INAF

ISTITUTO NAZIONALE DI ASTROFISICA
NATIONAL INSTITUTE FOR ASTROPHYSICS

Delibera n. 29/2016, Verbale CdA n. 03/16

Oggetto: approvazione dello schema di "Memorandum of Understanding for the establishment of the EU-T0 Data Research and Innovation Consortium".

IL CONSIGLIO DI AMMINISTRAZIONE

VISTO il Decreto Legislativo 4 giugno 2003, n. 138 di riordino dell'Istituto Nazionale di Astrofisica, pubblicato nella G.U. del 19 giugno 2003, n. 140;

VISTO il Decreto Legislativo 31 dicembre 2009, n. 213 di riordino degli enti di ricerca in attuazione dell'articolo 1 della legge 27 settembre 2007, n. 165;

VISTO lo Statuto dell'INAF, entrato in vigore il 1° maggio 2011, e ss.mm.ii.;

VISTO il Disciplinare di Organizzazione e Funzionamento dell'Istituto Nazionale di Astrofisica, entrato in vigore il 23 luglio 2012, e ss.mm.ii.;

CONSIDERATO che l'INAF, il CERN (Conseil européen pour la recherche nucléaire), il CIEMAT (Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas), il DESY (Deutsches Elektronen-Synchrotron), l'IFAE (Institut de Física d'Altes Energies), l'IN2P3 (Institut national de physique nucléaire et de physique des particules), l'INFN - Istituto Nazionale di Fisica Nucleare, il KIT (Karlsruher Institut für Technologie), il STFC (Science and Technology Facilities Council), il CEA - IRFU (Commissariat à l'énergie atomique et aux énergies alternatives - Institut de recherche sur les lois fondamentales de l'Univers) e SURFsara (SURF Stichting Academisch Rekencentrum Amsterdam) finanziano e forniscono alla comunità scientifica su scala globale ingenti risorse di calcolo per l'elaborazione dei dati provenienti dalle collaborazioni scientifiche esistenti;

RAVVISATA la volontà delle predette istituzioni di istituire a tale scopo un Consorzio denominato "EU-T0" per la creazione di un hub virtuale Tier-0 in grado di coordinare gli sviluppi tecnologici e l'integrazione dei dati e dei centri di calcolo esistenti;

VISTO lo schema di "Memorandum of Understanding for the establishment of the EU-T0 Data Research and Innovation Consortium";

VISTA la nota trasmessa per le vie brevi in data 3 febbraio u.s. dal Responsabile dell'Unità Scientifica Centrale VI "ICT - Information and Communications Technologies", con la quale il dott. Riccardo Smareglia sottolinea l'opportunità di addivenire alla stipula del Memorandum in questione;

ACQUISITO il parere favorevole del Direttore Scientifico alla sottoscrizione del MoU;

PRESO ATTO che l'approvazione di tale Accordo non comporta, in questa fase, alcun onere finanziario per l'Istituto;

RAVVISATE pertanto l'opportunità e la necessità di provvedere;

INAF

DELIBERA

con voto unanime dei presenti, espresso nei modi di legge:

- di approvare, nel testo allegato alla presente deliberazione in modo da formarne parte integrante e sostanziale, lo schema di "Memorandum of Understanding for the establishment of the EU-T0 Data Research and Innovation Consortium";

- di autorizzare il Presidente alla sottoscrizione dell'Accordo.

Roma, 21 marzo 2016

Il Segretario

Il Presidente

MEMORANDUM OF UNDERSTANDING

For the establishment of the
EU-T0 Data Research and Innovation Consortium

among

1. Centre National de la Recherche Scientifique (CNRS), 3 Rue Michel-Ange 75794 Paris, cedex 16, France

and

2. Istituto Nazionale di Fisica Nucleare (INFN), Piazza dei Caprettari, 70, 00186 Rome, Italy

and

3. Science and Technology Facilities Council (STFC),
Polaris House North Star Avenue, Swindon SN2 1SZ, United Kingdom

and

4. Deutsches Elektronen-Synchrotron (DESY) Notkestr. 85, 22607 Hamburg, Germany

and

5. Karlsruhe Institute of Technology (KIT), Kaiserstr. 12, 76131 Karlsruhe, Germany

and

6. Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT)
Avda. Complutense, 40, 28040 Madrid, Spain

