xeon E5-2697) with 36 cores each, connected with a Mellanox network.

**GATEWAY:** 28 nodes (Skylake Intel Xeon 8160) with 28 cores each, connected with a Omni-Path network. It is a dedicated interactive cluster for EUROfusion

**WLCG:** 288 Broadwell nodes (Intel Xeon E5-2697) connected with a custom internal network and linked with a high performance link with the Bologna INFN computing room. It is a dedicated cluster for INFN.

The computational capacity of these systems is usually expressed in terms of core-h, the number of computing hours that can be exploited for each single core of the system. In the picture below the total annual capacity is reported for each production system, for a grand total of 1.7 billion core-h.



Galileo	133 M core-h
Marconi A3	1'332 M core-h
Marconi 100	275 M core-h
<b>Total</b>	<b>1'740 M core-h</b>

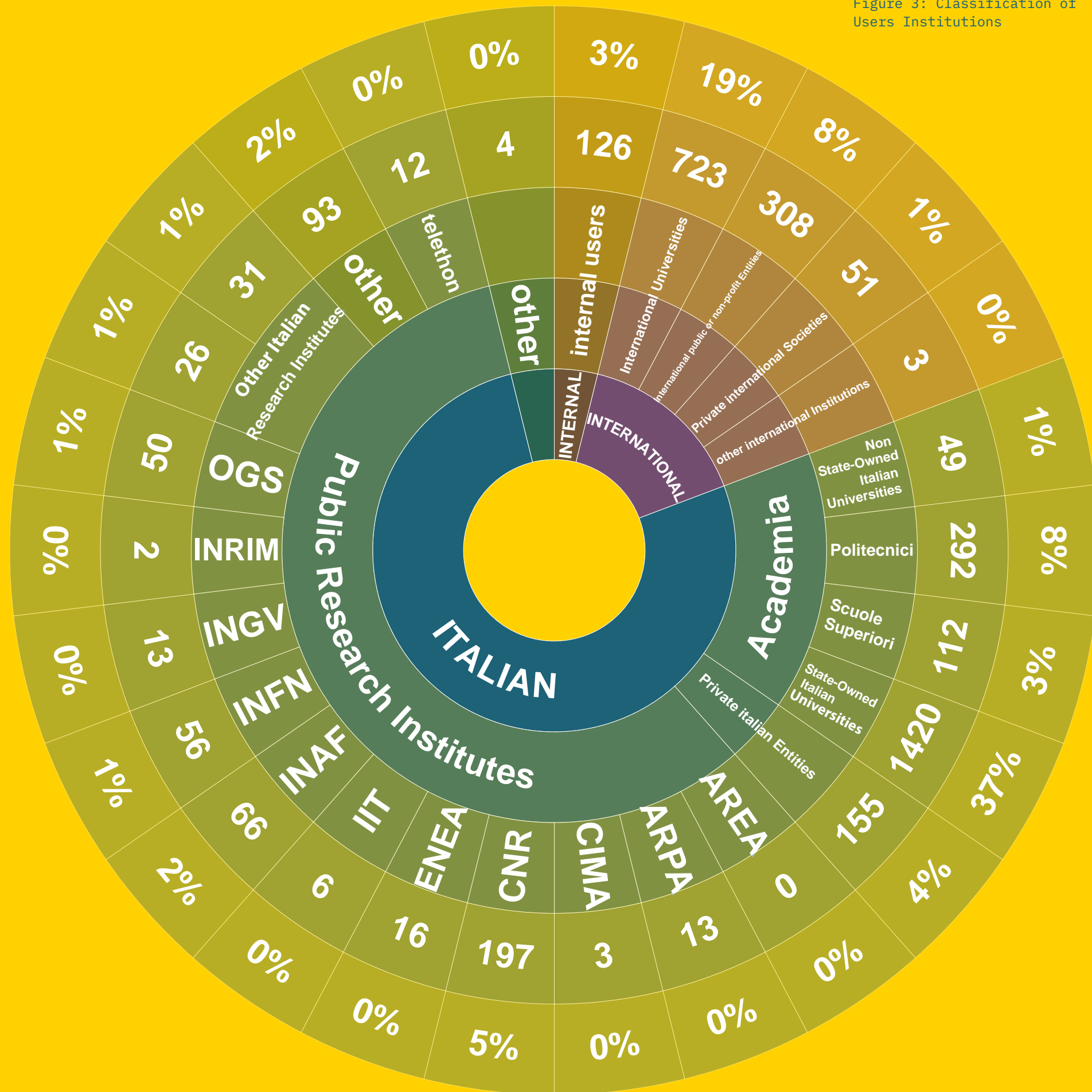
The total computational capacity is drastically reduced with respect of the previous year (3.6 to 1.7 billion core-h). This is due to the replacement of Marconi-A2 with Marconi100 cluster, made of 1/3 of nodes, each node containing half number of cores. This results in a dramatically lower number of core-h available yearly (1'800 to 275 million core-h). However, it is important to note that the presence of the GPUs in Marconi100 implies an important increase of the computational power in terms of PFlops, that is the number of operations the system is able to perform in one second (from 11 to 30 PFlops).

# Users' Statistics

Elda Rossi  
Cineca

At the end of year 2020 the active (personal) users on our HPC systems were 3.827. The great majority of the users are males (78%), aged 31 to 45 (46%), and working for Italian institutions (73%), mainly Universities or other public research bodies in Italy (38%) or abroad (27%). The large Italian cities, clustering multiple institutions, are well represented: the Milan area (495 users), followed by Rome (373), Trieste (330), Bologna (320) and Turin (168). Among the more represented foreign nationalities: Germany United Kingdom and France (thanks to the EUROfusion community), Spain, Switzerland and Neaderland. Our users represent scientists in all scientific disciplines that take advantage of computational resources, the most represented being: chemistry and material science (22%), Engineering and Computational Fluid Dynamics (23%), Astrophysics and Plasma physics (11%), Life science (11%).

Figure 3: Classification of Users Institutions





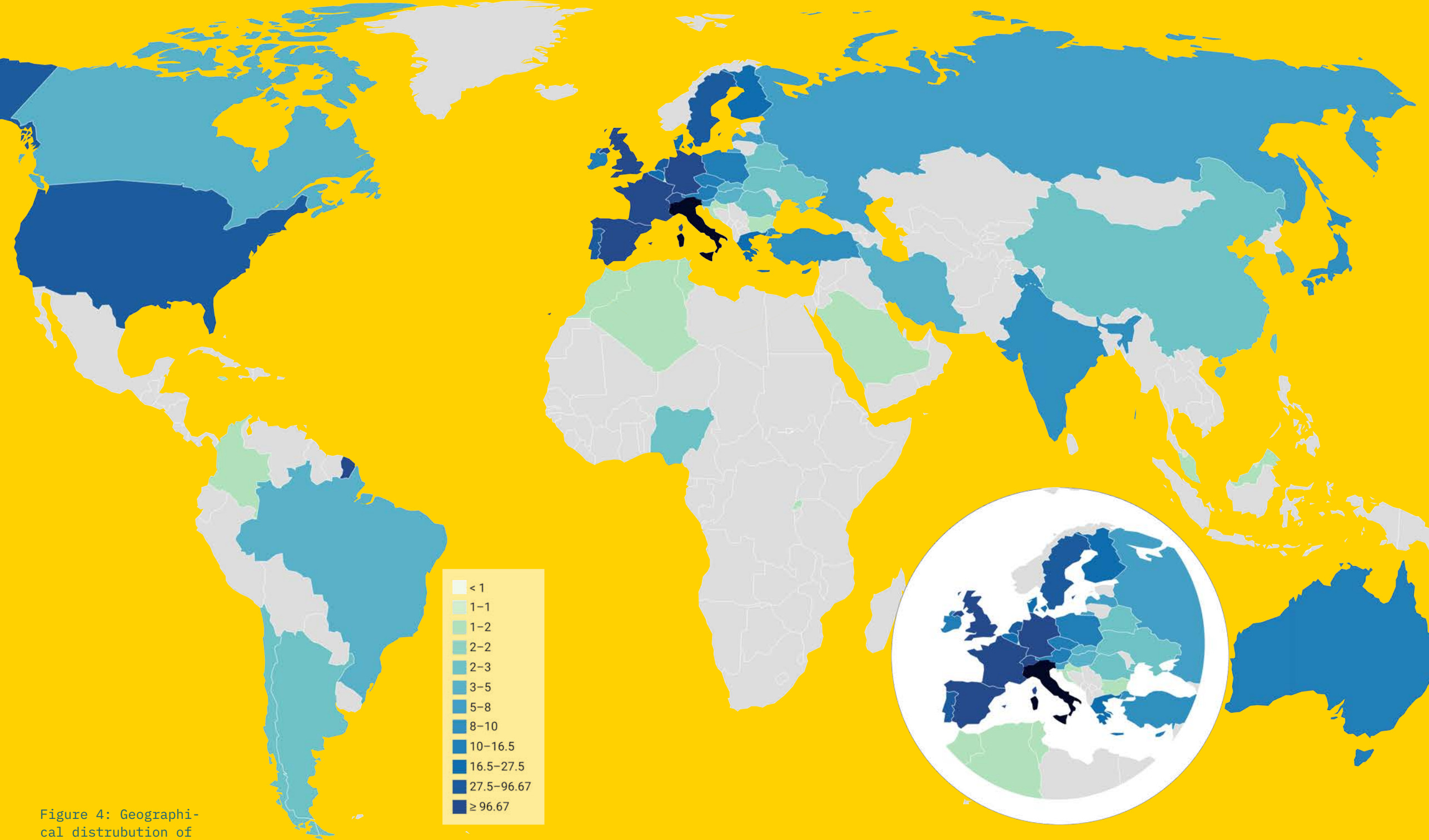


Figure 4: Geographical distribution of international users

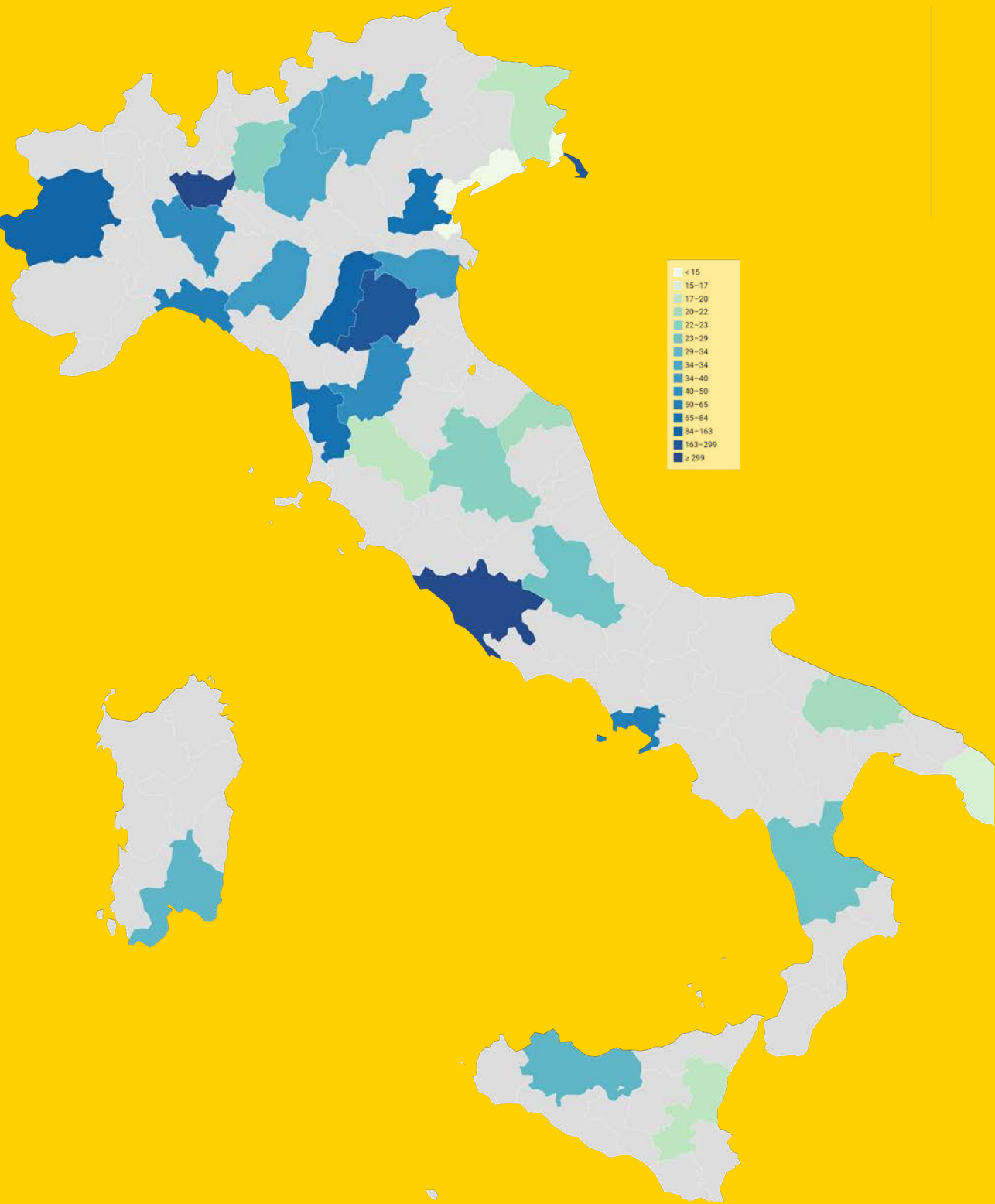
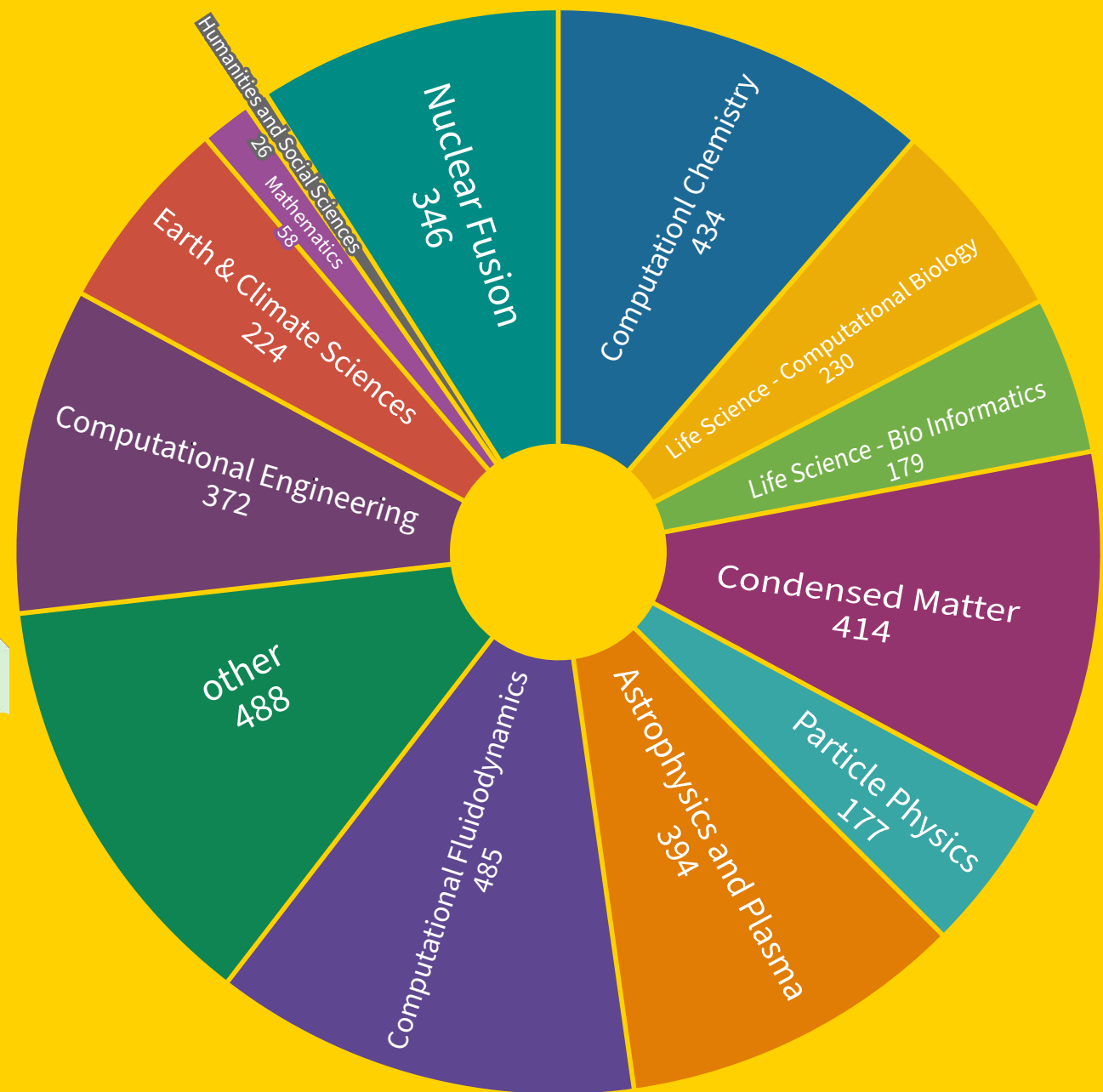


Figure 5: Geographical distribution of Italian Users

Figure 6: Distribution of Users by Disciplines



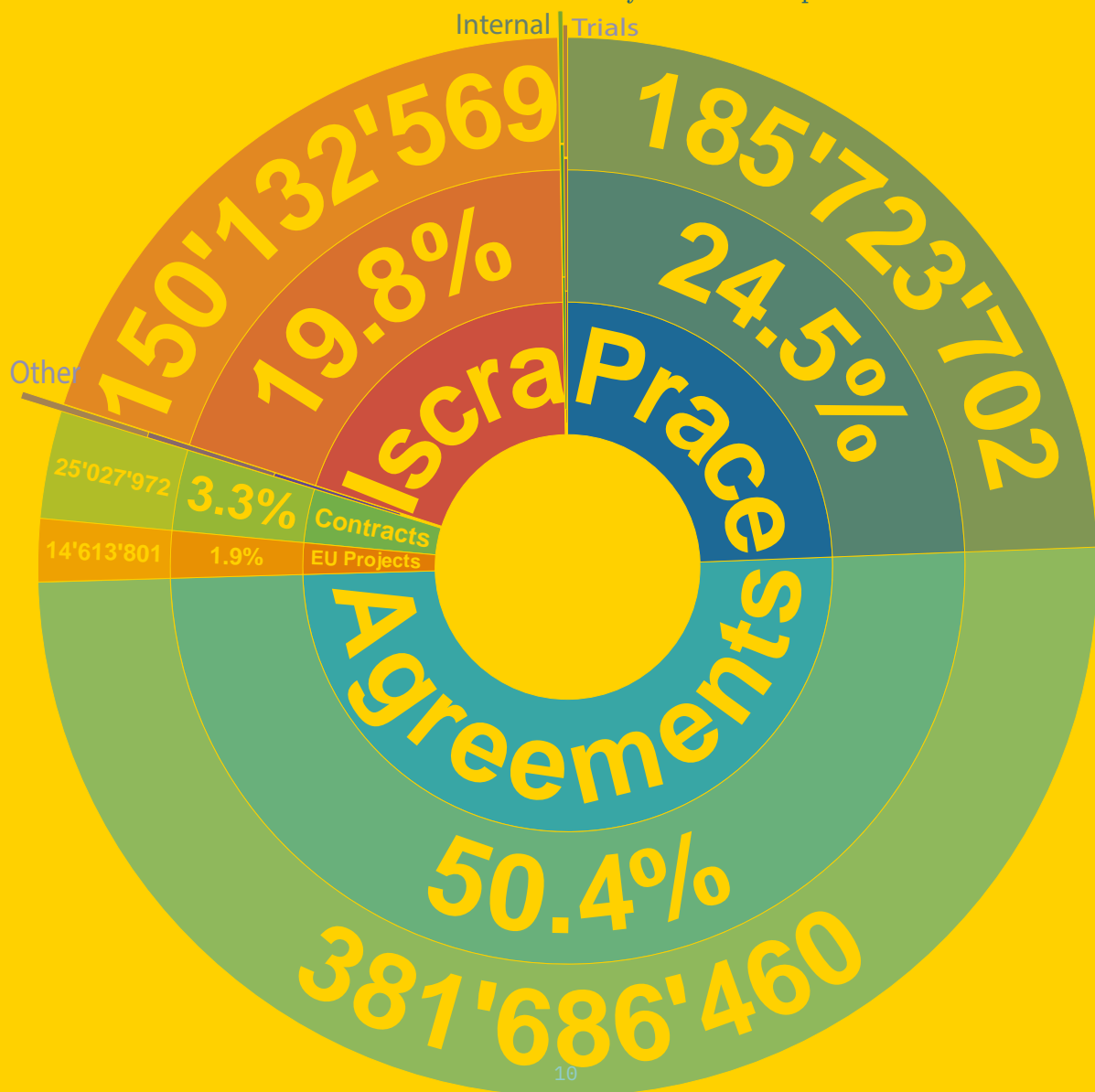
# Resource allocation: projects, agreements, collaborations

Elda Rossi  
Cineca

The HPC clusters in Cineca are used by Italian and European researchers for running their computational research. The access to the HPC facilities is based on “computational projects”. Some of them are based on peer-review mechanisms: PRACE is at European level; Iscra is for national researchers. Other projects, defined under the category “Agreements”, refers to collaborations with some major Italian research Institutions; “Industrial” represents activities for supporting the national industrial competitiveness. Each project has also a starting and ending date, a budget (in terms of core-h), a PI (Principal Investigator), several collaborators, and a science domain.