and

7. Institut de Física d'Altes Energies (IFAE) Edifici Cn, Universitat Autònoma de Barcelona (UAB),
E-08193 Bellaterra-Barcelona, Spain

and

8. CEA Institut de recherche sur les lois fondamentales de l'Univers (IRFU), Saclay Bât 141,
F-91191 Gif sur Yvette Cedex, France

and

9. European Organization for Nuclear Research (CERN), Route de Meyrin,
1211 Geneva, Switzerland,

and

10. SURFsara, PO Box 94613, 1090 GP Amsterdam, The Netherlands

and

11. Istituto Nazionale di Astrofisica (INAF), Viale del Parco Mellini, 84 - 00136 Rome, Italy

Hereinafter collectively referred to as the "Parties" and individually to as a "Party".

Preamble:

WHEREAS:

EU-T0 is a Consortium of funding agencies and research organizations which between them fund and provide very large scale computing and data processing resources for a set of scientific communities. The *Statement of Intent* (hereinafter referred to as "SID") of the Consortium is attached as Annex 3 and sets out the objectives of EU-T0.

THEREFORE THE PARTIES HAVE AGREED THE FOLLOWING:

ARTICLE 1
Purpose

The purpose of this Memorandum of Understanding (hereinafter referred to as "MoU") is to establish the EU-T0 Data Research and Innovation Consortium (hereinafter referred to as the "Consortium").

The Consortium embodies the common intention of the Parties to cooperate in areas of mutual interest in the above domain and to use reasonable efforts to carry out activities to achieve the objectives of the Consortium which are set out in Annex 3.

Participation in the Consortium by other similar funding agencies and research organizations is encouraged.

The provisions of this MoU are not intended to have legal effect, so they are neither legally nor financially binding on the Parties. The Parties agree that nothing under this MoU can be construed as the intention to create any liabilities between the Parties.

ARTICLE 2
Governance

The Consortium will develop and maintain a governance structure (including voting rights, rules, and membership conditions). The initial structure which has been agreed by the Parties is set out in Annex 1 to this MoU. This structure may evolve on the basis of future revisions of the Consortium's needs and/or vision.

ARTICLE 3
Activities

The initial activities (hereinafter referred to as "Activities") that are foreseen are specified in the *SID document* in Annex 3. The Activities to be undertaken may be modified and augmented from time to time subject to the agreement of the Consortium.

The Parties may agree to collaborate on common Activities or specific joint R&D projects. The execution of any common Activities or specific R&D projects will be governed by separate written agreements concluded between the Parties involved. Those agreements will, inter alia, included provisions concerning their scope, contributions (in kind or financial), intellectual property and publications rights.

ARTICLE 4
Resources

The Parties to the EU-T0 Consortium already fund, or provide, large scale computing resources for several scientific disciplines. These resources are listed in Annex 2. The aim is to make best use of these resources both to the benefit of the scientific communities already served by the Parties and to extend the reach to new scientific communities, by collaborating on the common development and provisioning of federated services. It remains up to each Party to determine what fraction of these computing resources will be made available as "common resources" to disciplines other than those already supported.

The large scale computing resources concerned shall remain the property of the Parties which own these facilities. Consequently their operation and management remain the responsibility of their respective owners.

The Parties to the Consortium are expected to make reasonable efforts to carry out the Activities in which they are involved (see Article 3), through using existing staff as well as through new staff recruited for new projects, supported either through their own or other funding sources. The Parties accept that whilst this is a best efforts commitment, the success of the Consortium depends upon it being achieved.

A common fund will be set up for the coordination and support activities needed by the Consortium, including meetings and material expenses for outreach and dissemination, to which all Parties will contribute. The definition and terms of any financial or in-kind contribution from Parties will be part of a further written agreement to be concluded among all Parties.

ARTICLE 5
Entry into force, duration and termination

This MoU will come into force on the day following the date of the last signature by the Parties. It will remain in force for two (2) years, corresponding to the expected period for the development of a strategy and implementation plan for EU-T0. The validity of this MoU will be extended automatically for successive one-year periods, unless and until terminated by joint agreement of all Parties.

Except as otherwise agreed by the Parties, and subject to Article 3, termination of this MoU shall be on completion of agreed Activities, or specific R&D projects, and shall be based on a best endeavours basis or as set out in any future addenda to this MoU.

Any Party may withdraw from this MoU at any time upon a six (6) months prior written notice sent to other Parties. In such a case and unless agreed otherwise by all the Parties, the MoU will remain effective toward the withdrawing Party.