In addition to this, there are dedicated partitions on Marconi A2, Marconi A3 and Marconi100 (EUROfusion consortium), and on Galileo (for weather forecast activities and industrial projects). On these partitions the allocation is made autonomously by the dedicated committee. A grantotal of 761 million core-h were allocated on 758 million actually available, that is 100,5% allocation. Half of the resources were allocated to agreements, 25% to Prace, 20% to Iscra, 3% to industrial.

The numbers below represent the allocated resources, in terms of million core-hours, for the different type of projects, and refer only to the shared partitions of the clusters.



## PRACE

Partnership for Advanced Computing in Europe

M. Guarrasi, E. Rossi, G. Scipione

Cineca

PRACE (Partnership for Advanced Computing in Europe) gives computational resources and technical support to European researchers to enable new high-impact scientific discoveries. It provides HPC resources through Preparatory Access (for code scaling and optimization) and through Project Access (for production) calls. Currently PRACE has 26 state members, representing European Union Member States and Associated Countries. CINECA represents Italy in PRACE as a Hosting Member. This means that CINECA Marconi100 system hosts both Preparatory Access and Project Access calls from PRACE. Currently only Italy, France, Germany, Spain and Switzerland are Hosting Members of PRACE. This is a great advantage for Italian-based researchers, since 75% of the total computational resources of PRACE are reserved for applicants with an affiliation coming from a Hosting members country. Furthermore, thanks to the availability of 7 different Tier-0 machines inside PRACE Research Infrastructure, the Italian researchers can choose the architecture that fits better the requirements of their codes/projects.

During 2020 two new calls (20 and 21) were launched, awarding on Marconi100 20 projects and making available for PRACE more than 3.1 Million of NODE hours. Furthermore, in the same period, a special fast track for COVID-19 was launched. In this framework CINECA allocates on Marconi100 further 8 projects for about 250'000 NODE hours.

At the same timeframe PRACE gives 5% of their available resources also to European Center of Excellence (CoE) for benchmarking and developing activities.

Italian-based project represents a considerable part of the total budget allocated by PRACE in Europe. Indeed in 2020, considering the entire PRACE Tier-0 infrastructure, 18 projects on a grand total of 105 were allocated to Italian based PIs. Furthermore, considering also the COVID-19 fast track, in the same timeframe 25 projects on a total of 138 have an Italian-affiliated PI.

In addition to computing time, PRACE also supports the selected research projects with assigning them a High-Level Support Team (HLST), that helps to further utilize the capabilities of PRACE Tier-0 through code optimization. In 2020 Cineca was one of the major contributors of PRACE systems and for the HLST available in combination to Marconi100 system.

Besides their specialized user support, the HLSTs located at the 5 different hosting sites (GENCI for France, GCS for Germany, CINECA for Italy, BSC for Spain, CSCS for Switzerland) work together for the continued testing and improving of the libraries and applications which run on their Tier-0 systems, and for ensuring early transitions to new technologies. They provide dedicated application support to assigned projects through benchmarking, code optimization and scaling-out of applications.

The researchers of PRACE awarded projects and HLST members work together providing scientific expertise and HPC expertise respectively. This collaboration ranges from 1 to 6 months. It is possible to have collaboration for more extended period (from 6 to 12 months) for projects that require refactoring of codes.

Accepted	Call 20 20 April 2020 to 26 April 2021	Call 20 20 April 2020 to 26 April 2021
Regular call	11	9
Reserved to CoE	6	10
COVID-19 Fast Track	8	-
Total projects for call	25	19
Total accepted projects	44	
Marcni 100 Node Hours / Local Core Hours	1.742.832 / 55.770.639	1.758.553 / 56.272.727



## Iscra

Italian SuperComputing Resource Allocation

Paola Alberigo

Cineca

ISCRA (Italian SuperComputing Resource Allocation) gives computational resources and technical support to researchers affiliated to Italian institutions, in order to support the Italian research. It provides HPC resources through Class C projects (for code testing and pre-production) and Class B projects (full production). IscraC projects

are evaluated and selected on a continuous basis, IscraB projects twice a year. In 2020 Cineca provided to the Italian research community 102 Mcore-h on Galileo and Marconi100. In particular two IscraB calls were launched as reported in the table below (calls 21 and 22).

	Call 21B 29 April 2020 to 29 October, 2021	Call 22B 4 August 2020 to 4 February, 2022
Accepted	41	29
Rejected	36	18
Total number submitted projects	124	
Total accepted projects	70	

	Call 21B	Call 22B	Local Core Hours
GALILEO	8.116.000	4.664.480	12.780.480
MARCONI 100	56.822.610	31.915.700	88.738.310
Total			<b>101.518.790</b>



# Agreements

**Maurizio Cremonesi**  
Cineca

Cineca computing platforms may be accessed also by agreements for collaborations. Indeed several Universities and Research Institutions decided to support their researchers by defining agreements reserving a certain amount of computing and storage resources for their researchers and technicians. These agreements grant access to resources from one to a few years, without the need of submitting a project proposal, so they are particularly suited for special projects and small experimental activities. The research and academic institutions with active agreements in 2020 are listed below.

In total 416 Mcore-h were allocated to agreement projects, 50% on off all allocated resources. The largest budget corresponds to INFN that, together with SISSA, ICTP and University of Milano Bicocca, in 2018 signed a three years agreement for the exploitation of Marconi A3. INAF holds an important agreement (64 Mcore-h in 2020) thanks to a three years activity on Marconi A3. The institutions in the field of Life Science research all together sum up to a 6 Mcore-h grant. A similar amount corresponds to the other 8 Universities and Polytechnics.

During year 2020 more than 20 agreements with Universities and Research institutions were active.

Institution	Award resources (Mcore-h)
INFN (Istituto Nazionale Fisica Nucleare)	183,10
SISSA (Scuola Internazionale Superiore di Studi Avanzati)	87,00
INAF (Istituto Nazionale di Astrofisica)	55,92
UniBicocca (Università degli Studi di Milano Bicocca)	35,85
ICTP (international Centre for Theoretical Physics)	28,89
Università di Trieste	5,42
IIT (Istituto Italiano di Tecnologia)	4,80
Elixir (Distributed Infrastructure for Biological Data)	3,52
Università Milano	2,94
Università di Genova	2,18
Politecnico Milano	1,83
OGS (Istituto Nazionale di Oceanografia e Geofisica Sperimentale)	1,71
Università di Bologna	1,00
Istituto Ronzoni	0,52
Telethon	0,49
Università di Bergamo	0,39
Università di Brescia	0,35
INMI (Istituto Nazionale Malattie Infettive Spallanzani)	0,31
OspPBG (Ospedale Pediatrico Bambin Gesù)	0,30

# Training HPC

**Francesco Falciano**  
Cineca

The pandemic COVID-19 had a huge impact on all the classes for all the education programs worldwide. The social distancing was one of the main resolutions taken by all governments trying to prevent the spreading of infection, so for guaranteeing the continuity in education programs, schools had to develop alternative ways of teaching.

The solution was found in online teaching platforms.

Cineca fortunately was ready to this new way of teaching with Cisco WebEX conference solution, so we could start to convert the huge amount of HPC teaching classes in online classes. During 2020 the teaching platform switched to the more suitable platform Microsoft Teams.

The pandemic arrived unexpectedly and many syllabus were not appropriate for being converted as they were in online event. For this reason, some of the courses programmed in 2020 before COVID 19 emergency had to be rescheduled or cancelled, and some teachers made their best to convert lessons in something feasible for being taught online.

This new way of teaching increased our competence both in terms of organization and in e-learning skills, and students expressed their satisfaction.

Many comments in the surveys expressed their likelihood in following the course being at home, not having to plan long travels and with the possibility to join the class at the very last minute.

This new way of teaching made the organization able to add to the classes some auditors, students with no rights to interact with teachers in the practical sessions.

Furthermore the new way of teaching allowed easily the recording of the sessions. They were published on the courses web site, building a rich library of teaching videos, many of them held in English language, available for the people worldwide.

During 2020 we reached these numbers: 12 online courses, and 4 schools. More than 50 people in the department were involved in

teaching, for a total of 56 days of lessons and highly specialised training. In total, more than 510 Italian and European researchers took advantage of our training program, with high satisfaction levels (8,6/10), as shown by the surveys we regularly submit.

In addition to the Cineca HPC Department classes, our expert were involved in partnership with external societies or Universities, giving their support in organizing or teaching 6 events (Masters and Doctorate schools, workshops, tutorials) on HPC, Digital Humanities and Big data topics. Other important initiatives were: Summer of HPC (a PRACE initiative offering young students from all over Europe to work with HPC teams); European Researchers Night (Cineca is the coordinator of Society, a European project organizing the Night in Bologna and surroundings in 2020 and 2021); HPC-Europa3 (we are the coordinator of the project offering transnational access to researchers visiting HPC centers).

The Training activity is ruled by two specific procedures that refers to the ISO standard: ISO 27001:2013 "Information Technology - Security techniques - Information security management system".

For 2021 Syllabus, we decided to plan events until July 2021 converting the 2021 YEAR program in ACADEMIC YEAR (2021-2022).

This for two reasons: for synchronizing HPC Department program with PRACE PTCs that is programmed in ACADEMIC YEARS instead of YEARS, and, because in July we can see how pandemic is going and, in case of positive evolution, trying to reschedule some in class-events.

For the period January-July 2021 we scheduled over 15 Courses and School.

Our effort is constantly focused in increasing the number of courses and Schools developed by Cineca HPC Department, teaching the newest technologies available and trying to admit the greatest number of students worldwide.

For that reason, we will try to admit the most numbers of students giving them access to



510



Italian and European Students

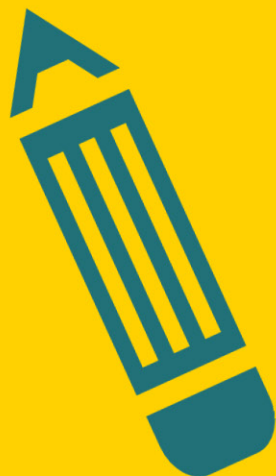
56

Days of lectures



50

Teachers



12

Classes

6

Collaboration in events

4

Schools



tutors' support, and for each course a reasonable number of auditors giving them access at least to an email support service.

For some courses we will open online forums so the students can discuss and support each other with the tutoring of our experts.

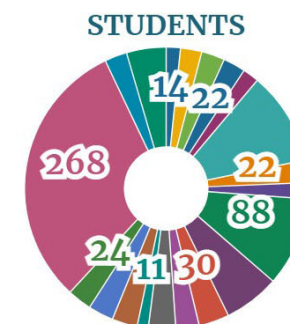
We will try to record and put online most of the lessons, tentatively in English language, making them feasible for the greatest number of students worldwide.

For some courses, we are working to implement a system of auto-valuation quizzes for the basic skills of the contents. Doing this quiz, a student can understand if he has the pre-requisites and leave in case his spot to a colleague left out that has a better background.

With the frequency certificates, we would like to release Open Badges, so we are working with colleagues in other departments to provide it for all the HPC Cineca Courses.

Of course, in 2021 we will work in PRACE project guaranteeing a new Syllabus and collaborating with the new Castiel project and, reasonably with new projects to come.

We are the reference for all the Training in HPC so we constantly receive calls and emails worldwide asking for collaboration and in the supply of new training events.



AVERAGE STUDENTS' SATISFACTION



- Introduction to Python
- Python for Scientific Programming
- Online Summer School on Parallel Computing
- Advanced School on Computer Graphics for Cultural Heritage
- Containerization in HPC
- Introduction to Fortran
- Introduction to Deep Learning and Tensorflows
- New accelerated partition of Marconi
- Programming paradigms for GPU devices
- High Performance Molecular Dynamics

Figure 7-8, numbers in training

# STAFF

In the HPC department of Cineca work about 95 people (33 females and 62 males), distributed among the three premises: 73 in Bologna, 7 in Milan and 15 in Rome.

Several new colleagues were enrolled in 2020 and 2021: we wish them a very fruitful and interesting job.

**Gianfranco Abrusci**

I obtained my PhD in “Computational Biophysics” at the University of Trento. In my research, I studied the dynamics of several proteins embedded in membrane using computational techniques like Molecular Dynamics. In CINECA, I work as HPC specialist with a focus on profiling and optimization of scientific applications.

**Paola Arcuri**

I received a bachelor’s degree in Computer Science from the University of Calabria, where I also obtained a Master’s degree in Scientific Computing. In Cineca, I work within the HPC department as a software developer; in particular, I focus on software performance and optimization.

**Tiziana Bassi**

I got my Ph.D. in Physical Sciences and Astrophysics at the University of Palermo and University of Toulouse. My research has been focused on the spectral study of accretion and ejection in bright X-ray binary systems, with particular interest to understand the geometry, mechanisms and physical processes playing a role in Galactic Black Hole Binaries. I have gained significant experience in Data Analysis, interpretation and presentation of the results.

**Gabriele Boga**

I graduated in the Master’s Degree in Aerospace Engineering at the University of Bologna with a thesis in Computational Fluid Dynamics. During the

research project that started with my thesis I had the opportunity to deal with some research topics concerning turbulence and I developed a deep interest in the fluid dynamics field. At Cineca I work in the HPC Department and I am mainly focused on CFD topics.

**Michele Bottazzi**

I graduated in Atmospheric Physics and Meteorology at the University of Bologna. I also got a PhD in hydrology at the university of Trento, where I contributed to develop an open source hydrological model. During my academic career I gained a vast experience in the environmental modelling and data. Currently in Cineca I work in the Data Management group and operational chain of the COSMO weather model.

**Marco De Pascale**

I have 7 years’ experience in Astronomy both as software developer and software engineer. My experience ranges from data science and machine learning, acquired during PhD research, to full-stack control-software development and design for multi-millions Euro instruments, gained in the following Post Doc contracts.

**Diego Molinari**

I got a master degree in Astrophysics at the University of Rome La Sapienza and a PhD in Astronomy at the University of Bologna. In my research activity I worked in cosmology in particular in the data analysis of the Planck satellite, an ESA mission. My studies focused on the anomalies of the cosmic microwave background, on testing the laws of fundamental physics and the statistical properties of the Universe. In CINECA I work as HPC User Support to assist users to efficiently run their codes on our clusters.

**Andrea Piserchia**

I got my PhD in “Methods and Models for the Molecular Sciences” at Scuola Normale Superiore di Pisa and I got both my bachelor’s and master’s degree degree in Chemistry at University of Padua. My interests range in the field of Computational Chemistry with particular emphasis to Molecular Dynamics simulations. At CINECA I work in the HPC specialist group where we provide support for scientific applications.

# Events

## The European Researchers’ Night of Society project

**Simona Caraceni**  
**Cineca**

For several months, the consortium worked to organize a small event with the presence of people in Bologna, the main venue. The organization of an event with the presence of the public posed significant challenges. Because of the date shift, from late September to the 27th of November, we could not organize outdoor activities, as we usually do. The safety measurements imposed by the pandemic posed severe constraints to the kind of suitable venues. But at the same time, the Steering Committee formed a working group to research the different kinds of online platforms and explore available options. This activity was crucial because of the worsening of the pandemic situation in Italy, and the decision to undertake only an on-line ERN, cancelling the in presence event.

The large number of people who were engaged by our web and social channels, especially during the week before 27 November, who participated in our pre-event program and in the online ERN main event (about 4.500 unique registered visitors), are all indicators that the overall awareness campaign was very successful. The exceptional situation due to the pandemic has created many difficulties. For several months we lived with doubts about what could actually be done and how. For example, such uncertainty is a reason for limited involvement of stakeholders, who were less present in the pre-event program than in previous years’ projects. Despite these difficulties and thanks to the effort of partners, we were nevertheless able to carry out an ambitious program of about 40 pre-events, 50% face-to-face and 50% online, plus 51 online lessons for the Schools. Supported by an intense communication campaign, these preparatory events, which involved about 10.000 participants, were the key element, which served to spread awareness of the European Researchers’ Night to the public.

At the end of the Night, the platform counted more than 5.000 registered users. Commonly, one registered user granted access to a group of people (family members, or roommates). Based on information gathered on access to

the website and to the Sharevent platform, we estimate the overall participation could have reached 10000 people. Although all activities were in Italian, targeting a local audience, 4% of visitors came from abroad. Among the Italian participants, 29% were from Bologna area and the cities involved in the project (for instance Rimini, 3%, Cesena 3%, Forli 2%), but we had visitors from many other cities as well (Milan, 10%, Rome, 5%, Florence, 2%, Parma, 2%, Naples, 2%, and the remainder distributed in numerous other places in Italy to the 100% of the national public).

Researchers involved in the Night event in Bologna and the nearby University Campus cities last 27 September 2019 welcomed more than 9000 visitors with a wide offer of activities, capturing different audiences. Many activities were designed for kids and youngsters, offering hands-on activities and entertaining science displays and games. The increase of the percentages of interviewees who attended ERN for the first time in 2019 was 64% in Bologna and 65% in the other cities. These numbers, which are the most significant indicators of the effectiveness of the event are even better than in 2018, where they were respectively 61% and 56%.

The success of SOCIETY is also testified by the increasing interest from Institutions/ Schools/ Companies, the so-called third-parties. The list of third-parties/collaborators has doubled from 43 in 2018 to a huge 86 in 2019. SOCIETY has become a reference to which an increasing number of public and private subjects ask for collaboration, in order to give greater value and larger diffusion to their own initiatives.

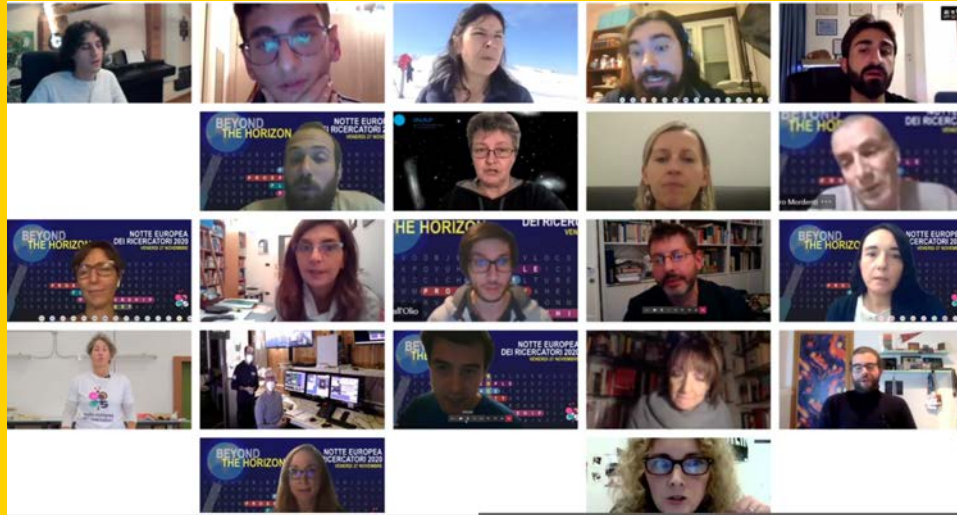
The rather ambitious program of about 30 pre-events distributed in the months preceding ERN has been widely successful, being attended by about 7300 people.

In Bologna most of the activities took place in via Zamboni, in the pedestrian academic district, right in the historical city centre. The inauguration ceremony took place in Palazzo Magnani, the stunning seat of Fondazione del Monte and Unicredit that host the important “Quadreria” and Carracci frescoes, where all the Night activities started at the presence of the most important personalities, students and researchers.

Almost 80 stands were distributed mostly in Scaravilli square and along the Zamboni street, where porticoes ensured protection in case of rain, that didn’t come gifting the event of a warm late summer starry night. Researchers were available at their stand throughout the entire evening (18:00-24:00), in most cases performing their activity on demand. In other



cases activities were organized in shifts and some required reservation. A limited number of activities were also held indoor, as the lucky one-to-one conversations (“A tu per tu”), held in Scuderie. Many high-school students were directly involved in outreach activities, introducing scientific contents to the general public, after long-term collaboration with scientists. Everybody could enjoy guided tours or simply wander through the many stands, exploring a broad spectrum of research disciplines, spanning from hard science to humanities, from outer space to human health, from applied technology to philosophy. For the entire duration of the Night, the stands were crowded with people making the 2019 edition a Night to remember.



## Events

### The European Researchers' Night of Net project

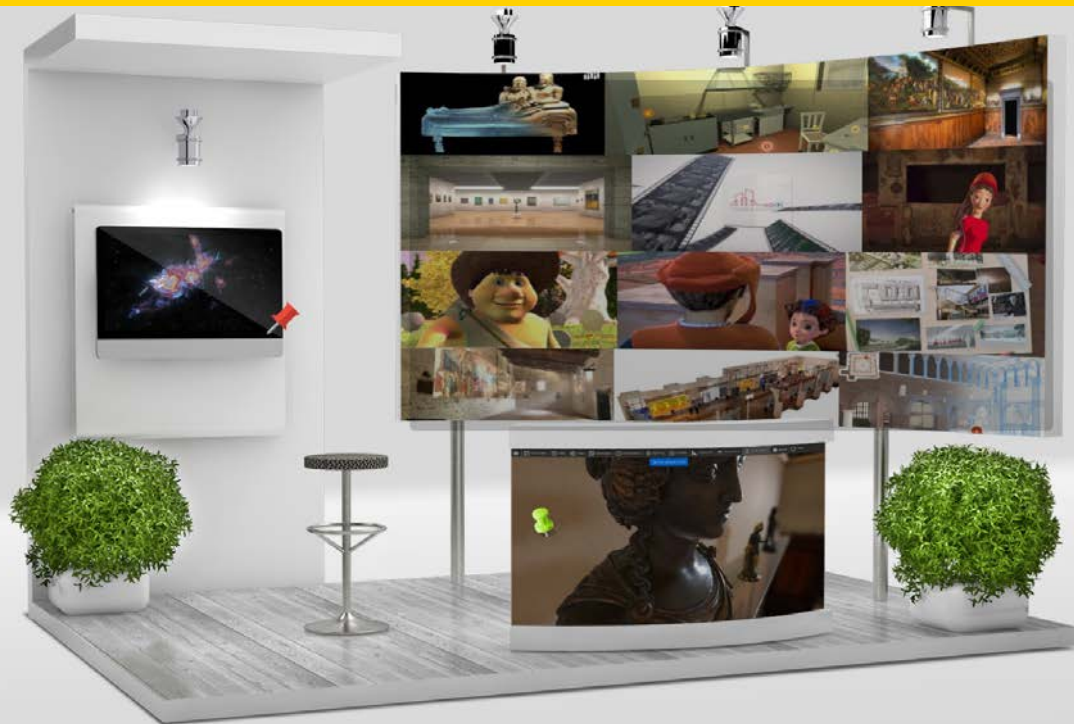
Neva Besker  
Cineca

The project NET- scieNcE Together organized its first edition of the European Researchers' Night, an initiative promoted by the European Commission since 2005 that involves thousands of researchers and research institutions every year in all European countries. In the 2020, due to the unpredictable evolution of the pandemia, the Night was organized on line at the end of November. The topic was the challenges given by climate changes and the goal was creating opportunities for researchers and citizens to meet and spread scientific culture and knowledge of the research professions in an informal and stimulating context. NET is a partnership of all the major local (Lazio) Research Institutes and Universities, such as CNR (coordinator of the project), CINECA, Ispra, INAF, INGV, ENEA, University of Rome Sapienza, University Tor Vergata, University of Tuscia and Telematic University UniNettuno. Cineca coordinated the events of the final European research Night as well as the activities with the schools. Some preevents were organized in presence in Rome and in many cities involved in NET project such as Livorno, Vulcano and Stromboli, La Spezia, Viterbo, Cerveteri, Tarquinia. The program of about 30 pre-events distributed in the months preceding ERN has been widely successful, being attended by about 1912 people in presence and 31687 people on line.

In order to allow the wider participation of the general public and the schools, offering an alternative to distance learning during the school hours, the ERN activities were organized over the entire week with the most significant events for the general public were organized for the final European Researchers Night, e.g. virtual tours and Raiplay (the web channel of the Italian public broadcasting company RAI - Radio Televisione Italiana), seminars and other activities for general public and families. A total estimated number of 4030 students were involved in all the events, together with their teachers. Researchers involved in the NET Night event welcomed almost 50000 on line visitors with a wide offer of activities, capturing different audiences.

Many activities were designed for kids and youngsters, offering games, escapes rooms, hands-on activities, contests, citizen science activities and talks with researchers. NET project made virtual tours of almost all institutions of the consortium in order to open virtually their scientific laboratories during the pandemia. It was a way to visit the environment of laboratories in which the researchers carry out their research daily and to permit to discover, even from home, the tools and technologies used. Everybody could enjoy also live meetings with researchers or simply wander through the many short documentaries and experiments videos made for the event, exploring a broad spectrum of research disciplines.

Figure 9,10,11: screenshots of the websites



### Visit lab Cineca - Laboratorio di Visual Information Technology

Benvenuti al nostro banchetto virtuale, in questo stand troverete alcuni dei nostri progetti e vari link per incontrarci dal vivo in questa piattaforma.

#### INCONTRA IL VISITLAB

- ✉ [visitlab@cineca.it](mailto:visitlab@cineca.it)
- 🌐 <http://visitlab.cineca.it/>
- 🐦 Twitter

#### RISORSE

informazioni generali

📄 Brochure Visit Lab





# Events

## SC20 Is Everywhere We Are

Daniela Galetti  
Cineca

The USA Supercomputing 2020 was held in virtual mode, as it was for the summer European edition ISC20.

Cineca attended, as usual, the exhibition also with a virtual booth, instead of the traditional one. The informations about Cineca projects and infrastructure were written in a sort of website specifically built for SC20, instead of printed materials in full bright colors on our temporary walls in Atlanta, used in standard conditions.

The number of virtual contacts was smaller than the number that usually we hit at the face-to-face conference. The mix of advanced scientific projects and italian styles is very appealing abroad, but it gives its best live.

On the other hand, we could transform this limitation to an opportunity to increase the number of people who attended the conference, thanks to the budget saved from the impossibility to travel.

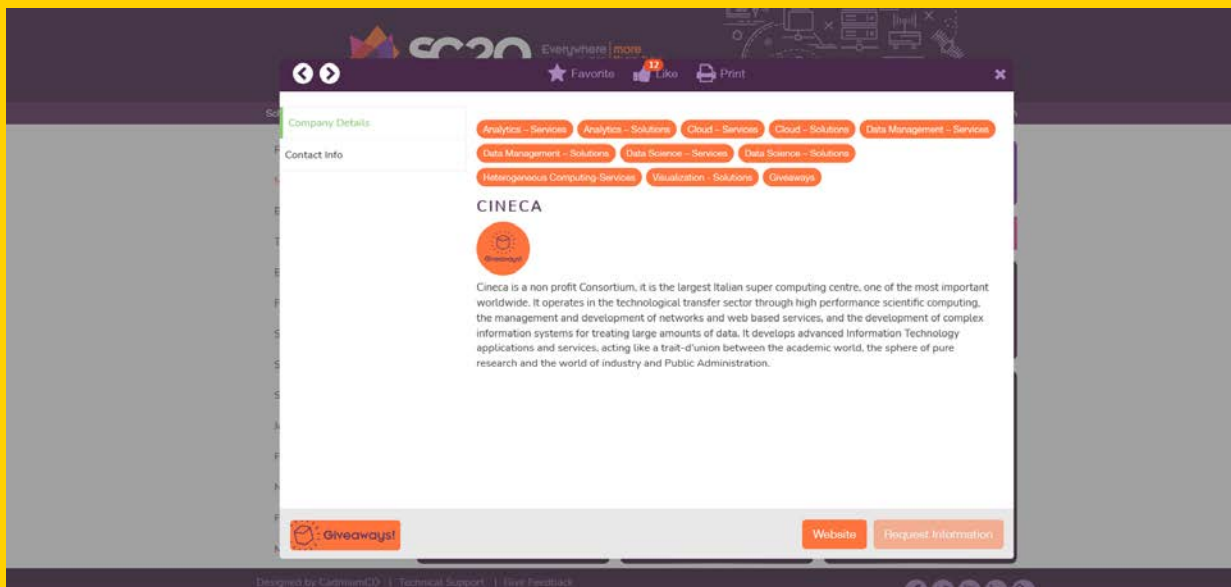
The 2020 motto was #MoreThanHPC, and in fact the Artificial Intelligence, machine learning, cloud items were increasing their presence not only in BoF (Birth of Feathers) presentation, but also in workshops and tutorials as well.

We were able to save the experience of single client meetings with vendors that allow the sharing of information about the roadmap of HPC hardware and software via video calls, but the participation to the exhibition in a visual way has limited sometimes the possibility to view and appreciate the products, especially if they are hardware ones.

The added value to go around and meet a mix of people (colleagues of other HPC sites, technical experts, students and company's vice presidents) sharing personal impressions and feelings about the HPC trends and exchanging opinions and technical experiences is unfortunately strong limited in virtual mode.

We like computers, nay, Supercomputer, but we are still humans!

Figure 12: a screenshot of the website



# Events

## Summer of HPC 2020

Massimiliano Guarrasi  
Cineca

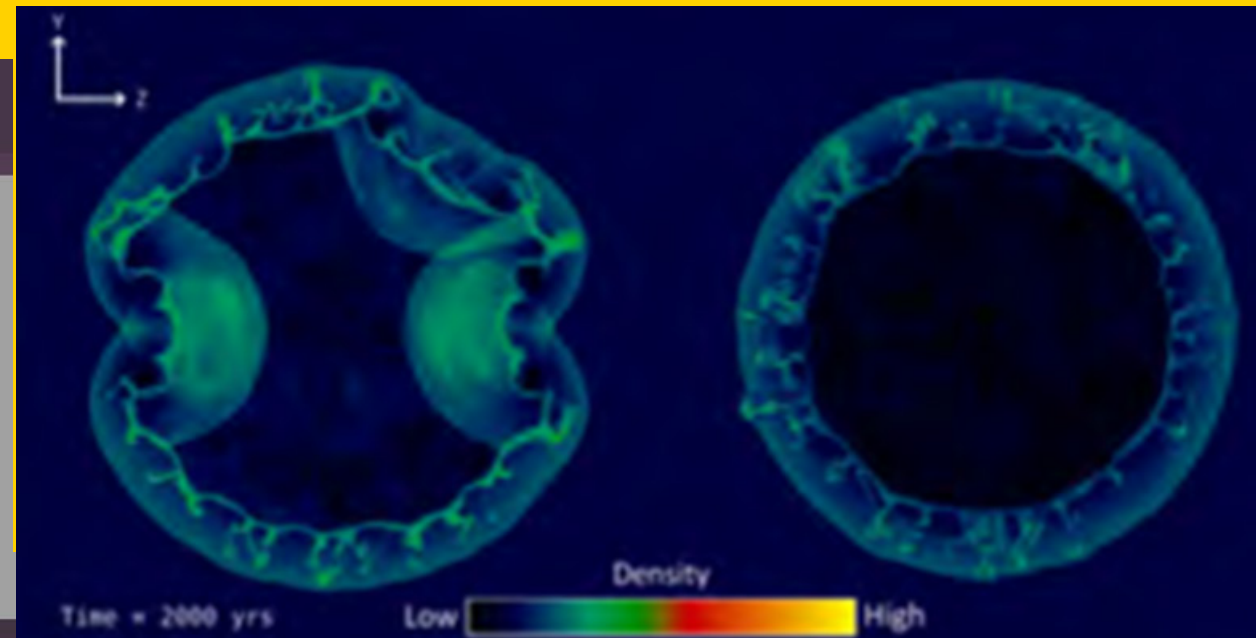
Summer of HPC is a PRACE programme that offers summer placements at HPC centres across Europe to late-stage undergraduates and/or master students. Participants will spend two months working on projects related to PRACE technical or industrial work. CINECA usually host two students per year in the framework of this program. During 2020, due to the COVID-19 pandemic, the internship was done remotely.

Particularly we hosted 5 students (Seán McEntee, Cathal Maguire, Aisling Paterson, Stefan Popov and Nathan Byford), that worked on two different projects:

- Anomaly detection of system failures on HPC accelerated machines using Machine Learning Techniques (Aisling, Stefan and Nathan)
- Visualization of supernova explosions in a magnetised inhomogeneous ambient environment (Seán and Cathal).

The five students spent July and August 2020 working remotely on their projects and obtaining very interesting results under the guidance of their mentors (Andrea Bartolini from University of Bologna and Salvatore Orland from INAF-OAPA) and site coordinator (Massimiliano Guarrasi from CINECA).

Figure 13: A 2D slice at the centre of Model B's 3D density profile (left), alongside a 2D slice at the centre of the symmetric model's 3D density profile (right), both at  $t = 2000$  years



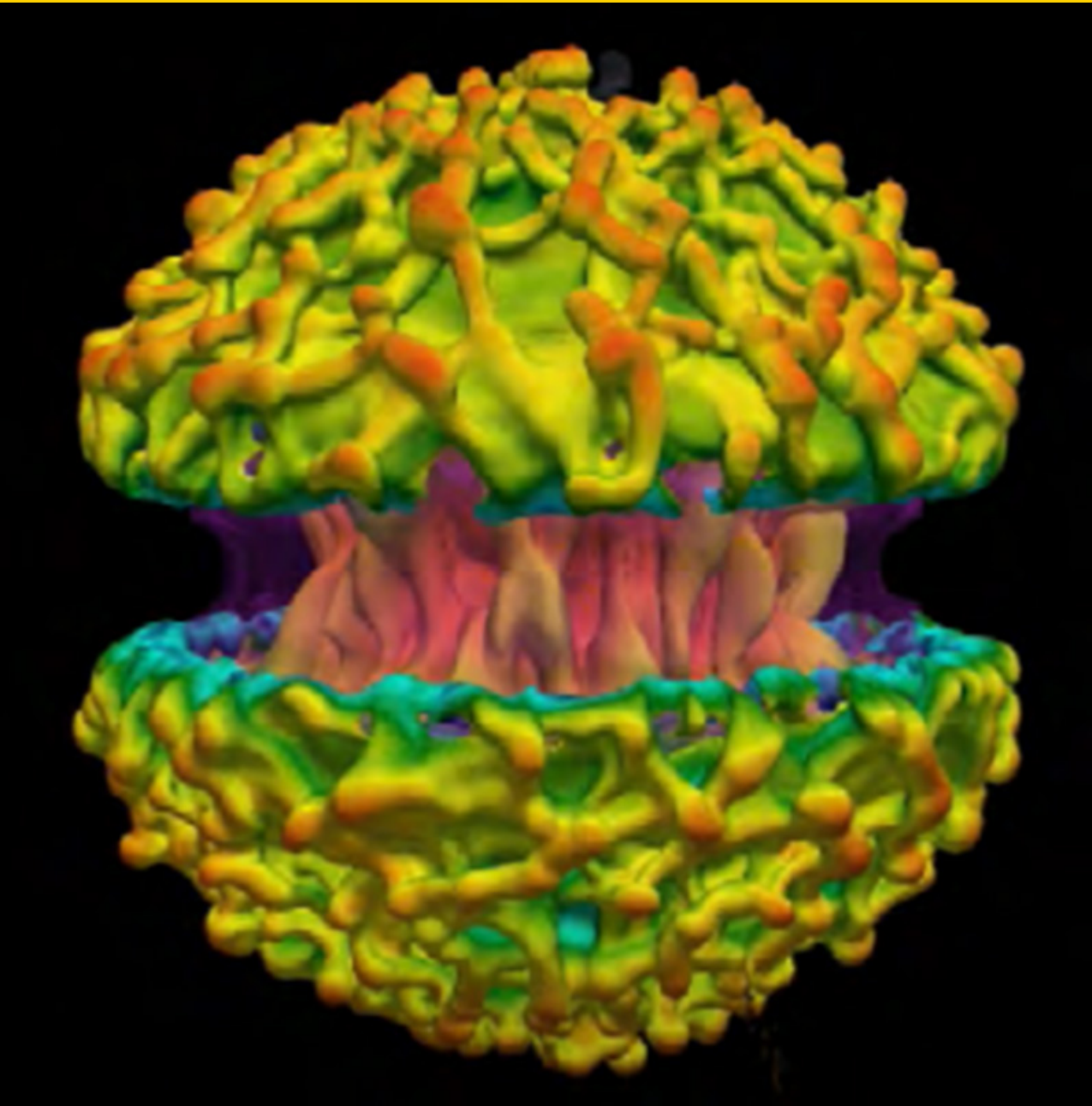


Figure 14: 3D visualisation of the density profile of Model B at a time of 2,000 years after SN explosion.

# HPC h-Index

Source; Web of Science Database: Web of Science Core Collection

Search: (FT=Cineca or FO=Cineca)

Time span: 2020-2021

Date: 23.07.2021



**Results found:**  
**858**



**Sum of the Times Cited:**  
**2662**



**Average Citations per Item:**  
**3,1**



**h-index:**  
**16**





# SCIENTIFIC OVERVIEW

Cineca is involved in a wide scientific field thanks to EU projects and collaborations. From artificial intelligence to urgent computing, bioinformatics and digital humanities, HPC is used to support them all.

# LIFE SCIENCES

## HPC and Life Sciences

Andrew Emerson, Maurizio Ortali, Elisa Rossi, Gabriella Scipione  
Cineca

Biomedical research, mainly through clinical trials, drug development and genomics data analytics, routinely collects, stores, analyzes, and shares massive amounts of data.

Thus, Life Sciences rely more and more on data analytics enabled by High Performance Computing (HPC), where both high compute and fast storage capabilities are immensely important.

Advancements in HPC are starting to reduce time to scientific discovery and generating the predictions necessary to make precision medicine a reality.

From our perspective, the “data deluge” in Healthcare is continuing to grow exponentially in recent years, by volume and heterogeneity, and it may be foreseen that disruptive technologies or breakthrough advancements will be needed in HPC arena, such as Quantum Computing, to actually be able to cater with the upcoming healthcare and genomics-driven research needs.

Life Science is evolving towards the need for fast storage and networking to handle the influx of real-time data from research, experimentation and collaborative organizations.

This includes: the processing of massive datasets for deep/machine learning, also considering the use of novel techniques such as Federated Learning; the optimization of modelling, such as a CADD (Computer Aided Drug Design) HPC-enabled infrastructures to accelerate the drug design process; the need for enhanced security and protection of healthcare data in HPC settings, in compliance with European GDPR regulations.

COVID research has even accelerated this process, and in fact Italy is leading some of the European reference projects in fighting COVID-19, such as the Exscalate4CoV for

the reuse of known and synthetic active ingredients in the preparation of new drugs to contrast therapy against COVID, as well as the Orchestra project, for the management and monitoring of contagion clusters in order to provide an innovative approach to learn from the SARS-CoV-2 crisis and derive recommendations to be better prepared in case of future waves and pandemics.

In this context, Cineca is being active and engaged in several fields, supporting the Life Science scientific communities, in different dimensions across several projects:

### **ORCHESTRA H2020 Project:**

CINECA leads the Work Package related to Data Management, in close collaboration with other HPC European Centers (HRLS in Germany and CINES in France) and experts in Healthcare Data Interoperability (Charité Hospital in Berlin) with the objective to set-up a common Data Architecture to allow data sharing among the largest Covid patient cohorts across Europe

**NIG**, the collaboration with the Network for Italian Genomes continues, with a new solution performed by CINECA to process and store omics data in accordance with GDPR. The aim is to create a database containing data from nucleic acids sequencing of Italian subjects and detect allelic variant frequencies

**Genomed4All** H2020 Project, in which CINECA will contribute to the set-up of a data sharing platform based on novel Federated Learning schemes, to advance research in personalised medicine in haematological diseases

**Exscalate4Cov** H2020 Project: CINECA is involved in coordinating the activities related to the largest virtual screening experiment ever performed (performing more than 1 trillion simulations in one single run), in close collaboration with other HPC facilities in Europe, namely

HPC5 (Eni), MARCONI-100 (CINECA), Barcelona Supercomputing Center and FZ Jülich.

**LIGATE EUROHPC** project: main goal of CINECA in LIGATE is to validate and deploy the CADD solution on its systems, including the Leonardo system, and make it available to academic and industrial users

**ARNO Observatory:** created at CINECA in 1987, it is an Observatory of real world data to support governance activities in the healthcare (expenditure control and appropriateness of drug prescription) of the individual LHUs, in 2016 was selected by the European Commission as one of the 20 examples of the use of Big Data in Healthcare in Europe

Cineca is also involved in scientific research offering its support, computational resources and storage to ELIXIR, TELETHON and AIRC.



## ORCHESTRA

Connecting European Cohorts to Increase Common and Effective Response to SARS-CoV-2 Pandemic

**Mattia D'Antonio, Chiara Dellacasa, Maurizio Ortali, Elisa Rossi, Gabriella Scipione**  
Cineca

The Connecting European COVID-19 Cohorts to Increase Common and Effective Response to SARS-CoV-2 Pandemic (ORCHESTRA) consortium brings together key European academic and research institution experts in infectious diseases and Data Management and High Performance Computing (HPC) from 26 organizations, in order to provide an innovative approach to learn from the SARS-CoV-2 crisis and derive recommendations to be better prepared in case of future waves and pandemics.

The ORCHESTRA project aims to respond to this need through the creation of a new pan-European cohort built on existing and new large-scale population cohorts in European and non-European countries.

Cineca in the project is the coordinator of the “Data Management” work package, in charge of the development of the platform that will collect data from the various cohorts involved.

Together with Cines and HLRS, HPC centers of respectively France and Germany, Cineca will be responsible to design and deploy a federated architecture based on three layers: National Data Providers, National Hubs (NH) and the centralized Orchestra

Portal. Moreover, in WP7, Charité of Berlin is in charge to identify and define common data standards to improve health data interoperability and enable data processing across cohorts and institutions and to ensure that all uses of personal/sensitive data comply with the EU GDPR, and with any additional derogations imposed by Member States.

The project started in December 2020 and in the first months of the project, most of the WP7 efforts have been devoted to gather the requirements from the different clinical settings across the clinical workpackages and to design a preliminary sustainable and privacy-compliant data architecture for the different cohorts, among the different HPC Centers.

In coordination with the European Commission, the ORCHESTRA team will consult with the European Centre for Disease Prevention and Control (ECDC) and the European Medicines Agency (EMA), in particular when it comes to making available data in real time that can be of value for shaping the continuously evolving public health and vaccine strategies.



ORCHESTRA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101016167.



# NIG

Network for Italian Genomes

Mattia D'Antonio, Silvia Gioiosa, Elisa Rossi

Cineca

Since 2016 CINECA supports the scientific research of the Network for Italian Genomes (NIG). The aim of NIG is to define an Italian Reference Genome for the identification of:

- genes responsible for genetic diseases and susceptibility genes for complex diseases in both basic and translational researches;
- genetic variants responsible for interindividual differences in drug response in the Italian population, and useful for population and forensic purposes;
- new target for diagnosis and treatment of genetic diseases.

CINECA is involved in the creation of a centralized repository for human genomic variant data produced by massive parallel sequencing, by the aggregation of genomic data from different Centers (Hospitals or University) in Italy, to obtain the typical genome of the Italian population to be used as a reference.