ARTICLE 6
Amendment

Any amendment to this MoU will be agreed upon by consensus of the Consortium Board and ratified by written agreement of all Parties.

The MoU will remain open to further legal entities to join, upon unanimous agreement by all Parties, and subject to the signature of an amendment.

ARTICLE 7
Settlement of disputes

In the event of any dispute or difference arising out of the interpretation or implementation or application of the provisions of this MoU, the Parties will settle it amicably through consultation or negotiation that may result in appropriate amendments to this MoU.

ARTICLE 8
Confidentiality and Intellectual Property

"Confidential Information" means any and all information and/or data in any form and of any nature whatsoever – including, but not limited to, all written or printed documents, samples, models, software, and/or information whether or not patentable - disclosed by a Party to one or more other Party(ies) under this MoU and which is explicitly marked as confidential. Oral or visual information shall be designated as confidential on disclosure and then confirmed in writing by the disclosing Party to the receiving Party(ies).

For the duration of this MoU, the Parties will share information and technology in a way that respects and preserves intellectual property rights and agree to abide by the following principles:

(i) The Parties treat any Confidential Information for the duration of this MoU and for a period of five (5) years after termination of this MoU, as strictly confidential.
(ii) The Parties take all appropriate steps to safeguard the Confidential Information.

The obligation of confidentiality shall not apply to Confidential Information that can be proved:

- To belong to the public domain as a result of publications, or
- To fall into the public domain without any fault of the receiving Party, or
- To have been disclosed to the receiving Party by a third party, without the obligation of confidentiality, or
- To have already been known to the receiving Party prior to disclosure by another Party, or
- To be the result of work of employees of the receiving Party, who had no access to the disclosed information, or
- To have to be disclosed due to a legal obligation or an order by court or an authority.

The Parties will not make any warranty with regard to the Confidential Information disclosed, and in particular they will warrant for neither the correctness and usability of the Confidential Information disclosed nor the non-infringement upon rights of third parties.

Activities concerning the execution of specific R&D projects, including rights to confidentiality and intellectual property, will follow these principles but be governed by separate collaboration agreements concluded between the Parties involved and added to this MoU.

EU-T0 Consortium Governance

The initial EU-T0 governance structure will comprise:

(i) A Consortium Board (CB).
(ii) A Spokesperson.
(iii) An Executive Board (EB).

Consortium Board

The Consortium Board (CB) will consist of one representative from each Party signatory of EU-T0. Representation will be at funding agency level or equivalent.

Each Party will designate its representative (hereinafter referred to as "Member") authorised to negotiate and decide on all matters covered under this MoU. In cases where a Member is unavailable for a CB meeting, each Party can designate a substitute for that meeting.

The CB Chair is elected by the Board from among the Parties' representatives, for a term of two (2) years renewable once. On election, the Chair ceases to be a representative of his/her Party and can be replaced on the CB by another Member chosen by his/her Party.

A scientific/technical expert may be invited by each Member of the CB meetings in an advisory role, within a limit of one (1) expert per Party.

The CB is the top-level decision making body and is the dedicated forum for discussions between the Parties. The CB will provide strategic direction and oversight and decide by consensus, providing a quorum of at least two thirds (2/3) of voting Members are in attendance, on all matters concerning the implementation of the Purpose and Activities of this MoU, including but not limited to the following matters:

- Development of a future roadmap and strategy for EU-T0.
- Approval of the plans and Activities to implement the strategy.
- Oversight of progress and performance in delivering agreed plans and Activities.
- Nomination of expert panels to provide a scientific evaluation, when necessary.
- Approval of the accession of new Party to the consortium.

The CB will be convened whenever necessary, but at least once a year. Meetings will have a formal agenda and will be minuted. The agenda of each meeting will be sent to the Members at least seven (7) calendar days before the meeting.

Spokesperson

The Spokesperson is responsible for the coordination of the Activities of the Consortium.

The Spokesperson has as primary responsibility pursuing the purpose set out in the SID, and for the organisation and execution of the Activities.

The Spokesperson represents the Consortium on behalf of the CB, and is the primary point of contact for external bodies.

The CB appoints the Spokesperson for a term of two (2) years.

Executive Board

The Executive Board (EB) is the body in charge of steering and ensuring progress on agreed EU-TO Activities in a timely way.

Members of the EB will be representative of the Parties to the Consortium.