The introduction of General Data Protection

Regulation (GDPR) has lead CINECA to develop a new solution to manage personal and sensitive data, in particular genomic data (Next Generation Sequencing Exome and Genome Sequencing). For NIG, the main challenge involved sensitive data from different Institution which had to be managed and analyzed in compliance with current law.

The architectural solution is based on a Virtual Machine with encrypted volume and the analysis is executed with a workflow manager system.

This solution allows the creation of a shared database (NIG-db) containing data from nucleic acids sequencing of Italian subjects. NIG-db will only publish aggregated variant frequencies from a pool of data provided by the submitters.

# GENOMED4ALL

Genomics and Personalized Medicine for all through Artificial Intelligence in Ematological Diseases

Maurizio Ortali, Elisa Rossi

Cineca

GenoMed4All - Genomics and Personalised Medicine for all through Artificial Intelligence in Haematological Diseases - is a new ambitious European project, recently selected and granted by the European Commission under the Horizon 2020 Research & Innovation programme.

GENOMED4ALL will support the pooling of genomic, clinical data and other “-omics” health data (data EHR, PET, MRI and CT , Next Generation Sequencing, Microarray, Genome Wide Association, Copy Number Variations, DNA sequencing, RNA sequencing, including single cell, etc.) through a secure and privacy respectful data sharing platform based on the novel Federated Learning scheme, to advance research in personalised medicine in haematological diseases thanks to advanced novel AI models and standardized sharing of cross-border data. In the current data economy, artificial intelligence relies on creating sophisticated models based on existing data and using the trained models to predict an output when new data is encountered. However, there are cases as in clinical research where data cannot be freely exchanged or uploaded to the same machine where the model is to be created and trained. Federated Learning (FL) techniques allow researchers to develop sophisticated models using data that is distributed (i.e. stored on different platforms in clinical hospitals across Europe), and never exchanged. In this sense, Federated Learning is designed to respect data privacy and security, while still allowing users to carry out sophisticated modelling.

GENOMED4ALL will demonstrate the potential and benefits of trustable and explainable AI technologies, with a novel approach to AI models and algorithms using AI advanced deep learning, variational

autoencoders, generative models, besides combining with advanced statistical and Machine learning processes approaches to exploit the powerful set of “-omics” data which will be at researchers’ disposal.

The project has started in January 2021, will span 4 years and mobilize a large consortium of 23 key partners from Spain, Italy, Germany, France, Cyprus, Greece and Denmark, covering the whole value chain of clinical, regulatory and ethics research, academia, healthcare, disruptive tech and digital service providers. Most of clinical centers are already part of an European consolidated research network of excellence centers in hematological diseases (EuroBloodNet).

In the GENOMED4ALL project, Cineca will contribute with its own reference HPC infrastructures supporting the project and with expertise in the field, in particular CINECA will liaise the Consortium to the European Open Science Cloud initiative. First discussions related to the design of the technical Architecture of the project have just started. For the GENOMED4ALL project, CINECA is interested in continuing the extension and the reach of his own shared HPC infrastructures, supporting the Clinical research community. The set-up and scale-up of a dedicated AI European Infrastructure will broaden the user base of computing resources and will allow to develop and validate AI models in the field for the migration from project oriented to service-oriented compute resource provisioning. The experience gained with the definition of the GENOMED4ALL architectural design will serve as a potential input for the characterisation of new services to be exploited by the scientific and industrial community at large.

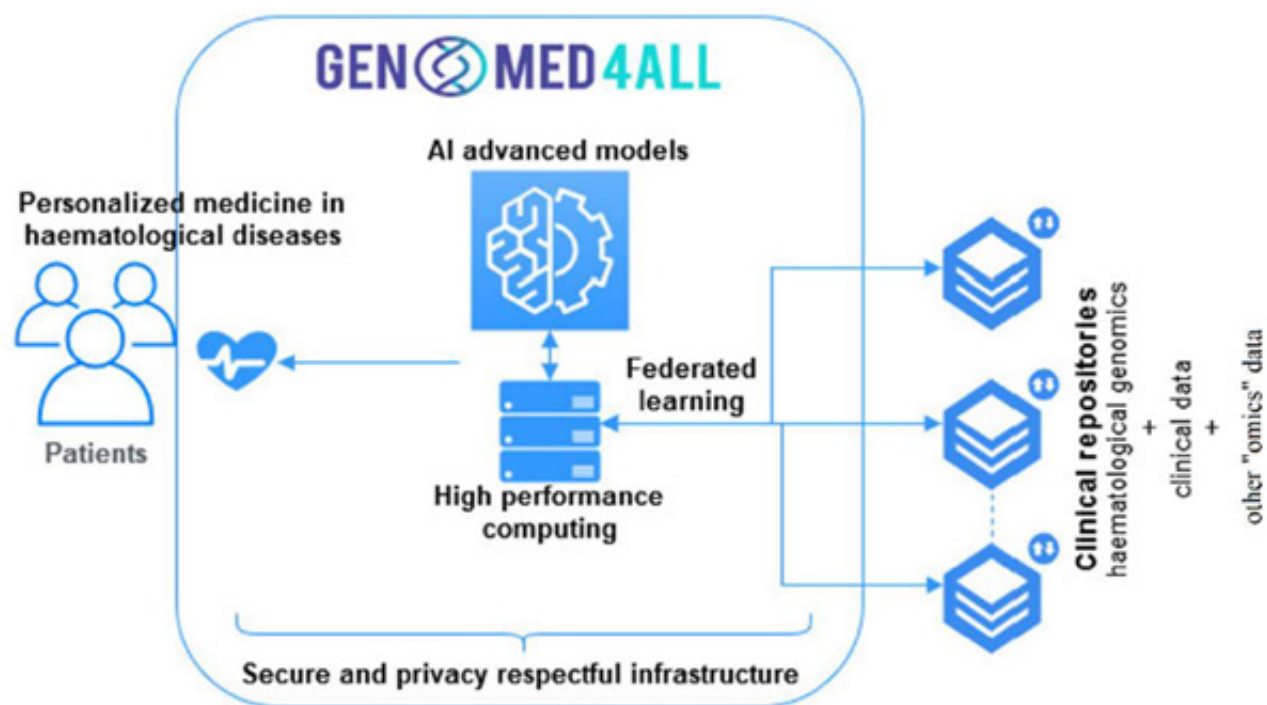


Figure 15: Genomed4a11

## Exscalate4Cov

EXaScale smArt pLatform Against paThogEns

Andrew Emerson, Federico Ficarelli, Maurizio Ortali, Neva Besker, Alessandro Grottesi  
Cineca

The response from researchers worldwide to the growing pandemic in 2020 was extremely rapid and supported by the European Union with funding initiatives such as the H2020-SC1-PHE-CORONAVIRUS-2020 call with a fast-track mechanism to rapidly review proposals and activate projects.

One of these is Exscalate4Cov (“EXaScale smArt pLatform Against paThogEns for Corona Virus”) ([www.exscalate4cov.network](http://www.exscalate4cov.network)), which combines the expertise of research labs across Europe with supercomputer facilities in Italy, Spain and Germany. The project started as early as April 2020 with 18 partners from all over Europe and will continue until October 2021. The initiative is centred around a virtual screening infrastructure based on the LiGen™ (Ligand Generator) software, which is owned by Dompe’ (a major Italian biopharmaceutical company) but has been co-developed with CINECA and other partners over the last 15 years.

The computational pipeline of the project involves running very long molecular dynamics simulations of viral proteins with the GROMACS software. The outputs from runs are then used as targets in the virtual screening procedure of Ligen. The software components have been optimized to make very efficient use of both Cineca’s M100 supercomputer and the HPC5 system of Eni S.p.A., two of the most powerful supercomputers in Europe. The result is that we can perform very long 10 microsecond molecular dynamics simulations in just a few months and docking runs capable of processing millions of molecules a second. An outcome of the docking experiments using a database of compounds known to be safe for humans (i.e. either known drug molecules or from natural products) has been the identification of a set of molecules

with a potential inhibitory effect on virus replication. Laboratory experiments confirmed the inhibition influence of some of these and one compound in particular, the osteoporosis drug Raloxifene, was selected for application to the European Medicines Agency (EMA). The application was successful and is currently in a clinical trial in Italy. Subsequently, a further two compounds have been identified and are awaiting approval from the EMA as clinical candidates.

In the short time available Exscalate4Cov has been remarkably successful, with over a trillion potential drug molecules and 45 proteins simulated and has made all the data generated within the project freely available outside the consortium, either directly from open-source repositories or via the associated partners programme. So far, the project is generating valuable data: more than 400 active molecules have been identified so far out of >30000 experimental data generated, and 29 peer review papers with a global impact factor > 154 point in 1 single year.

In addition, the Exscalate4Cov Team has deployed valuable web sites to support the global research community with bioinformatics and simulations tools, with the release of the most complete (> 40 simulations) and the most informative (> 10 μs) set of SARS-COV-2 molecular dynamics simulations.

Although the project will not extend beyond September 2021, it is hoped that the legacy of resources, know-how and data, will be of practical use in the current crisis and for future pandemics. Some of the computational components of the Exscalate4Cov project, will be anyway further extended and developed in the LIGATE EU Project, co-funded by the EUROHPC JU.



GENOMED4ALL has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101017549.



EXSCALATE4COV is supported by the European Union’s Horizon 2020 research and innovation programme under grant agreement No 101003551.

**EXSCALATE  
4COV**

# LIGATE

Ligand Generator and portable drug discovery platform AT Exascale

Andrew Emerson, Federico Ficarelli

Cineca

A remarkable success story of the current pandemic has been the speed with which safe and effective vaccines have been created and distributed to millions of people. However, despite this it is concerning that we have very few effective drug treatments for people already suffering from COVID-19. This is in fact in line with the observation that due to long and expensive development costs the number of new drugs coming to market in the last few years has been in constant decline. Using experience gained in projects such as Antarex and Exscalate4Cov, the LIGATE project will develop a CADD (Computer Aided Drug Design) infrastructure to accelerate the drug design process.

The LIGATE project aims to integrate and co-design best in class European open-source components together with proprietary (European) IP to keep worldwide leadership in CADD solutions exploiting today's high-end supercomputers and tomorrow's Exascale resources, fostering European competitiveness in this field.

The proposed LIGATE solution, in a fully integrated workflow, enables a drug design campaign to be delivered with the highest speed together with the highest accuracy by implementing the auto-tuning of the solution parameters to meet the time and resource constraints. This predictability,

together with the full automation of the solution and the availability of the Exascale system, will allow the full in-silico drug discovery campaign to be run in less than one day to respond promptly, for example, to a worldwide pandemic crisis.

In the LIGATE project, CINECA aims to deploy leading-edge HPC systems, including the Leonardo pre-exascale system which is expected to provide a peak performance of 250PFlops. These systems will be used as a test-bench for LIGATE technologies and will evolve during the duration of the project towards the Exascale level.

The main goal of CINECA in LIGATE is to validate and deploy the CADD solution on its systems, including the Leonardo system, and make it available to academic and industrial users. To reach these goals, CINECA will expand its training program on HPC technology to include the CADD solution. For industrial users, CINECA will also demonstrate the CADD solution by specific actions carried out by the Italian EuroHPC co-funded HPC Competence Centre (e.g. workshops on Drug Design targeting specifically industrial users). Moreover CINECA will offer to interested industrial users, support of the use of the CADD solution and the usage of its own HPC resources, including Leonardo.



LIGATE has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 956137. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Italy, Sweden, Austria, Czech Republic, Switzerland.



## EARTH AND ENVIRONMENTAL SCIENCES



# The operational weather forecast in Italy

to support the country's multi-sector users.

The opportunity offered by the ItaliaMeteo Agency

**Carlo Cacciamani**

**Head of the IdroMeteoClima structure  
of the Regional Agency for Prevention, Environment and Energy  
of the Emilia-Romagna Region**

## 1. THE ROLE OF METEOROLOGICAL INFORMATION IN TODAY'S TECHNOLOGICAL SOCIETY

The meteo-climatic information, consisting of the set of monitoring activities of the evolution of "weather", meteorological forecast and climatic evaluation, represents an important step in support of a vast spectrum of public and private users who, thanks also to an adequate "weather" support, they can create specific services for the benefit of the "chain" of activities they carry on.

One of these users is certainly the national civil protection system. In this case the weather-climatic information makes possible to evaluate downstream risk

scenarios induced by extreme weather events hitting vulnerable and exposed areas. Weather information is essential also for the "energy sector", to optimize the production and consumption of energy in "real time" and plan its storage and distribution through networks. Agriculture is another important "user" of "meteo". Specific meteo-climatic services are needed for planning and executing field operations, as well as the agronomic practices for the crop defense. The list can continue with the management of water resources (production and distribution), support to transport and tourism sectors: climatic weather information are necessary for ship routing services, management of rail and road traffic and planning of winter and

summer tourism.

## 2. HOW IS OPERATIONAL METEOROLOGY STRUCTURED IN ITALY TODAY? WHAT NEEDS IMPROVEMENT?

Faced with the great need for "weather" that the country is asking for, in Italy today the quality of the weather-climate offer is still too limited. Unlike most countries in the world, Italy does not yet have a civil national meteorological service that provides an adequate response to the demand for "weather" towards public and private users. In Italy only the Meteorological Service of the Air Force operates, providing support to aviation and other strategic defense functions. Due to the lack of a multi-purpose national weather service, a considerable number of regional weather services have flourished over the years to meet the needs of regional and local administrations, as well as a public and private market which has developed significantly over time and needs these information.

Unfortunately, there is no unitary governance in this system distributed throughout the territory, and this lack has produced the absence of common monitoring standards, duplication of services and therefore economic disadvantages and poor penetration capacity of the country into the European meteorology market.

Only through the construction of a national meteorological structure, necessary to coordinate such a system, national meteorology will be able to provide greater support to multiple sectors of activity, making better use of new meteorological monitoring technologies and resources of supercomputing (HPC) necessary to run the increasingly sophisticated meteorological modeling chains and manage the massive amounts of data available today.

## 3. THE ROLE OF "ITALIAMETEO": THE NEW NATIONAL WEATHER AGENCY SUPPORTING THE COUNTRY

The implementation of the ItaliaMeteo Agency, approved at the end of December 2017 (Law n. 205 / 2017), certainly marked a turning point in the more than twenty years of reorganization of national meteorological services, paving the way for the integration of military and civil, state and regional knowledge and capabilities in the sector. The Law 205/2017 established also the National Steering Committee for Meteorology and Climatology, a body composed of thirteen experts in the sector in charge to ensure coordination and strategic direction and, on an operational level, the birth of the ItaliaMeteo Agency, for meteorology, climatology and the state of the sea.

Today the construction process of the Agency is completed. The two founding documents, provided for by Law 205/2017, the Statute and the Implementing Regulation of the Agency, have been approved by the Government and now only the appointment of the Director of the Agency is awaited in order to be able to start it up.

The Agency, with headquarters in Bologna, will distribute weather and marine forecasts, homogenize communication languages to citizens, promote applied research in weather forecasting and climatology, manage measurement networks, organize and disseminate observed data through telecommunication channels. Furthermore, the Agency will have to provide training, participate in European and international projects and programs and promote collaboration with public and private entities. In addition to these tasks typical of a national weather structure, the ItaliaMeteo Agency will also have to coordinate the varied System of Weather Bodies operating mostly at the regional level.



# The Italy Flash Flood use case in MISTRAL

Meteo Italian Supercomputing Portal

Tim Hewson, Estibaliz Gascon and Andrea Montani  
ECMWF

ECMWF (the European Centre for Medium range Weather Forecasts) was a key participant in the Mistral project, which ended in January 2021. Its role was to create a real-time stream of new and innovative rainfall forecasts for Italy. The scientific goal was to deliver forecasts that are much more accurate than raw model forecasts, by exploiting and expanding upon pre-existing forecast post-processing techniques. These techniques are applied to the raw output of ensemble prediction systems, with CINECA HPC facilities being the platform for the post-processing. The practical goal

was then to deliver these new products to forecasters across Italy, primarily to help with the challenge of flash flood prediction. Commonly flash floods arise as a result of short-period extreme rainfall, and so the products were created with this in mind, by focussing on overlapping 6h forecast periods.

The ensemble inputs were provided by two very different systems. We had the world-renowned ECMWF ensemble, with 51 members, running at 18km horizontal resolution and covering the globe, and also the COSMO-2I-EPS (CONsortium for Small-

percentile 95 (mm) ; fcst step: 42-48 hours; VT: Friday 22 January 2021 0000 UTC

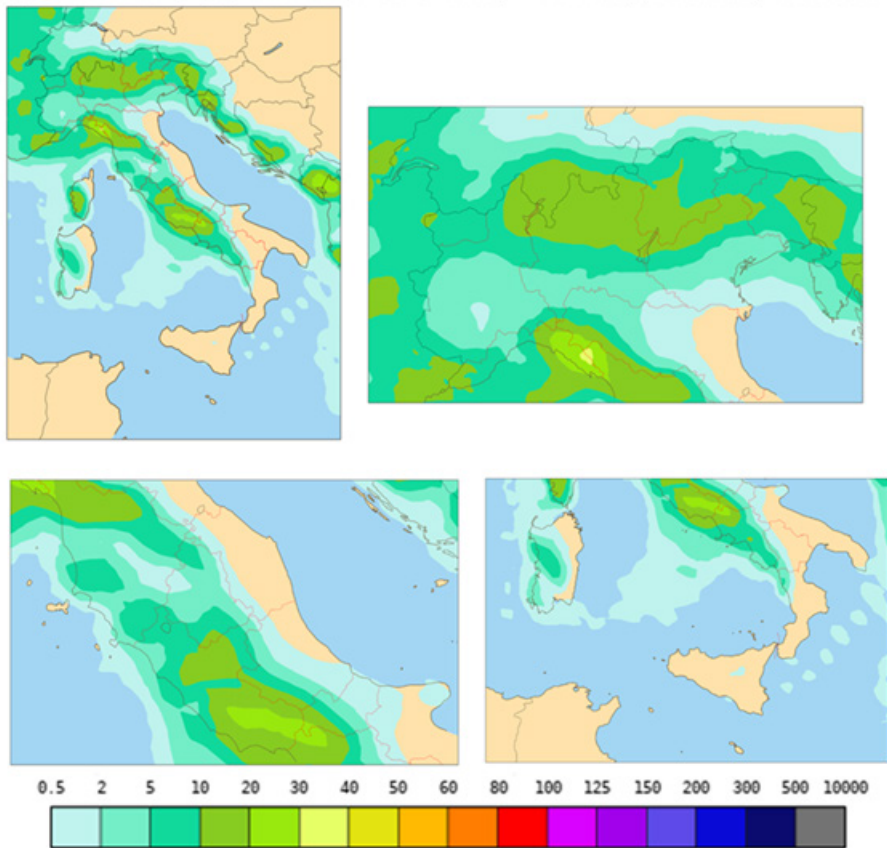


Figure 16: 95th percentiles (in mm) of 6-h precipitation forecast from flash flood products for lead times T+42-48h and initialized on 19th January 2021 at 00 UTC. Top left corresponds to the product covering COSMO-2I-EPS, top right, the North of Italy, bottom right, central Italy and bottom left, the South of Italy.

scale MOdelling) system, which runs at 2.2km over Italy and some neighbouring regions, with 20 members. In the ECMWF ensemble convection, which can often be source of flash floods, is parametrised, whilst in the latter it is explicitly resolved. This needed the use of two different post-processing approaches, tailored accordingly.

To post-process the ECMWF ensemble we used the innovative ecPoint technique, pioneered at ECMWF, which has been delivering real-time experimental global products since April 2019. These products have a 12h prediction period. For Mistral's 6h period we re-engineered the post-processing decision tree by bringing in new branching to help correct for diurnal cycle errors, and to handle the influence of complex topography, which Italy has in abundance. Ultimately ecPoint aims to introduce weather-dependant sub-grid variability, and weather-dependant bias correction. The sub-grid variability facet enables to anticipate very localised extreme rainfall, which tends to drive flash floods. Sometimes predicted values for points can be as much as 5x more than is shown by the wettest ensemble member (albeit with low probabilities of such totals at a given site).

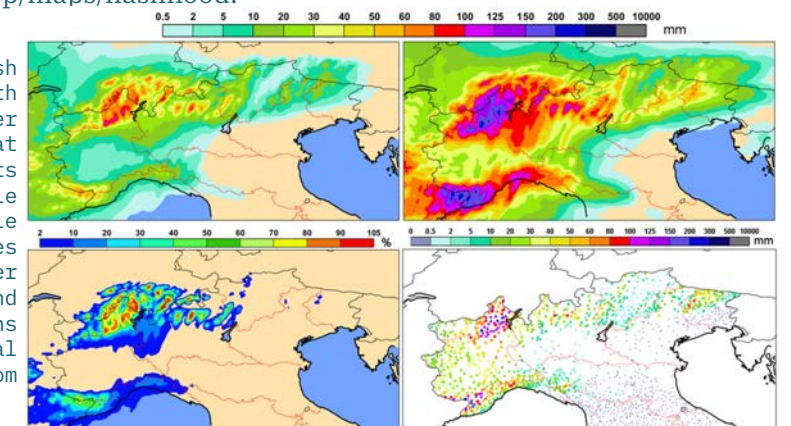
In post-processing the COSMO-2I-EPS output focus instead on compensating for the lack of members. Conventional wisdom is that one would need well over 1000 members to properly represent all the degrees of freedom arising in convective cases in particular, rather than the 20 we had. So we use a scale-selective neighbourhood post-processing approach, which can work well at shorter lead times in particular. This aims to retain fine-scale signals in the ensemble where the members' rainfall totals agree well (e.g. for rainfall arising from strong topographic forcing), and to spread out signals across much larger "neighbourhood" regions where agreement is poor (e.g. when there are scattered heavy showers).

Following the post-processing activities is used a weighted blending approach to create the final products, giving increasing weight to ECMWF output as lead times lengthen.

Although not as computationally intensive as running numerical models themselves, the post-processing still provides a challenge for HPC, especially when the operational necessity of short, reliable run times is recognised. Task parallelisation helps us achieve this.

Routine verification shows that the post-processed output improves markedly upon the raw model output, with both components contributing. Flash flood case studies also highlight clear benefits. Similarly, initial feedback from some users at DPC in Rome, who provide warning guidance, has been very positive, and with the suites continuing to run now that the project period is over we hope and expect the user-base to grow further. Products are updated once per day, and can be found in the Mistral Open Data Portal here: <https://meteohub.mistralportal.it/app/maps/flashflood>.

Figure 17: Italian Flash Floods products example with initialization on 2nd October 2020 at 00 UTC and valid at 21 UTC the same day. The plots represent the 50th percentile (top left), the 99th percentile (top right), the probabilities of 6-hours precipitation greater than 50 mm (bottom left) and the corresponding observations available in the MISTRAL portal for the same period (bottom right).



Co-financed by the Connecting Europe Facility of the European Union

Co-financed by the Connecting European Facility Programme of the European Union



# The new climate datasets in HIGHLANDER

HIGH performance computing to support smart LAND sERvices

Paola Mercogliano, Monia Santini,  
CMCC

Gabriella Scipione, Cinzia Caroli,  
Cineca

Estibaliz Gascon, Augustin Vintzileos, Tim Hewson,  
ECMWF

One of the objectives of the HIGHLANDER project is to generate, through the use of High Performance Computing, new climate datasets that support a better and more sustainable management of natural resources and the territory and consequently help reduce the risks associated with climate change.

Now the first important results have been published and others will be shortly.

On March 31, 2021, the new dataset VHR-REA\_IT ([https://doi.org/10.25424/cmcc/era5-2km\\_italy](https://doi.org/10.25424/cmcc/era5-2km_italy)) for the recent climate was released. The data were produced, as part of the project, by dynamically downscaling the ERA5 reanalysis (<https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.3803>) over Italy with the COSMO-CLM model (<https://www.cmcc.it/models/cosmo-clm-climate-limited-area-modelling-community>), to go from the native ERA5 resolution ( 31 km) to a resolution of 2.2 km for the period 1989-2020.

The climate simulation was configured and performed by the CMCC Foundation (<https://www.cmcc.it/>) on the GALILEO supercomputer of the Italian computing centre, CINECA. GALILEO is equipped with 1022 36-core compute nodes, each one containing 2 18-cores Intel Xeon E5-2697 v4 (Broadwell) at 2.30 GHz. The model run was conducted using 60 nodes, corresponding to 2160 cores, and employed about 61 hours to simulate 1 year. The simulation produced

8 TB of outputs, which are now available to support project applications and case studies, as established in agreement with HIGHLANDER project partners.

The typical use of such a dataset is indeed research and downstream services, like calculation of indicators and indices, as well as feeding impact models, to serve decision support systems concerning several land components.

At the same time, the partner ECMWF

(<https://www.ecmwf.int/>) has produced the first release of statistically downscaled sub-seasonal forecasts (Sub-SEA) using the ecPoint post-processing technique (<https://www.nature.com/articles/s43247-021-00185-9>).

“The sub-seasonal ecPoint forecasts are post-processed products delivering probabilistic forecasts at point-scale (instead of at grid-box scale) of global 24-hour precipitation and minimum, maximum and mean temperatures in 24h, for weeks three and four of the forecast (days 16-30). They are based on the ECMWF ensemble (ENS) forecast system with 51 ensemble members (1 control and 50 perturbed members), run at ~0.35° horizontal resolution (~36km), with the final outputs for ecPoint precipitation and temperature provided as percentiles for each grid box.” (Deliverable D4.3 of the Highlander project, available at: [https://highlanderproject.eu/wp-content/uploads/2021/07/D4.3\\_Downscaled\\_subseasonal\\_forecast\\_update20210629.pdf](https://highlanderproject.eu/wp-content/uploads/2021/07/D4.3_Downscaled_subseasonal_forecast_update20210629.pdf)). The ecPoint technique adjusts the raw model gridbox forecasts to cater for both systematic biases at grid-scale, and sub-grid variability; such adjustments depend on the forecast weather type (in a gridbox) as well as other factors.

The Sub-SEA products are being computed twice a week on the CINECA HPC facility of GALILEO and are now available to the Highlander project partners. Their application area includes agricultural planning and animal and human welfare. Initial user feedback has highlighted dramatic improvements in forecast quality for 2m temperature at sites near coasts; this is reassuring as poorer near-coast performance in raw model output was one of the aspects targeted in the first iteration of the ecPoint post-processing decision trees.

These two “climate datasets”, and others, produced in the first phase of the HIGHLANDER project, will be used by the

other partners (Arpae, FEM, Arpap, CIA) for the implementation of various Downstream Applications and (pre-)Operational Services (indicated with the acronym DApOS). The design and development of the DApOS are already underway and we are looking forward to integrating them in the Highlander portal hosted on the CINECA cloud.

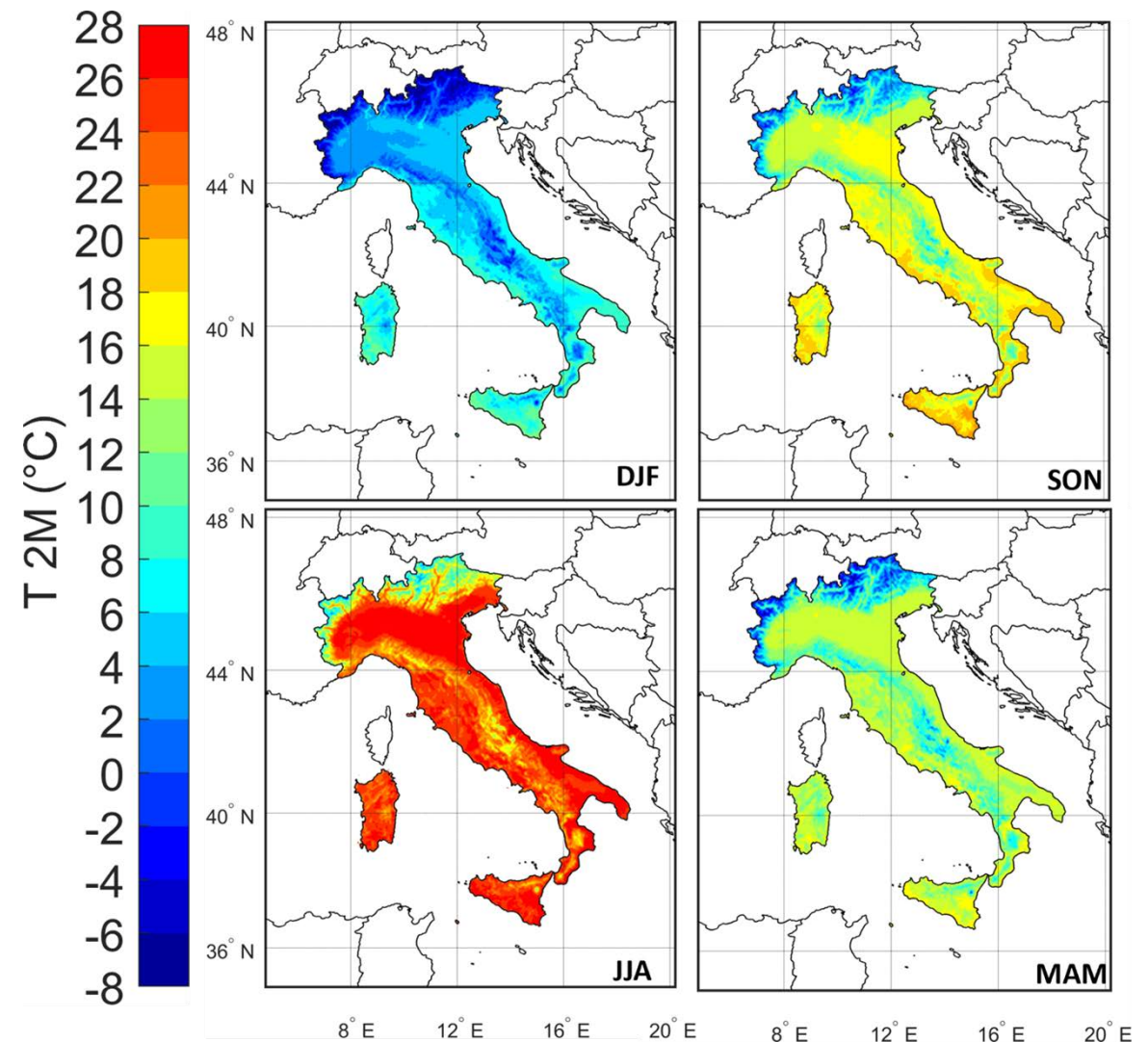


Figure 18: Seasonal temperature trend at an altitude of 2 meters for the period 1989-2020 obtained from the downscaling of the ERA5 re-analysis to about 2.2 km.



Co-financed by the  
Connecting European  
Facility Programme of  
the European Union  
Grant agreement n°  
INEA/CEF/ICT/  
A2018/1815462



# AIDA

Artificial Intelligence Data Analysis

Francesca Delli Ponti  
Cineca

The H2020 AIDA project uses some of the most advanced data analysis techniques to study large streams of data available from space missions and space simulations. The aim of the AIDA project is two-fold, targeting scientific discoveries in the space science as well as it is directed at transforming the way society can access the results of space missions and simulations, making the data accessible and giving, including non-expert citizens, access to the utmost performing tools, such as machine learning techniques, to analyse the data. Just as amateur astronomers have in the past discovered moons, comets and astronomical processes, the open-source software AIDApY will empower citizens to use the most advanced tools in data analysis and AI to explore space science.

Part of the project, two main products are defined:

- AIDAdb, a collection of simulation output data as well as high-level data stored on CINECA computers and accessible via iRODS metadata;

- AIDApY, a high-level Python package for the analysis of spacecraft data from heliospheric missions using modern techniques. In particular, AIDApY centralizes and simplifies access to spacecraft data from heliospheric missions, space physics simulations, advanced statistical tools, Machine Learning, Deep Learning algorithms, and applications.

For more information on them, how to use and to be update on advancements, see the dedicated webpage:

[http://aida-space.eu/AIDApY\\_AIDAdb](http://aida-space.eu/AIDApY_AIDAdb)

The dissemination of the project is conceived mainly to involve a larger community to use these new instruments. To this end two schools were planned. The first one was on January 2020 in CINECA, dedicated to Machine Learning and data analysis together with a first introduction on the two project's developed tools. We received excellent feedback from the 40 students attending the school at CINECA.

The students appreciated very much both scientific quality and organisation. To enlarge communication, the content of the school was recorded and set available on the AIDA project official website.

The AIDA project arriving to its close in 2022, a list of activities are yet to be finalized as they were postponed due to the pandemic situation worldwide, particularly activities regarding the second summer school and dissemination.

The second AIDA school program will be dedicated to the products, AIDApY and AIDAdb, in order to broadcast the open-source tools that are made available to the community.

Despite the end of the AIDA project in 2022, the official website will remain active and will be regularly updated.

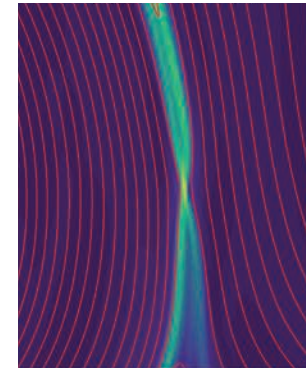
Stay in touch with us!



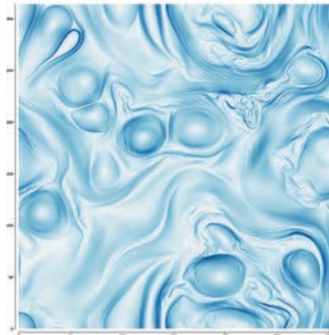
The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.



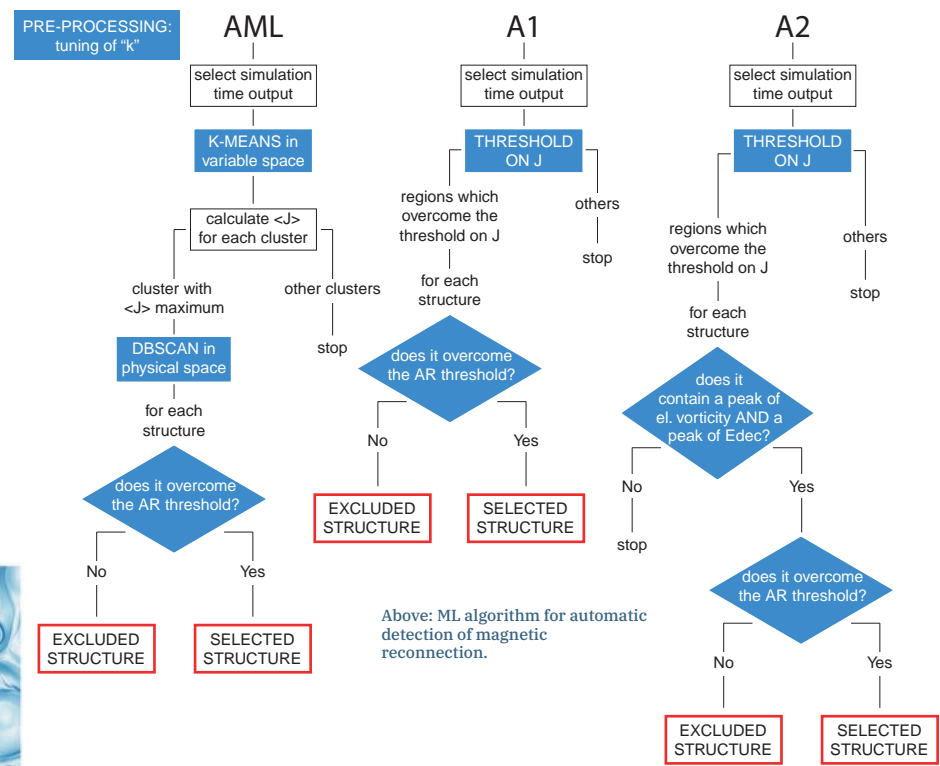
## AIDA PROJECT aida-space.eu



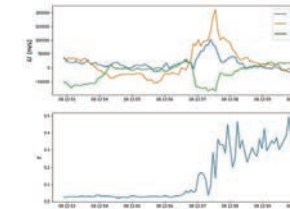
A characteristic reconnection structure as emerging from numerical simulations and routinely detected in space by satellites.



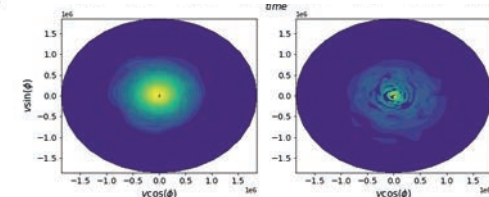
Plasma turbulence kinetic numerical simulation (AIDAdb) showing the formation of reconnection at electron scale see Califano et al. Frontiers in Physics 8 (2020)



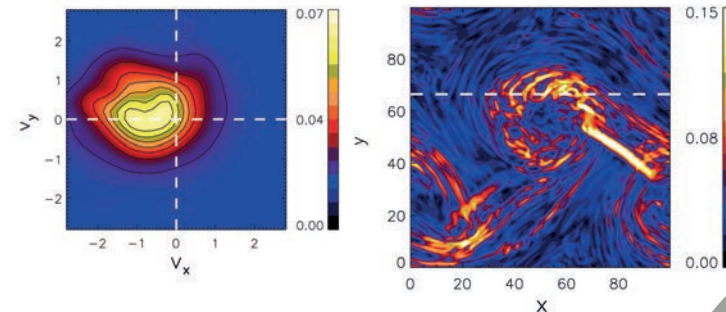
Above: ML algorithm for automatic detection of magnetic reconnection.



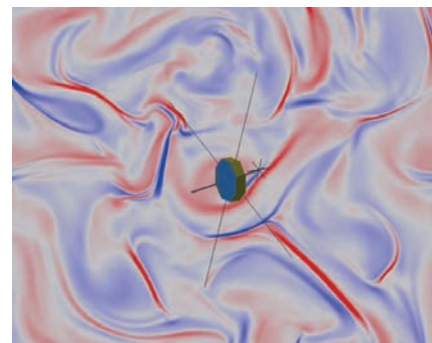
On the left three ion bulk flow speed components, in GSE coordinates, obtained from one of the MMS satellites for the time interval considered and the epsilon parameter, i.e., an estimator of non-Maxwellianity for distribution functions is computed using the extreme events package of AIDApY, and shown for the same time interval. We see a consistent increase in epsilon in the stream. Below, a VDFs, collected from MMS, for epsilon close to zero ("quiet phase", left), and for epsilon around its maximum ("bursty phase", right).



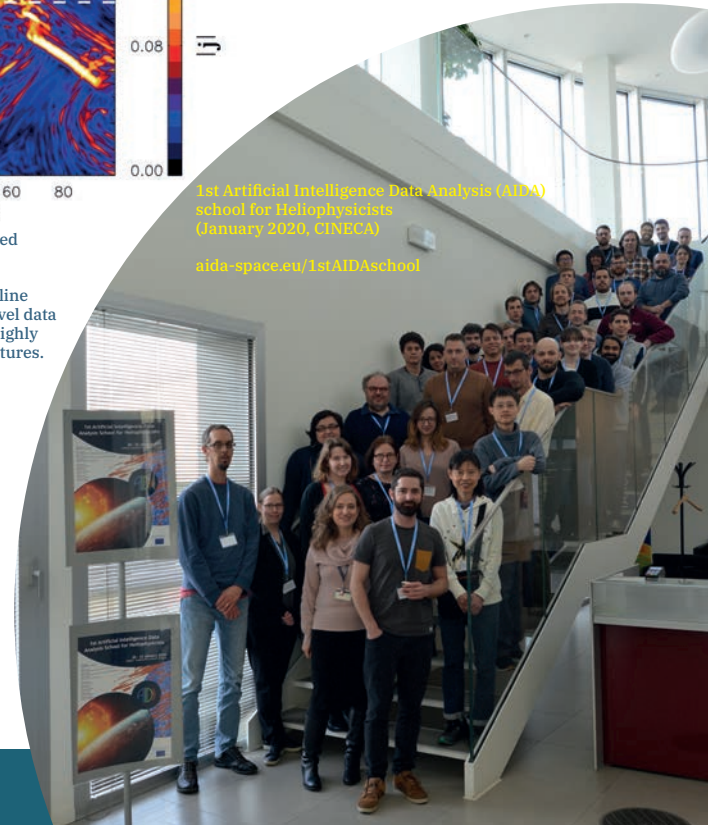
For more info on AIDApY and AIDAdb:  
[aida-space.eu/AIDApY\\_AIDAdb](http://aida-space.eu/AIDApY_AIDAdb)



Above a turbulent stage of the Kelvin-Helmholtz instability in magnetized collisionless plasmas. On the right, an example of the turbulent 2D configuration with the development of the instability is depicted by the contour plot of the current density in the (x,y) plane. The white dashed line indicates the virtual satellite trajectory that can be found in the high-level data directory. On the left, a cut plot of the proton VDFs, the figure shows a highly distorted VDF and the presence of both anisotropic and agyrotropic features.



Visualisation of virtual spacecraft travelling in a simulated, turbulent plasma environment.



1st Artificial Intelligence Data Analysis (AIDA) school for Heliosphericists (January 2020, CINECA)

[aida-space.eu/1stAIDAAschool](http://aida-space.eu/1stAIDAAschool)

# Into the (un)known of the Cosmo

Giannandrea Inchingolo

Dipartimento di Fisica e Astronomia, Università di Bologna

Istituto Nazionale di Astrofisica

Fiber of the Universe is a project between Giannandrea Inchingolo of the Physics and Astronomy department of the University of Bologna and Cineca VisitLab. This collaboration aims to create a pipeline to visualize astrophysical data using Cinematic Scientific Visualization tools.

In collaboration with the (AM)<sup>2</sup> research group of Prof. S. Morigi of the Applied Mathematics department of the University of Bologna, we offered two master internships and one master thesis aimed to transform astrophysical simulation data output, as .hdf5 and .dat data formats, in 2D meshes and 3D volumes editable in Blender.

We choose the .vtk and .vdb formats to take advantage of the scientific visualization tools and CG expertise available at VisitLab (e.g. BVTK Nodes).

The output of this project is the creation of new visualization techniques for astrophysics analysis and the production of short videos for the dissemination of state-of-the-art astrophysics research to the general public.

A case of study is by the video "The VLA shedding lights on the origin of radio relics", obtained combining 3D cosmological simulations and 2D observation images from GI's research group led by Prof. F. Vazza.

For this video, we combine volumetric rendering of cosmological simulations with spherical projections of 2D observation. The video won the first award of the NRAO Image competition for the VLA 40th anniversary celebration.

Thanks to the possibilities given by the cinematic scientific visualization, we combined representation of multi-scale dynamics, showing both large scale cosmological simulation and small-scale particle evolution of the same scenario. This peculiarity offered a new visualization of the scientific results that inspired

a further research data analysis, currently under publication in a peer-reviewed paper.

For the dissemination of the astrophysics contents, we produced a series of 4K UHD3DVR 360-degree short videos and 8K videos. The videos are part of a Science-Art project developer by GI named Into the (un)known.