1. Members of the EB are proposed by the Parties and appointed by the CB
2. The EB meets 3-4 times per year.
3. The EB receives progress reports on agreed Activities and gives feedback, which takes into account strategic, funding and "political" aspects.
4. The EB will be responsible for:
    - The identification of future opportunities for partnerships or collaborative projects in scientific computing, data processing and e-infrastructure area including resources required
    - The preparation of plans and Activities to implement the agreed strategy
    - The implementation of the approved plans and decisions of the CB
    - The monitoring and reporting of the agreed Activities to CB, including collaboration on specific common Activities or R&D projects
    - The representation of the Consortium and its interactions with other organisations under specific mandate from the CB
    - The establishment and operation of any coordination and support activities under specific mandate of the CB
    - The preparation of clear proposals for decisions to be made by the CB.

The Executive Board may assign coordinators to those Activities that do not (yet) have their own project structure.

Annex 2 to EU-T0 MoU

e-Infrastructures and related computing centres of supporting EU-T0 Members

The EU-T0 Parties own existing large-scale physical infrastructure, as follows:

- CERN is the world's largest particle physics lab and home of the Large Hadron Collider (LHC), the world's most powerful particle accelerator. CERN provides research facilities for HEP researchers across the globe. LHC experiments will run up to 1 million computing tasks per day and generate around 15 petabytes of data per year over a 10- 15 year period. Analysis of the data relies on the combined resources of some 200 computer centres world- wide. CERN leads the World- wide LHC Computing Grid project (WLCG), to provide computing resource to store, distribute, analyse and access (in near real- time) LHC data for a community of more than 10,000 physicists worldwide.

- CIEMAT (http://www.ciemat.es) is a public research body assigned to the Spanish Ministry of Economy and Competitiveness under the Secretariat of State for Research, Development and Innovation. CIEMAT is involved in basic and applied science and technology projects in the fields of energy, environment and the technologies related to them, at national and international levels. CIEMAT collaborates as well with other institutions, universities and business in the sector, to transfer the knowledge and technology that it has generated. The Scientific Computing Unit works on advanced scientific computing techniques, providing infrastructure and support to the huge and complex requirements of the aforementioned projects, using distributed computing, grid and cloud technologies. The Unit participates in the Port d'Informació Científica (PIC), a joint undertaking with IFAE (see PIC description under IFAE). Past and current projects include WLCG, EU DataGrid, Enabling Grids for e-Science (EGEE), EGI-Inspire, EELA, EDGeS and the nuclear fusion related EUFORIA project.

- DESY, the "Stiftung Deutsches Elektronen-Synchrotron DESY" in Germany, is one of the world leading laboratories for particle physics and photon science and has long-standing experience in the design and operation of accelerators for high energy physics and synchrotron radiation research. DESY operates one of the most brilliant synchrotron radiation sources world wide (PETRA III), two Free Electron Lasers (FLASH and FLASH-II) and will operate the European XFEL accelerator, a project from the ESFRI list. DESY holds a 50% share of the European XFEL, an ESFRI-project. As a member of the Helmholtz Association in Germany, DESY is a non-profit research organization funded by public funds from the government and the states.

- IFAE, the Institut de Física d'Altes Energies (http://www.ifae.es/), is a consortium of the Universitat Autònoma de Barcelona (UAB) and the Government of Catalonia, part of CERCA, the public Catalan network of research Institutes (www.cerca.cat). IFAE conducts experimental and theoretical research at the frontier of fundamental physics, including Particle Physics, Astrophysics and Cosmology. IFAE also works at the cutting edge of detector technology, applying its know-how to Medical Imaging, Data-Intensive Computing and projects in other research and engineering domains. IFAE was granted in 2012 the Spanish Government's Severo Ochoa award. IFAE's Experimental Division currently participates in ATLAS at the Large Hadron Collider, in the T2K neutrino experiment, and in the MAGIC and CTA Cherenkov gamma-ray telescopes. In the field of sky-surveys searching for dark energy, IFAE leads the Physics of the Accelerating Universe (PAU) survey, and participates in DES and in ESA's EUCLID mission. IFAE manages Spain's largest scientific data centre, the Port d'Informació Científica (PIC). Created in 2003, PIC is a joint undertaking of the Spanish and Catalan governments through CIEMAT and IFAE, with additional support from UAB. PIC has been designated by the Spanish government as its LHC Tier-1 centre, and it is the main (Tier-

0) data centre for the MAGIC telescope and the PAU dark energy survey, as well as some of the largest sky-survey catalogue simulations performed on supercomputers. PIC maintains a transversal innovation activity with many significant results over the years, related to software, hardware, monitoring and energy efficiency. It also seeks collaborations to transfer the knowledge and technologies to other fields, including aerodynamic design and medical imaging analysis.