Into the (un)known is a multi-sensory, immersive exploration of the Cosmos, where the most recent astrophysical data are heightened into real works of art, so to allow the public to access scenarios - until now reserved only to researchers - with emotional immediacy.

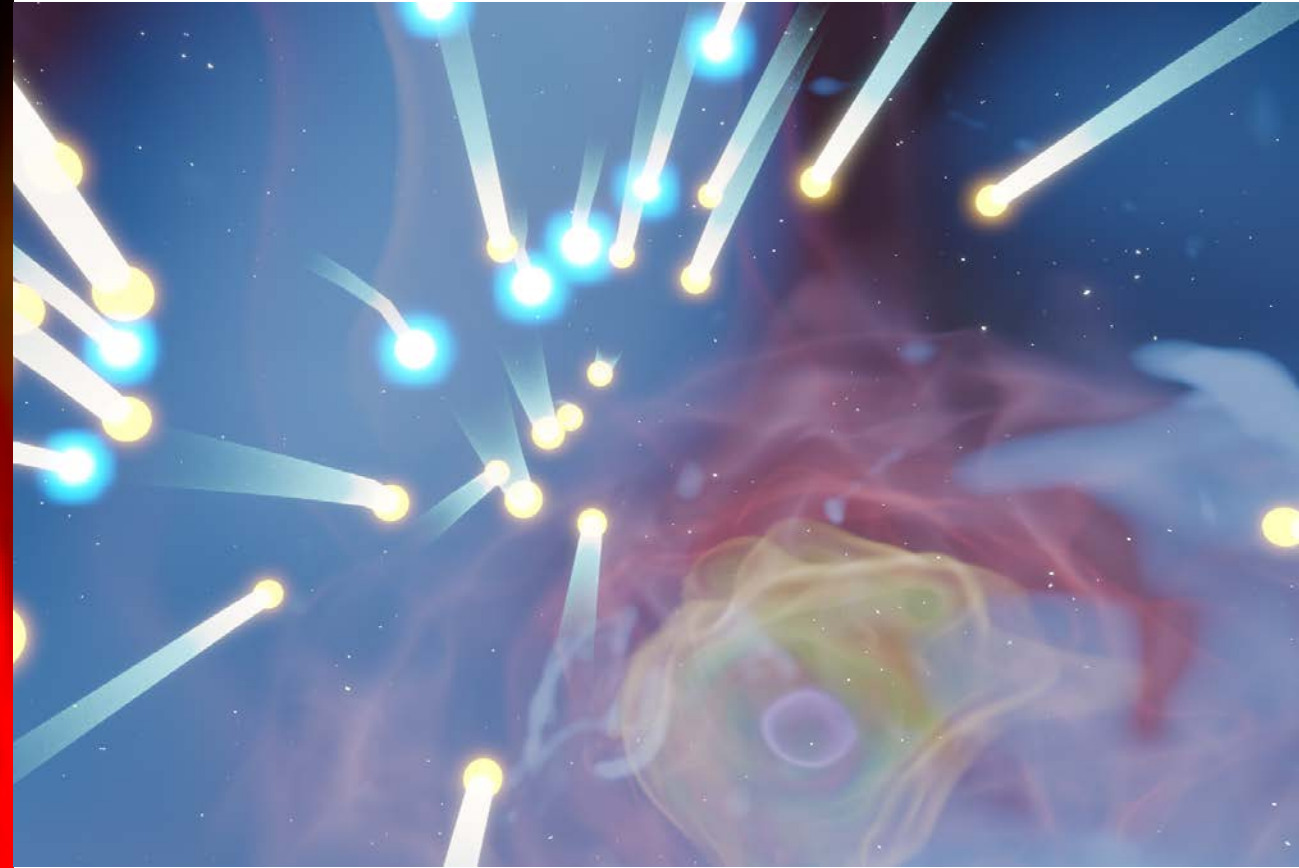
The 8K and the 4K VR videos produced were part of an in-presence museum exhibition of Into the (un)known at the 2020 edition of Festival della Scienza in Genoa, organized in collaboration with the Istituto Nazionale di Astrofisica (INAF) and the Media Lab of the Massachusetts Institute of Technology in Boston (data sonification).

We decide to distribute the VR contents using Youtube, so they are easily accessible using personal devices (smartphones) and Cardboard for the 3D stereoscopic view.

The pipeline allows to easily reconfigure the workflow according to the requirements of the different output video content to be produced. In particular, the short videos produced will be composed in single storytelling for the exhibit of Into the (un)known at the MEET Digital Culture Center in Milan in Fall 2021.

With Fiber of the Universe, we combine Art and Science, using an artistic expression to improve the scientific visualization tools and to communicate Astrophysics and Science in general to a broader and diverse audience, where usually state-of-the art research does not enter easily.

Figure 19,20, 21:  
into the (un)known  
of the Cosmo





---

BIG DATA  
AND  
ARTIFICIAL  
INTELLIGENCE

# Big Data in the analysis of TV ratings

Donatella Sforzini  
Cineca

For more than 20 years, CINECA has been collaborating with the Rai Marketing Department within the activities of the “Public Service Detection and Monitoring Systems” structure.

One of the tasks of this structure is to supervise surveys on media audiences (TV, radio and web), updating the techniques and methodologies for surveying and statistical representation of vision behaviours according to the evolution of supply and consumption, but also by analyzing surveys results to obtain the maximum information contribution and the evidences to support of Rai’s (editorial and advertising) offer.

Auditel, the official information source for estimating TV audience in Italy, produces every day, only for TV sets, the minute-by-minute audience data of around 150 national broadcasters (DTT and SAT, free and pay) and around 190 local broadcasters. These data are estimated on the basis of actual consumption recorded by the Auditel “sample system”, a panel of more than 40,000 individuals aged 4 years and

over belonging to approximately 16,100 households in whose homes there are the so-called “meters” that record the acts of viewing on each TV set. All the listeners’ socio-demographic information is then linked to the acts of listening.

The complexity increases if we consider that the surveys do not stop at traditional TV ratings but also extend to what Auditel calls the TV beyond the TV set, i.e. the TV enjoyed both at home and on the move, on digital devices such as Smart TVs, Personal Computers, Smartphones, Tablets and Game Consoles. Auditel’s “census scheme” is used to assess these audiences, measuring, directly from the content provider, the use of any editorial or advertising content issued and actually used by users.

CINECA has provided methodological support for the Auditel research system over the years, as well as specific studies such as audience clustering, analysis of listening flows (i.e. how the audience moves between broadcasters and times of day rather than between different periods) and forecasts.

Starting from 2020, a series of proof of concept (PoC) have been launched with the goal of integrating Auditel data with new sources (internal and external to Rai) and exploiting the potential of artificial intelligence applied to the television audience domain.

The current PoC concerns the development of predictive models that provide estimates on TV listening to support the definition and analysis of alternative competitive scenarios.

The idea is to estimate the audience ratings of a program by providing a set of information that characterizes it both from the editorial point of view, or the “structural elements” of the program (genre, conduction, duration, topics covered, talents in video, social interactions,...), both as regards its expected positioning in the schedule (season, network, type of day, time slot).

The procedure involves two steps:

- the initial phase of program definition, for the selection of the “mix of ingredients” (structural and positioning elements) that can maximize the expected return in terms of audience;

- the operational definition phase of the schedule, considering the hypotheses on the programming / counter-programming choices that could be made by the competitors.

The complexity of this proof of concept lies in the fact that the final result of the evening to be estimated will have to be constructed by integrating, in the appropriate way, the forecasts from various predictive models which, at this stage, are assumed to be declined according to the different TV genres to be considered (film, entertainment, information, etc.).

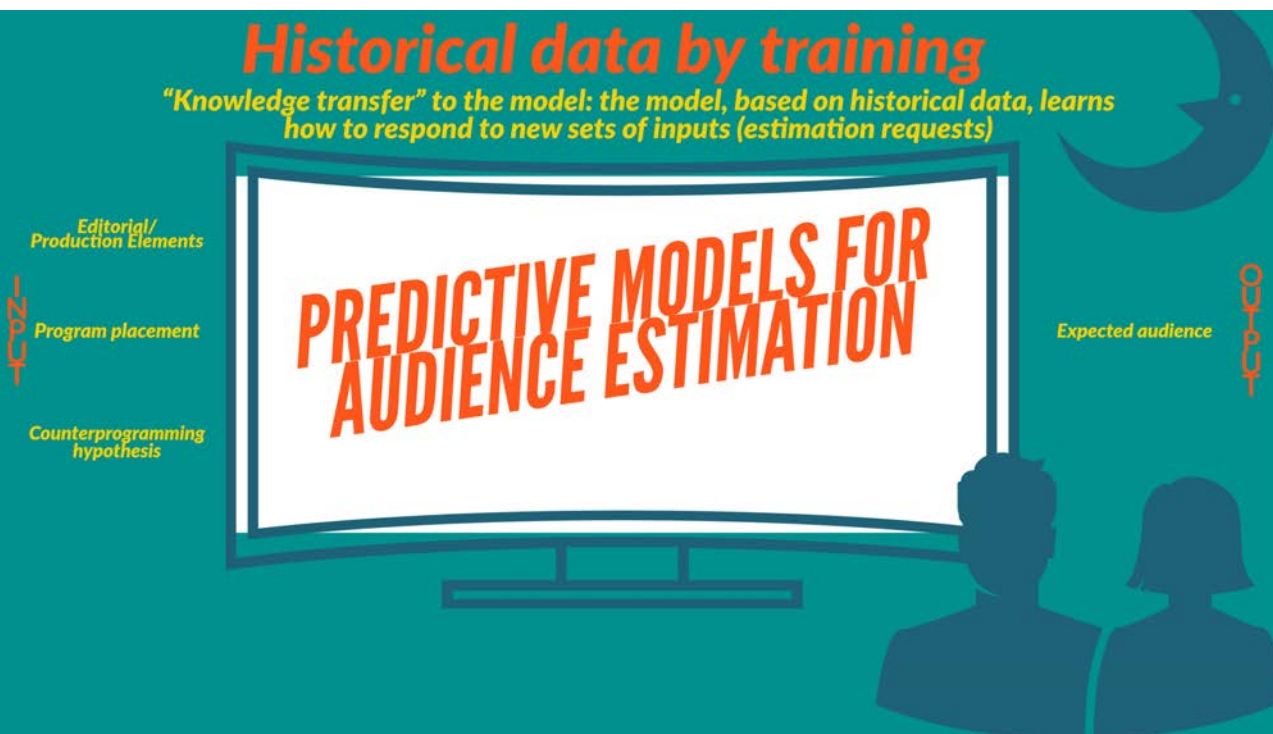


Figure 22: predictive models for audience estimation

# Fostering AI innovation

Marco Rorro  
Cineca

The NVIDIA AI Technology Centre (NVAITC), established on January 2020 and forged on the agreement between NVIDIA, CINI and CINECA, is a program to enable and accelerate AI research projects in Italy. The program is a national-level agreement built around project-based collaborations with institutions member of the CINI network and fosters the collaboration of CINECA for computational resources. It aims to enabling academic institutions at all levels to conduct their research more efficiently by collaborating into research projects, training students, nurturing startups and spreading adoption of the latest AI technology throughout Italy. Areas of contribution include DL/ML frameworks adoption, software development and environments set-up, performance optimization and hyper-parameter tuning, support for efficient data loading, mixed-precision training, inference, multi-GPU scaling, research participation, and dissemination of results. NVAITC projects can enable academics at all levels to do their research more efficiently as NVIDIA's GPUs are largely used for AI model training and inference.

After only one year of activity the center has helped the AI community to grow significantly leading to seven scientific publications in

co-authorship published, more than one million of GPU hours consumed at CINECA, 700 students trained and one workshop (CADL) about computational aspects of deep learning organized in combination with ICPR conference. Furthermore, the center will officially contribute to the National PhD program in AI delivering lectures, supporting students and providing computational resources to most relevant research projects. In order to participate to the program and receive support, interested PIs can submit at any time a proposal using this template (<https://www.conorzio-cini.it/index.php/en/labaiis-home/labaiis-researchandprojects/1629-nvidia-ai-nation-program-and-nvaitc-proposal-template>) via email to [segreteria.aiis@consorzio-cini.it](mailto:segreteria.aiis@consorzio-cini.it). The local lead engineer (Giuseppe Fiameni – [gfiameni@nvidia.com](mailto:gfiameni@nvidia.com)) can be contacted for any input, suggestion, or advice before submitting it. Proposals are reviewed with the help of fellow NVAITC engineers on a first come-first-serve basis. Evaluation is based on NVAITC criteria (target publication, technology stack and computing scale), rules of engagement and shared realistic expectations and usually takes less than two weeks.

# MATERIAL SCIENCES



# TREX and MaX supporting science to exascale

Targeting Real Chemical accuracy at the Exascale - Materials at Exascale

Fabio Affinito  
Cineca

One of the most important activities carried out by Cineca consists in supporting the scientific communities in their computational activities. This involves the technological evolution of the scientific codes and the direct support to the users in order to improve the correct exploitation of the computational resources. In many cases, this work is performed under the hat of the EU Centres of Excellences (CoE). For the field of Material Sciences and Computational Chemistry, Cineca takes part to two CoEs: MaX, Center of Excellence for Material Sciences, started in 2016, coordinated by the Italian CNR. Among the partners of MaX there are HPC centers (JSC, BSC, CSCS, Cineca) and the developers of the most important scientific codes on the domain of material sciences (SIESTA, Quantum ESPRESSO, Yambo, BigDFT, CP2K). The synergy between code developers, scientist and HPC technologists drives the evolution of the codes in order to make them ready for the exascale transition. The same strategy has been addressed by the TREX CoE, which started its activities in the 2021. Also in this case, among the partners there are HPC centers and international institutions of research. For TREX the focus is on Quantum Chemistry codes, with a particular interest on Quantum Monte Carlo codes, which are typically the most suitable codes for scaling up to a huge number of computing nodes, re-vesting a particular importance as a case of interest for the exploitation of the new HPC hardware.

Both for TREX and MaX, the role of Cineca is twofold. On one side we contribute to the activities of co-design of the applications. This means to study the approaches that will permit to better exploit the hardware

enabling kind of calculations relevant for the most innovative scientific challenges. A particular role is also the work committed to reduce the “energy” spend for running a simulation, by adopting new algorithm or making the code able to run on devices able to reduce the energy consumption. On the other side, we are committed in the activities of user support and dissemination. The support that we give to the users of MaX and TREX permits to find the weakest spots in the calculation and to individuate improvements that can be made to the codes to make them more efficient. Together, we are part of an activity of dissemination based on tutorial and schools. This aspect, more than others, was difficult in the next year, due to the COVID-19 pandemics. This led to a complete revolution of the tutorial and workshop, in particular for the MaX project, where all the plans needed a deep change. Nonetheless we were able to transform all the planned activities in virtual ones. This was quite difficult, but at the end permitted to carry on all the activities with an equal success and, in some cases, with the benefit of reaching people that in “ordinary times” would have not been able to afford long travels. Also TREX is walking on the same path, and the first large event, the “QMC school” organized with the support of Cineca received a general appreciation by the attendees.

For the next years, all the CoEs, including MaX and TREX, will be called to interact with the new EuroHPC infrastructures, and in particular with Leonardo. This will be a brand new challenge for which we will commit to work together with the partners of MaX and TREX.

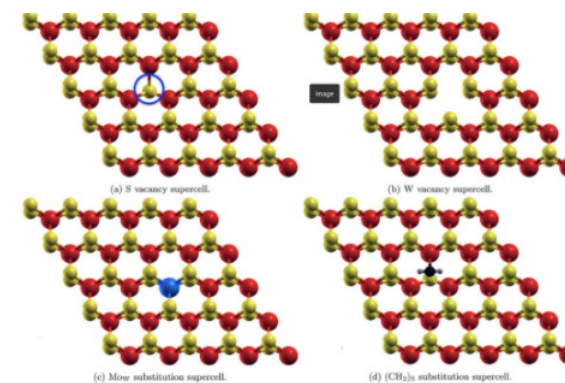


Figure 23: Effects of impurity and vacancies in 2D materials calculated with Quantum Espresso and Yambo, by De Melo et al (<https://doi.org/10.1002/qute.202000118>)



MAX has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 824143.



The TREX: Targeting Real Chemical Accuracy at the Exascale project has received funding from the European Union's Horizon 2020 - Research and Innovation program - under grant agreement no. 952165.



# QUANTUM COMPUTING

## CINECA Quantum computing perspectives

**Daniele Ottaviaani**  
Cineca

Throughout its decades of history, CINECA has been always providing state-of-the-art High Performance Computing solutions. In order to meet this target to the highest standards, CINECA meticulously keeps track of the latest advances in emerging technologies.

Quantum Computing represents not only a novel technology but also a totally new approach to computation, based on the principles of quantum mechanics, which has the potential to address and solve problems that current supercomputers would not be able to handle. Quantum computing has truly revolutionised the core aspects of computation by replacing the concept of a binary logic unit, the classical bit, with a new building block exhibiting a quantum mechanical behaviour, notably the qubit. From theoretical arguments, this striking conceptual difference could lead to an exponential enhancement of the computational power of quantum machines with respect to traditional computers. The surge in the application of quantum computing technologies has also led to a fruitful relationship with HPC resources, pioneering the way to a radical paradigm shift in computational techniques.

CINECA is ready to face these computational challenges and drive this process by leading the development of a national quantum computing hub alongside the existing HPC center. The core idea behind this effort is to push and facilitate the Italian community, both public and private, to embrace the new computing paradigm proposed by quantum computing and become a leader in this field over time.

In order to achieve such goal, CINECA intends to operate at different levels, setting itself to various objectives:

Study, assess and monitor the progress of this state-of-the-art, promising technology fostering interactions between quantum computing experts;

Support research by providing access to

both HPC computational resources and quantum devices that will be made available to researchers thorough agreements with partner providers;

Test the applicability of the different quantum devices currently available by developing algorithms capable of addressing problems of great interest, such as optimization, chemical simulations and machine learning;

Develop hybrid protocols where quantum computing is used as a support and accelerator for Cineca's HPC Systems;

Facilitate the discussion between universities, research centers and private parties interested in quantum computing applications;

Assist private companies in their pivotal efforts to initiate the quantum computing transition;

Stimulate both HPC and quantum communities through a series of conferences, educational courses and technical reports on the usage of the latest quantum computing softwares and hardwares.

The journey of CINECA into the field of Quantum Computing starts in 2018. In 2021, CINECA took an important step towards the concrete realization of this vision through the foundation of its Quantum Computing Lab, whose role is to develop the above objectives and hence exploit quantum computing as a new paradigm for information processing and high-performance computation.

### 2021 TRAINING ACTIVITIES

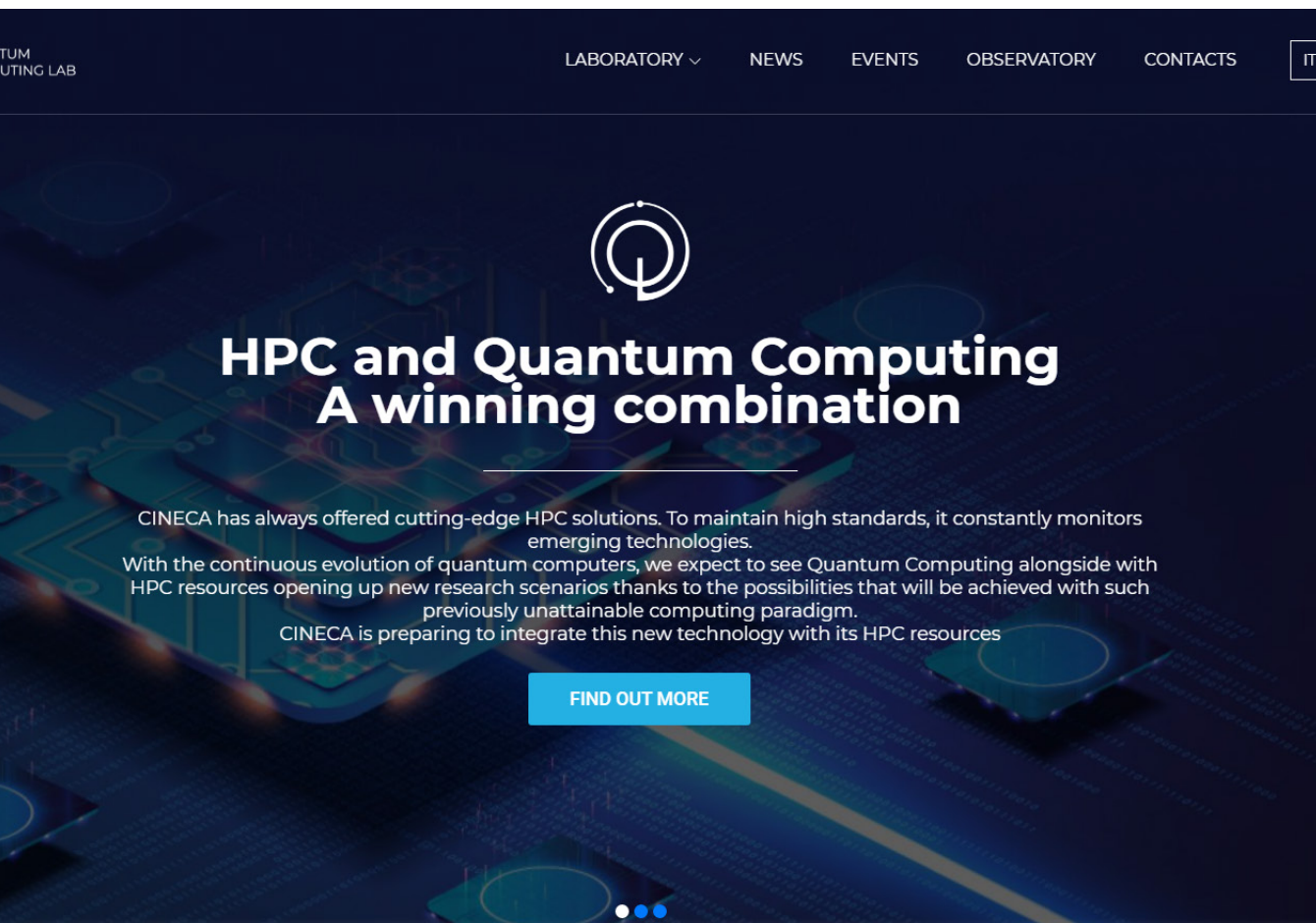
CINECA's training activities in the field of quantum computing have been going on since 2018, when the "HPC and Quantum Computing" workshop was held for the first time. HPCQC is a workshop dedicated to various purposes, including the dissemination of scientific results and encouragement for the creation of new interpersonal relationships between scientists and manufacturers. In addition

to the workshop, now in its third edition and ready to be able to return at the end of 2021, this year CINECA has also decided to organize its first school “Introduction to Quantum Computing”, an introductory school lasting 5 days, in English and sponsored by PRACE, which purpose will be to educate pupils in the correct use of quantum computing resources, explaining their use starting from the basics of the new computational science. Moreover, CINECA is also continuing to actively participate in the dissemination and support of scientific research, participating as protagonists in various webinars and conferences

## AVAILABLE QUANTUM COMPUTING RESOURCES

Starting from March 1, 2021, CINECA has made several hours of calculation on D-WAVE’s most advanced Quantum Annealers available free of charge to Italian universities and research centers through the IS CRA project. Moreover, thanks to an agreement with the European manufacturer Pasqal, it was also possible to start a university collaboration in order to test the promising capabilities of the innovative “neutral atoms” machines of the French company. The picture is completed by open-source emulators pre-installed on the “Marconi100” HPC machine capable of emulating “gate-model” quantum circuits. The future sees the creation and installation of hpc-ready emulators, able to make the most of the power of CINECA supercomputers, and the expansion of the quantum offer by inserting other promising technologies to allow students and researchers to conduct their research in the field of quantum computing.

Figure 24-25: screenshots of the QuantumLab website



## Quantum Computing as a Service: how to access the quantum resources made available by CINECA

From March 1st 2021, CINECA will make available, **free of charge**, computing hours selected from the most advanced and mature quantum computing technologies currently available on the market **to all Italian universities and research centers**.

In particular, it will be possible to access:

- **Quantum annealing resources**, the use of which is possible thanks to the **recent collaboration agreements between CINECA** and the leading company in the quantum annealing sector, **D-WAVE**, which currently owns **one of the most powerful quantum annealer in the world**. Through the agreement, CINECA will make calculation hours available to the Italian scientific community both on the **new D-WAVE computer, Advantage**, which had a QPU of **over 5000 qubits**, and on the **hybrid D-WAVE machines**, modern supercomputers that combine the power of quantum annealer with HPC resources, increasing the computing power of the machines to allow users to solve combinatorial optimization problems using a maximum of **one million binary variables** (equivalent, therefore, to a quantum annealer with one million qubits).

The agreement provides for a **total of 55 hours to be distributed monthly** between now and the end of the year (approximately **5 hours per month**). Each hour is **interchangeable with 20 hours of calculation on hybrid systems**. Quantum annealing resources will be awarded upon **submission of an IS CRA-C project**. It will be possible to request these resources **starting from the first IS CRA-C call in March**, whose deadline is scheduled for March 15, 2021; subsequently it will be possible to **renew** or open new projects **for each IS CRA-C call**, until the end of the hours available.

[For more information on D-Wave](#)

- **Quantum computing resources on Pasqal machines**. French start-up born in early 2019, Pasqal is a spin-off of the **Institut d’Optique (IOGS)**, which has been studying physics problems for years, in particular many-body problems **intractable with current HPC resources** available worldwide. The **first prototype** of a quantum computer with **neutral atom technology** was born in the IOGS laboratories as early as **2011**, thanks to the studies of Professor Antoine Browaeys, Professor Thierry Lahaye and their team of experts.

In this first **exploratory phase**, CINECA will give Italian universities and research centers the opportunity to **test one or two projects of high scientific relevance and applicative interest**, chosen through the requests received through the IS CRA portal, with the support of Pasqal. Applications must be received by **March 15, 2021**.

[For more information on Pasqal](#)

- **HPC resources for emulation of quantum computing environments**: to efficiently emulate a quantum computer, it is necessary to have a very powerful computer with a lot of RAM. From this point of view, CINECA has the most powerful public supercomputer in Europe (source: <https://www.top500.org/>), **Marconi100**. For this, the CINECA quantum computing team has created a new profile, called **profile/quantum**, which will allow all CINECA users to use a series of emulators pre-installed by us. To find out which emulators are currently installed on our systems and what their performance is, please refer to the following technical attachments:

[Report: emulatori Quantum su Marconi-100 \(HTML Version\)](#)

[Report: emulatori Quantum su Marconi-100 \(PDF Version\)](#)

For more information on how to access Quantum Computing resources, please refer to the [IS CRA project page](#).

Search ...

### LATEST ARTICLES

New demonstration of Quantum Supremacy from CINECA and D-WAVE

CINECA and D-WAVE Quantum Computing Resources in Italy

Webinar: Drug Discovery with Quantum Annealing

Cineca Quantum Computing Service

### WORKSHOP

Edition 2020

Edition 2019

Edition 2018

# HUMANITIES AND SOCIAL SCIENCES, DIGITAL ECOSYSTEMS

## Isabella d'Este Virtual Studiolo

Deanna Shemek  
University of California - Irvine

Cineca's VisITlab is bringing powerful computing and design skills to Isabella d'Este Archive's Virtual Studiolo (<https://isabelladeste.web.unc.edu/>), a cultural heritage project for study and exploration of one of the most significant domestic interiors of the Italian Renaissance. The magnificent rooms in the Gonzaga court of Mantua where Marchesa Isabella d'Este displayed her collection of art, books, musical instruments, and antiquities were celebrated during her lifetime as the signature space of one of Italy's most powerful and creative women. The original rooms within the Museo del Palazzo Ducale di Mantova have undergone architectural alteration, and what remains of Isabella's famed collection is now dispersed in museums around the world. The goal of the project is to reassemble Isabella's rooms virtually, opening them to anyone with access to the Internet, a smart phone, or a tablet device. Even richer immersive options will be available through the use of head-mounted VR devices. This Public Humanities project is conceived for researchers, students, educators, artists, and other curious explorers.

Taking our cues from the history of performative arts, we are assembling the Virtual Studiolo not as a re-creation, but as a historically-informed presentation. Given the high level of uncertainty surrounding the arrangement of the various pieces of Isabella's studiolo collection as well as the numerous questions that remain regarding both the timing and the acquisition of some of studiolo pieces, virtual reconstruction may provide a possible, and powerful, tool for exploration of the studiolo's development. As we continue to specify agreements with the museums that host pieces of the collection, the virtual versions of these objects, realized with photogrammetric technology, are being added to the virtual

environment and may be arranged at will by users.

2020-21 was a period of intense creativity and labor for the Virtual Studiolo's team. Thanks to the generous support of the National Endowment for the Humanities (<https://www.neh.gov/>), Cineca creators and US-based historians built on previous awards from the Samuel H. Kress Foundation, the University of California, and the University of North Carolina as well as the participation of the Kunsthistorisches Museum of Vienna and the Museo delle Ceramiche di Faenza, employing photogrammetry, 3D modeling, VR animation, and data programming to begin assembling our immersive environment.

We completed phase one of the project with a first-release 3D model of the studiolo and grotta for desktop navigation. The model features 3D assets representing several items from Isabella's collection; tools for measuring, lighting, and first- and third-person navigation; a floor plan and compass; a time bar tracing the arrival objects in Isabella's collection; an inspection gallery for close-up viewing and manipulation; and a powerful catalogue linked to historical data, bibliography, and other information for study. Also completed in this phase was our VR concept video, the preliminary testing ground for our fully immersive, Virtual Reality experience of the studiolo and grotta.

Now that a philologically correct 3D Web version for desktops and an emotional VR version for experiencing it with VR headsets are available, the team is ready for the next challenges. The work of this phase will serve as a basis for future enhancements of the project, both for scholarly users and for the general public.

# Memory and archive in motion

Francesca Zanella  
CSAC, University of Parma



Sottsass Virtual Exhibitor is a project started in 2016 and still undergoing thanks to the collaboration between CSAC of the University of Parma and VisitLab CINECA, with the contribution of Emilia Romagna Region and Parma Municipality.

The main goal of the project is the valorisation of visual and design archives through a multiplicity of actions: study of design histories through the archives, cataloguing and digitization of archival fonds, valorisation and dissemination through exhibitions activities and digital environments.

In 2016 CSAC launched its triennial research projects focused on the study and valorisation of a selection of funds and collections. The Ettore Sottsass jr. was one of the selected funds, because of its relevance: more than 14.000 drawings, images, maquettes, objects and documents donated in 1979 by the authors, and not yet catalogued online. In the same year started a PhD research project titled Memory and archive in motion financed by Emilia Romagna Region as part of the broader project Computing, Culture and Society: percorsi di studio e formazione nell'epoca dell'umanesimo digitale, coordinated by the University of Bologna. The focus of Memory and archive in motion was the reactivation of an archival 'treasury' such as the documentation of the work of Ettore Sottsass preserved by CSAC of the University of Parma, through an exhibition (Ettore Sottsass. Oltre il design 2017, CSAC, Parma) and a digital 'environment' realized with the collaboration of VisitLab Cineca. The keywords of the project

are Accesibility, Narratives and Digital Archives as the focus is the enhancement of Cultural Heritage, tangible and intangible, as the design archives are, thanks to the design of knowledge production and sharing centered on the research, the cataloguing, the digitizing and the design of a digital environment. Thanks to the 3D representation of the 2017 exhibition Ettore Sottsass. Oltre il design, held at the Cistercian church of the Valserena Abbey, venue of the CSAC, we are designing a gateway to the University of Parma online catalogue and to external resources, mixing heterogeneous contents, amplifying the 2017 narratives.

Through our new environment we can launch queries to extract information about a single drawing or item exhibited, reassembling the design process, or deepening the research through the online catalogue. But we can also look 'outside' the CSAC archives, consulting external resources and collections where the work of Ettore Sottsass is documented: the Centre Pompidou collection in Paris, and in company archives too, such as the Olivetti historical archive, the Poltronova archive or the Bitossi Museum, and so on.

With Sottsass virtual exhibition, CSAC and VisitLab CINECA are sharing their competencies on historical research and cataloguing, and on the other hand on computer graphic, designing an open source framework with a cross-medial approach with which to promote data interoperability in the cultural heritage domain.

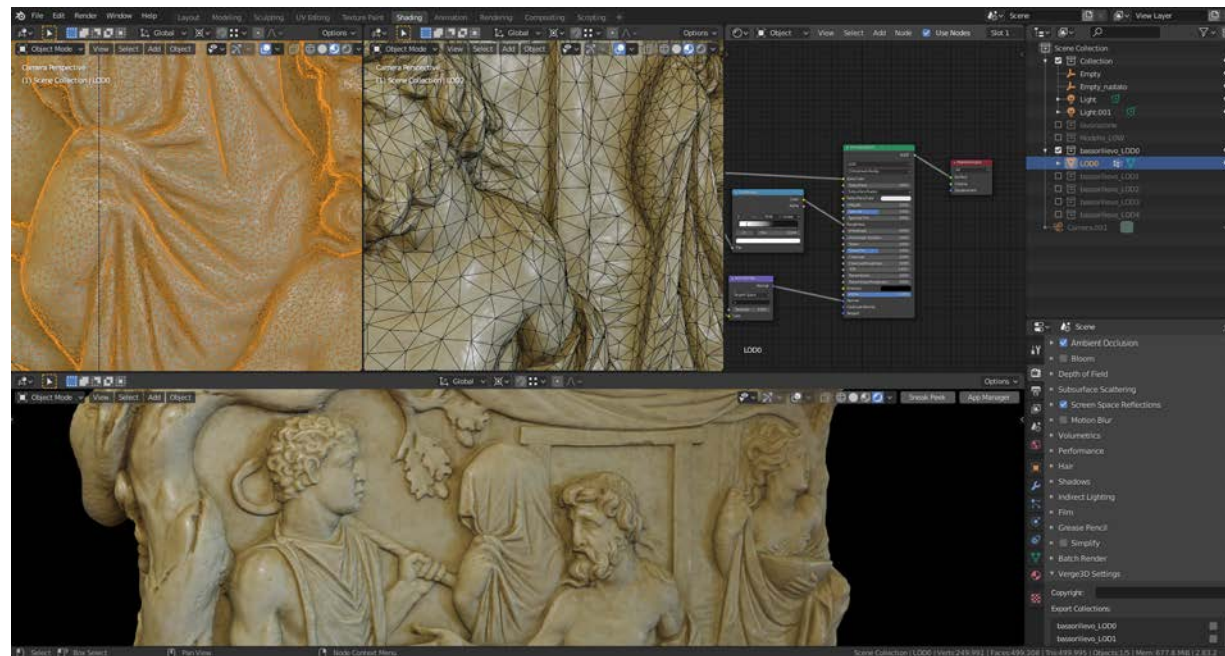


Figure 26  
Photogrammetric processing of a bas-relief belonging to the collection of Isabella d'Este. Ducal Palace of Mantua. Comparison view between a maximum definition model (LOD 0) and an intermediate definition model (LOD 2). In the latter, the fine details are reproduced using textures and normal maps. In particular, the normal maps were created starting from the high definition version, in order to display the fine details even in the absence of a high poly mesh.

Figure 27:  
Frame from the 360° VR video, available on the Cineca VisitLab YouTube channel. Use the QR Code to get there. For an immersive experience it is possible to enjoy the video with a VR cardboard.

Figure 28: CSAC Parma virtual environment



## TÜBKE MONUMENTAL

Franz Fischnaller

Digital Art Director at the Kunstkraftwerk Leipzig

“TÜBKE MONUMENTAL” is an immersive digital art show experience, based on the Monumental Artwork: THE EARLY BOURGEOIS REVOLUTION IN GERMANY, also known as the Peasants’ War Panorama. It depicts a circular panorama of the Battle of Frankenhausen, fought on 15 May 1525 during the German Peasants’ War, with no beginning or end, transcending the historical reality “into the timelessness of the apocalyptic creation of the world or its downfall”.

The painting, located in the Panorama Museum in Bad Frankenhausen, Germany, is one of the largest monumental panorama oil on canvas painting dome in the world. It spans 14 metres by 123 metres (1,700 square meter canvas), with more than 3000 characters with 75 key scenes. It was created between 1976 to 1987 by one of the founders of the Leipzig School, the East German painter, Werner Tübke (1929-2004), recognized as one of the most important East German artist and one of the most controversial painters of the second half of the 20th century.

It is a great figurative painting, an infinite landscape in which everything seems to be at the same time: the seasons, movement and calm, abundance and emptiness, life and death, poly- and monochrome, dreamlike fiction and tangible reality.

A high-resolution digitalization of this artwork has been done for the first time and stitched to a unique 110 Gigapixel image for the realisation of this project. TÜBKE MONUMENTAL will be exhibited in 2021 at the Digital media Center of the Kunstkraftwerk Leipzig in Germany.

The digitization of 1200 high-resolution image data was carried out by Centrica from Florence in close cooperation with the Panorama Museum in Bad Frankenhausen and KunstKraftwerk, Leipzig.

For the “TÜBKE MONUMENTAL” immersive Art show, the visualization in 3D of the Panorama-painting has been realised and adapted to the Maschinenhalle of the Kunstkraftwerk, an exhibit space of 30 x 15 meter and 8 meter heights. The digital content in size of 18000 x 2000 pixels is divided into 6 videos, that are distributed on 25 projectors and distributed on the walls through the Watchout system setup in the Kunstkraftwerk.

Franz Fischnaller has developed the content in 12 scenes, elaborating the painting together with the Cineca team in a creative combination in-between the 2D gigapixel images and the 3D visualization, to create the immersive show.

Creating a CG movie is always a challenge, and besides the traditional activities of modelling, texturing, animating, lighting, many technical issues were addressed. One example is the reconstruction in 3D of some portion of the landscape shown in the painting.

The Visit Lab team of Cineca has created and modelled a 3D landscape that surrounds in 360 degrees the viewer, but at the same time closely matching the 2 dimensional painting so that the painting in/out on the 3D model can fade seamlessly.

While there are techniques to recreate the 3D from pictures, here an automatic approach wasn’t feasible because the author cheated a lot with the perspective, thus an ad-hoc approach was developed using Blender and tools coming from the scientific visualization field.

A Gigapixel Immersive digital art show experience to be exhibited in the Kunstkraftwerk Leipzig, in Germany

Script, Storyboard, Art and Tech Direction by Franz Fischnaller

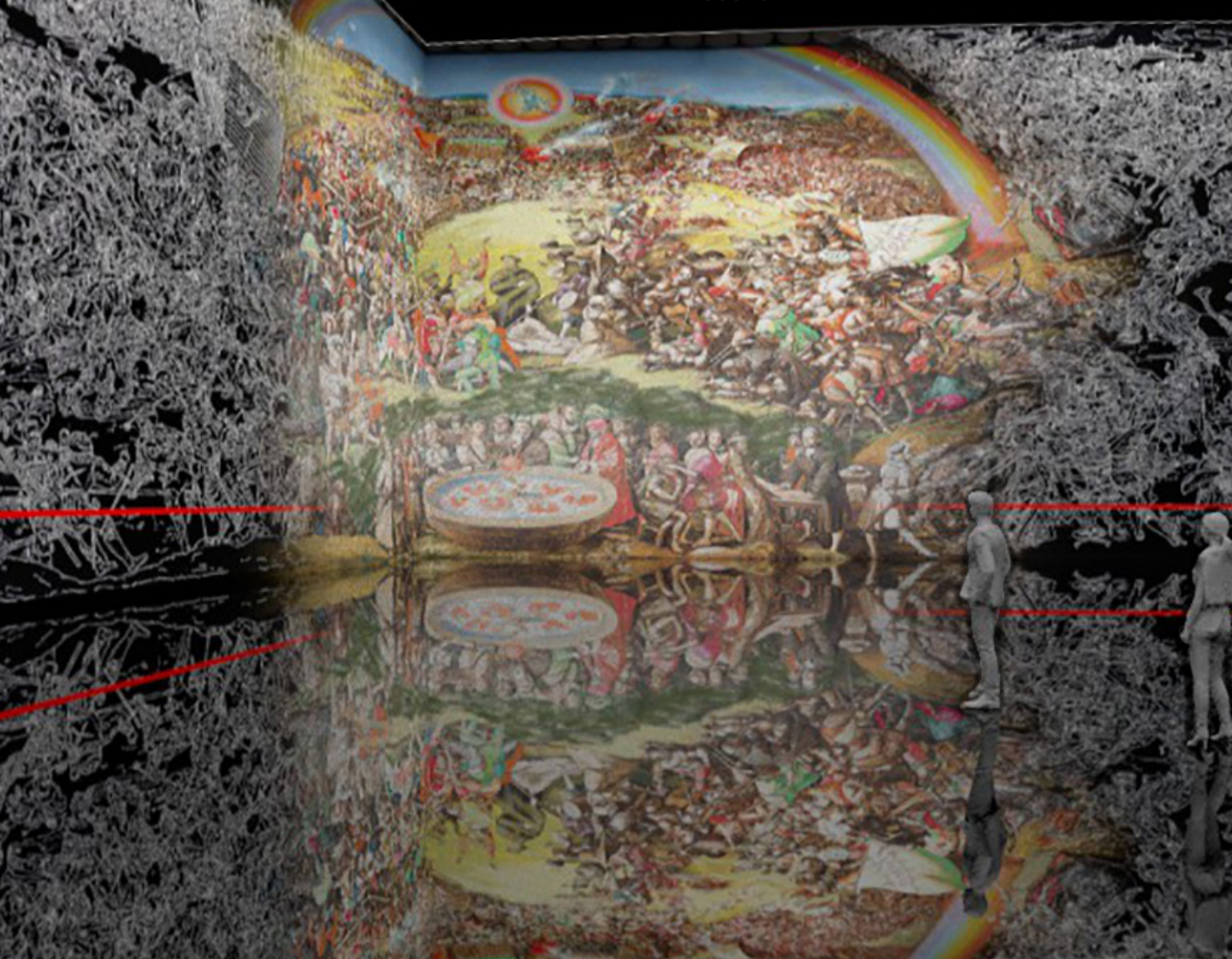
Implementation/Visualization: Antonella Guidazzoli, Silvano Imboden, Daniele De Luca, Maria Chiara Liguori, Visit Lab CINECA, Bologna, Italy

Music, Original Composition and musical arrangement by Steve Bryson, USA  
Editing & Postproduction: DIAC, Turin, Italy  
Digitalisation of the the Panorama Museum Bad Frankenhausen by Centrica srl, Florence, Italy  
HighRes 3D scanned image of the grapefruit

captured by CT scanner, courtesy by Dr Satre Stuelke and Prof. Jessica Rude  
3D modeling & texturing: Bingqing Dong  
Project Coordination: Prof. Markus Loeffler  
Curation by: Franz Fischnaller

A production of Kunstkraftwerk Leipzig with the generous support of: Ostdeutsche Sparkassenstiftung

Figure 29: Installation preview  
Figure 30: Preview of the second scene



# From I-Media-Cities to DARE Ravenna project

Antonella Guidazzoli, Maria Chiara Liguori  
Cineca

Recognised in 2020 by the European Innovation Radar as a market ready product [1], the I-Media-Cities platform [2], developed by Cineca in a Horizon 2020 project has, since then, been deployed for another European project: DARE - Digital Environment for Collaborative Alliances to Regenerate Urban Ecosystem in middle-sized cities [3].

The new instance relies upon Open Street Map for georeferencing contents, features controlled vocabularies in Italian, a customised virtual collection creator tool and metadata management of 3D assets. The new back end was presented during the events celebrating the opening of the DARE web portal [4].

[1] <https://www.innoradar.eu/innovation/34051>

[2] <https://www.imediacities.eu/>

[3] <https://www.darsenaravenna.it/>

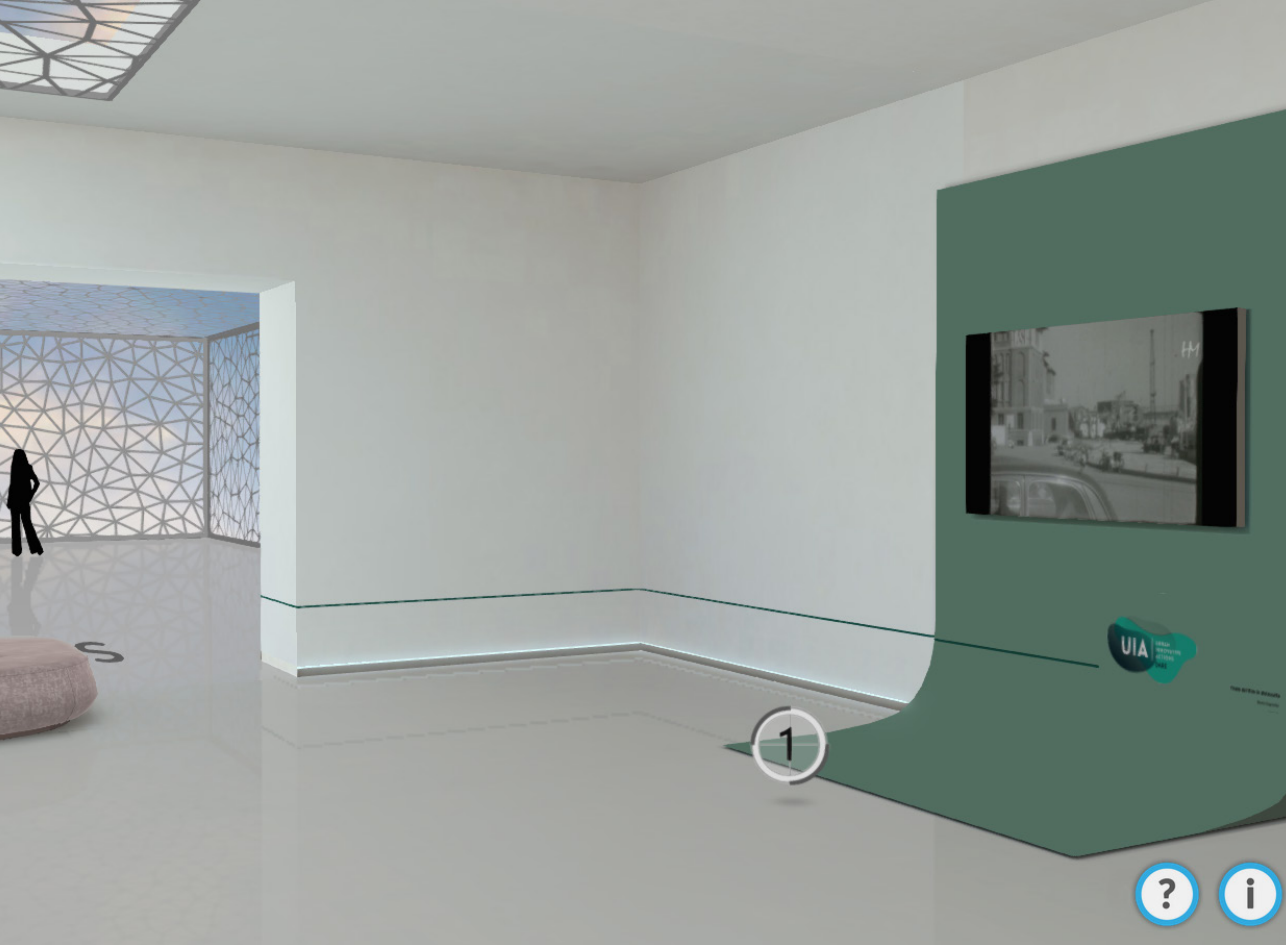
[4] <https://www.darsenaravenna.it/objects/la-vigilia-dell-imbarco crcsmn69s58e783q>



**EUROPEAN UNION**  
European Regional Development Fund

Urban Innovative Actions is an Initiative of the European Union promoting pilot projects in the field of sustainable urban development.

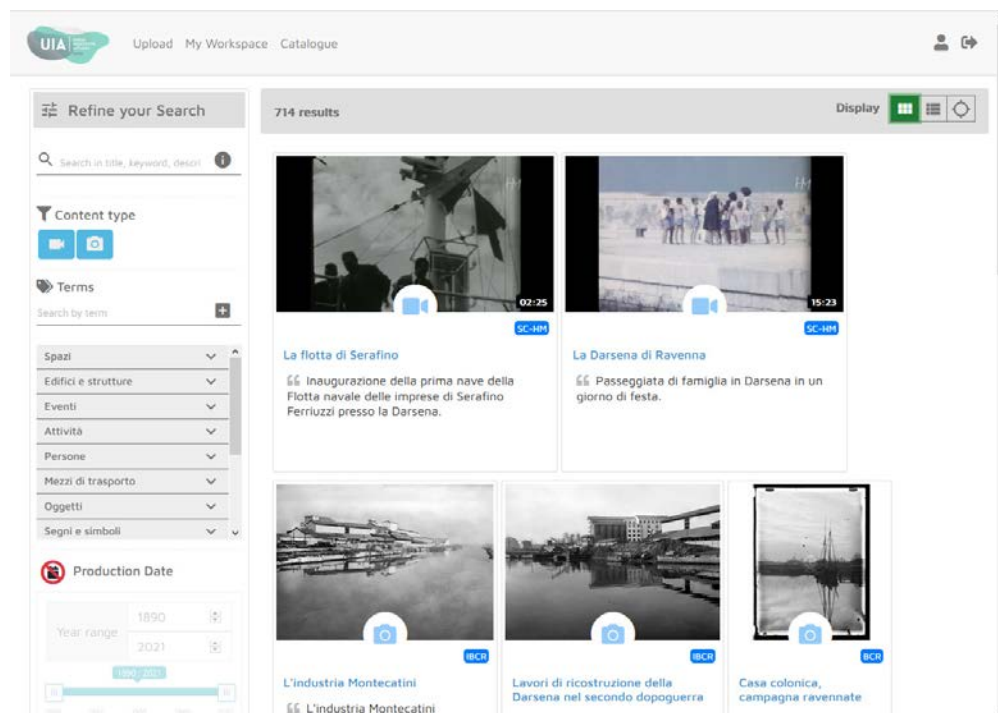




# COMPUTING PROVISIONING SERVICES

Figure 31: Virtual gallery

Figure 32: Screenshot of the website





# INDUSTRIES

## Supporting the National Industry through HPC-powered Proof-of-Concepts

Claudio Arlandini  
Cineca

The “Test before invest” approach is one of the pillar strategies that the EU recommends to Digital Innovation Hubs and Innovation Centres in supporting industries, especially SMEs, that want to understand the benefits in adopting new technologies like High Performance Computing, Data Analytics and Artificial Intelligence in their business. The advantages are obvious: experimentation with new digital technologies – software, hardware, business models – allows to understand new opportunities and return on investments, also including demonstration facilities and piloting.

It is an approach that Cineca found particularly useful in the past years, that led to long-standing collaborations and a proven impact, and that is realized mostly through the opportunities proposed by EU projects implementing proof of concepts experiments via cascaded funding mechanisms.

The year 2020 was particularly rich for this kind of activities in Cineca, thanks to the start of new projects that launched their first open calls in the second half of the year. The selected winning proof-of-concepts will be the care of our innovation-support team for 2021, together with those selected through additional open calls, and new projects like EOSC Future.

We provide here a quick overview of the end-user companies and subjects of the new proof-of-concepts.

### FF4EuroHPC

- Axyon AI – Leveraging HPC for AI & DL-powered solutions for asset management

(Fintech / AI)

- InSilicoTrials Technologies - Cloud-based HPC platform to support systemic-pulmonary shunting procedures (Bioengineering / Multiphysics simulation)
- Manta Group - HPC-based Multi-Physics and Multi-scale modelling of the autoclave process for aeronautical components (Aerospace / Multiphysics simulation)
- Green Tech Solutions - HPC-Based navigation system for Marine Litter hunting (Environment / AI)
- Aidro - TPOpology Optimization of Micro-channel Heat-Exchangers (Manufacturing / Multiphysics simulation / AI)

### EUHubs4Data

- GMATICS - Data-driven Model for the Analysis of Sea-state (Energy, Geophysics / AI)

### EuroCC Italy

- Optimad Engineering – Immerflow GPU porting (CFD / HPC)
- Axyon AI - AUDeep Learning-based software for time-series application (Fintech / AI)
- Arianet - QualeAria-Local: an air quality forecasting system on a national and micro-urban scale based on dispersion models and AI models.(Environment / AI)
- Cherrydata - Real-time Satellite Data Analysis (Environment / HPDA)
- Waterview – WeatherAI (Environment / AI)
- Aresys - HPC for Earth Observation Processing (Oil&Gas / Image processing)

# EuroCC bio

Cinzia Zannoni  
Cineca

The EuroCC-Italy Competence Center addresses the continuum of HPC, HPDA and AI technologies and knowledge, delivering an open service platform to facilitate access to: certified training material, on-line and on-site training programs; field work opportunities for young researchers and professionals; high HPC/HPDA/AI technologies and methodologies; proof of concepts to evaluate the impact of the technology on business and operations; computing resources; technology transfer actions between academia and industry.

EuroCC-Italy leverages on constructive collaborations between industry and academia, and relies on gender equal working groups spanning on different generations, where people can express their knowledge, skills and creativity, and make new ideas flourish from experience.

The EuroCC-Italy Competence Center open service platform, is a scalable project. It can be extended with the contribution of current and new partners belonging to the HPC/AI/HPDA national ecosystem.

Starting from industry business needs, the competence centers offers, to its users, resources and skills to set up proof of concepts, to assess the effectiveness of an HPC based solution. Success stories arising from the proof of concept results, help to sustain and increase awareness on the positive applications of HPC/HPDA/AI technologies to serve industry and people needs.

EuroCC-Italy belongs to the EuroHPC Competence Center Network, (link EuroCC Access) 33 competence centers, one for each country, exchanging knowledge,

experiences, use cases and success stories of HPC/HPDA/AI applications.

The pillars of EuroCC-Italy are the HPC competences and infrastructure available in CINECA (link), including pre-exascale system Leonardo(link); Associazione Big Data (link) a scientific network of Research Centers and Universities; BI-REX network with industry associations, and services for innovation, to reach SMEs in the manufacturing sector; corporates like Eni, Dompé and Leonardo Finmeccanica that have a long-lasting experience in using and applying HPC, HPDA and AI, in their research and business processes, willing to support innovations in SMEs belonging to their supply chain, and to the green economy, national and European ecosystem, through concrete application examples and needs. EuroCC-Italy is an open project to be extended in its due course to new partners belonging to the HPC/AI/HPDA national ecosystem.

The set up of the EuroCC-Italy Competence Center has received funding from the European High-Performance Computing Joint Undertaking (JU) <https://eurohpc-ju.europa.eu/> under grant agreement No 951732. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Germany, Bulgaria, Austria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, the United Kingdom, France, the Netherlands, Belgium, Luxembourg, Slovakia, Norway, Switzerland, Turkey, Republic of North Macedonia, Iceland, Montenegro.



This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under Grant Agreement 951732



# Training for industries

Eric Pascolo  
Cineca

Industrial processes and services are evolving fast thanks to innovative technologies and workflows that are revolutionizing traditional schemes. To keep up with these changes and match the increasing need of highly specialized education that enterprises of all sizes face nowadays, the Italian National Competence Center (NCC) set up three training levels will be carried out for companies:

1. “Curiosity and Interest”: is the first approach on HPC, HPDA and AI technologies
2. “High performance innovation”: aims to provide company personnel with skills on the technologies to be applied in the innovation processes of services and products
3. “Fieldwork”: internship program dedicated to young people and give them the possibility to develop skills by participating to the projects of the competence center.

The first level events, which represent the first approach to these innovative technologies, will be carried out in collaboration with other training institutions, even within existing educational paths.

The second level courses, aimed at satisfying specific training needs, will be created according to the needs that emerged from the national mapping of skills or from specific courses created in agreement with the companies that will be protagonists of the innovation projects. The internship program will allow young graduates or PhDs to deepen their

knowledge of HPC / Big data / AI technologies by working on the projects developed within the competence center.

The training activities have officially started in May 2021 with 3 “Curiosity and Interest” courses.

“Supercomputing for Industries and Services” is 24-hour course, delivered in 6 meetings, has been organized in collaboration with the Italian Competence Center on Industry 4.0 BI-REX, also part of the NCC Italy network, and aimed at showing the key importance of supercomputing for companies: from the speedup of designing of prototypes to the ability to deal with the increasing volume of data, from optimization of processes to artificial intelligence algorithms.

On the 18th and 19th of May the Italian NCC delivered a 12-hour laboratory within the IFTS course organized by IFOA Bologna “Expert on data analysis and visualization”. During this lab we gave students the opportunity to familiarize with high performance computing and gave them access to CINECA supercomputing resources for a hands-on session on parallelization of algorithms.

In collaboration with IFOA, the Italian NCC delivered a 4-hour workshop entitled “HPC for innovation” that targeted people working in SMEs to show them the benefits of supercomputing and the basics of HPC infrastructures.



# Supporting ENI HPC Infrastructure

Cinzia Zannoni  
Cineca

For over 20 years, Cineca has been collaborating with Eni (the most important Italian Energy Company) in the set up of the company HPC ecosystem, implementing the most advanced algorithms developed by Eni researchers, developing applications and managing the company HPC infrastructure. From 2007 till 2012, Cineca hosted Eni supercomputers in its own datacentre. In 2013, Eni set up its own HPC infrastructure in its Green data Centre (GDC) and Cineca continues to manage Eni's HPC infrastructure in GDC.

On Cineca side, the collaboration involves different teams and people with different skills, able to work in research projects, and in the development and management of HPC applications and systems.

In march 2020 Eni set up its fifth generation High Performance Computing System, HPC5, located in the Eni's Green Data Centre (GDC). The system has a peak performance of 52 PFlops, it is currently ranked in the ninth position of the June 2021 TOP500 List, emerging as the most powerful HPC system World Wide installed by an industrial company. HPC5, associated to the other HPC4 system, takes the whole Eni's HPC infrastructure to the availability of 70 PFlops.

This infrastructure provides a strategic support to Eni's digital transformation process along its whole value chain, from exploration to reservoir management to the analysis of all the data generated during the operations.

In 2021 Cineca collaborated with Eni in the evaluation of novel HPC technologies, for the upgrade of HPC4, expected in June 2021, and on the evolution and maintenance of HPC applications, as well as on new Eni research projects developing new applications for reservoir simulation, data analysis and interpretation.

In the course of 2020, specific effort was devoted to the consolidation of the software development and testing process, to support the continuous development and engineering of new functionalities and algorithms developed in Eni research project. Cineca is currently involved in many different projects with Eni concerning:

- the parallelization, optimization and portability of hybrid codes on different platforms,

- the evolution of the HPC production environment,

- the development and maintenance of production applications,

- the development of new applications, new components and new solvers, to be integrated into the production applications.

The collaboration with Eni allowed Cineca staff to face the problems and the typical needs of industrial production, to grow in the understanding of the specific domain and in the ability to work in highly specialized and effective multidisciplinary teams.

## FUTURE TRENDS

Cineca high performance computing (HPC) infrastructure is uniquely positioned to promote and foster the national and European scientific application ecosystem.

# What's next: evolution of the computing infrastructure

Mirko Cestari, Elda Rossi  
Cineca

Cineca's mission is to support scientific research by providing access to advanced computational systems. Technologies such as Deep Learning, Quantum Computing, accelerated computing, will gradually revolutionize the way scientists conduct numerical experiments. It is part of our mission to design Cineca systems by incorporating technology changes and ensuring users to exploit them in a profitable manner. Cineca high performance computing (HPC) infrastructure is uniquely positioned to promote and foster the national and European scientific application ecosystem. Some important changes are foreseen in the HPC infrastructure, leading to an increased diversity of hardware available and allowing to better address the variety of user workloads.

## DGX

Three NVIDIA DGX A100 appliances complete the Marconi100 supercomputing system. They will start the production phase in early 2021. This infrastructure – given the high number of GPUs per server – is particularly devoted to accelerating deep learning and artificial intelligence applications. Each DGX-A100 includes 8 Ampere-based A100 accelerators, two 64-core AMD processors (AMD Rome 7742), 1 TB of RAM memory. All servers are interconnected through the NVIDIA Mellanox HDR InfiniBand. The DGX infrastructure will be part of the ISCRA offer.

## Galileo100

In 2021 it is foreseen a significant upgrade of the tier-1 system Galileo. The new system – Galileo100 – will start the production in the third quarter of the year. It is a DELL system, co-financed by the Italian Ministry of University and Research and the European Commission through the ICEI-Fenix project. All the computing nodes are based on two Intel CascadeLake 8260 processors at 2.4 GHz, for a total of 48 cores per node, 384 GB memory. Galileo100 architecture design followed the guidelines and agreements

of the ICEI-Fenix project (<https://fenix-ri.eu/>) and will allow to share among main HPC European partners the strategy for the provisioning of federated services, including computing resource provisioning, data storing and sharing, cloud and interactive computing. The latter aims to provide users with the ability to interactively elaborate, process and visualize data, as well as the ability to drive non-standard HPC workloads that – for instance – want to process live-data continuously coming from data sources. The interactive computing service is clearly relevant in the attempt of closing the gap between infrastructure dedicated to high-performance computing and infrastructure dedicated to data analysis and AI. In this regard, Cineca has taken the leading role of conducting the tender procedure on behalf of the partners of the European ICEI-Fenix project. The Cloud infrastructure and its integration with the rest of the HPC ecosystem will be enhanced with latest generation servers, dedicated high-performance full flash storage, and a new installation of the most recent stable Openstack version. This will allow to provide Cloud services that complement the HPC offer and better serve the requests coming from all the scientific communities.

## Leonardo

Leonardo is a precursor-of-exascale class system, one of the three funded by the European Commission with the EuroHPC programme. The tender for the acquisition of the system started in 2019 and completed in late 2020. The delivery, installation and acceptance of the system is expected in the first half of 2022, in order to start the production in the second half of 2022. Leonardo is based on Atos XH2000 platform technology and will provide to users two main partitions: “Booster” with GPU accelerators and “DataCentric” with more traditional X86 technology. The technological continuity from Marconi100 to Leonardo constitutes an essential asset for all scientific communities and private



Figure 1-2: Leonardo overview

## Leonardo specifications



### Booster

- 116 BullSequana XH2000 Direct Liquid cooling racks
- 240 PFLOPs HPL Linpack Performance (Rmax)
- 3456 computing node servers
- 13824 NVIDIA Ampere next-gen GPUs
- 884 TB of HBM2



### Data Centric

- 16 BullSequana XH2000 Direct Liquid cooling racks
- 9 PFLOPs HPL Linpack Performance (Rmax)
- 1536 computing node servers
- 3072 Intel next-gen CPUs
- 786 TB DDR5 Memory
- 5.8 PB of local NVM

[www.cineca.it](http://www.cineca.it)



companies. In fact, users can focus their efforts on porting the applications on accelerated heterogeneous systems knowing that they can count on Leonardo hosting agreement until 2027. With Leonardo, Cineca users will be able to access a system - that is the natural evolution of Marconi100 - and is equipped with new generation NVIDIA Ampere accelerators. In principle, therefore, it will be possible for leading projects to access Leonardo resources after demonstrating the efficiency and ability to exploit Marconi100's computational resources. To this end, the synergy between the HPC data centers hosting Marconi100 and Leonardo system will play a significant role: the increase in computing power will lead to the production of large amounts of data and will entail the need for communication channels that allow high bandwidth data transfers. In this regard, Cineca will start in 2021 all the activities required to enhance the performance and security of the high-performance communication network involving both the Casalecchio and Tecnopolo data centers.

## **Cybersecurity and new technologies**

Cineca has recently encompassed a series of activities aimed at strengthening the

cybersecurity of the supercomputing infrastructure. An overall reconfiguration of the HPC service provisioning is under way. In this regard Cineca is running a proof of concept to include in the infrastructure a bastion host, a data mover, and to provide to users a secure 2-factor authentication. Besides, a feasibility study for data at-rest encryption and data segregation will be carried on by the ICEI-Fenix infrastructure (Galileo100). Another strategic objective is the scouting of the new technologies available on the market to define, together with the scientific communities and end users, the requirements of the systems for the next cycle of evolution. In fact, Cineca participates in the main national and European projects with architectural and infrastructural development objectives (EPI, EPI2, EuPEX, Regale) to define the characteristics of the next Exascale systems. In this regard, Cineca believes that is important to acquire a prototype to study the new emerging processors architecture. For example processors equipped with SVE (Scalable Vector Extension) instruction set and HBM (High Bandwidth memory). In this way we will monitor and evaluate the most promising technologies for the future of HPC.



**EuroHPC**  
Joint Undertaking