- The CNRS - Centre National de la Recherche Scientifique (National Centre for Scientific Research) is a government funded research organization, under the administrative authority of France's Ministry of Research. As the largest fundamental research organization in Europe, CNRS carried out research in almost all fields of knowledge, through its eight research institutes among which the National Institute of Nuclear and Particle Physics (IN2P3). Founded in 1971, the aim of IN2P3 is to promote and unify research activities in the fields of nuclear physics, particle and astroparticle physics. It coordinates programmes within these fields on behalf of the CNRS and universities, in partnership with CEA. Whilst these main fields represent the core of the discipline, IN2P3 also has several additional vocations: enabling other scientific domains to benefit from its competencies and solving certain problems posed by society, and accompanying universities in contributing to youngsters' training. IN2P3 is also responsible for running several major national facilities including particle accelerators. It also supports several international facilities. As the experiments require significant investments in terms of financing and personnel, from the outset the Institute took the structure of a limited number of large laboratories and Infrastructures or technological platforms in order to facilitate the pooling and optimization of its resources and competencies. IN2P3 also shares with the CNRS/INSU Institute for Astrophysics the support to some projects in Astrophysics and Cosmology. CNRS, together with CEA-IRFU, fund almost all national e-Infrastructures for these domains. These e-Infrastructures provide data archive and computing resources also to research projects in other fields such as biology and life science, humanity and social science.

- INAF, the National Institute for Astrophysics in Italy. INAF is currently operating several observing facilities - both space- and ground- based - producing a huge amount of data. About a thousand of staff and contract research people are working with these data acquired over the world. These data is going to increase by several orders of magnitude thanks to the next generation of observatories in which INAF is involved as Euclid, CTA, E-ELT and SKA. This wide and heterogeneous community will need to access data in a multiwavelength and multimessenger framework. At the same time there is also a wide community working on data analysis and cosmological simulation using local data center or Tier-0 external infrastructure like CINECA. Data archives are spread over INAF structures with mainly two data center: IA2 (Italian Center for Astronomical Archive) that is an INAF data center and ASDC (ASI Science Data Center) owned by ASI, but with a component on INAF staff. Dedicated data center are under development to fit the peculiarity of the huge data producer instruments. For example, the Cherenkov Telescopes Array is a large observatory devoted the Very High Energies gamma ray astronomy that will dominate the most part of this century. It will produce a huge amount of data (from 9 PB up to 25 PB per year) and will be operated for at least 30 years after its deployment. Computing resources with very different approaches are under study and development: HPC and HTC for the data reduction to be performed directly at the observatory site, a distributed archive and distribute computing both in a Cloud framework for the data centers. Also 10 Gbps links is the next target to connect observation facilities and all INAF structures to best optimize INAF data and computing infrastructures.

- INFN, the National Institute for Nuclear Physics, is the Italian research agency dedicated to the study of the fundamental constituents of matter and the laws that govern them, under

the supervision of the Ministry of Education, Universities and Research (MIUR). It conducts theoretical and experimental research in the fields of subnuclear, nuclear and astroparticle physics. All of the INFN's research activities are undertaken within a framework of international competition, in close collaboration with Italian universities on the basis of solid academic partnerships spanning decades. Fundamental research in these areas requires the use of cutting-edge technology and instruments, developed by the INFN at its own laboratories and in collaboration with industry. Groups from the Universities of Rome, Padua, Turin, and Milan founded INFN on 8th August 1951 to uphold and develop the scientific tradition established during the 1930s by Enrico Fermi and his school, with their theoretical and experimental research in nuclear physics. In the latter half of the 1950s, INFN designed and built the first Italian accelerator, the electron synchrotron developed in Frascati, where its first national laboratory was set up. During the same period, INFN began to participate in research into the construction and use of ever-more powerful accelerators being conducted by CERN, the European Organisation for Nuclear Research, in Geneva. Today, INFN supports the research activity of about 5,000 scientists whose work is recognised internationally not only for their contribution to various European laboratories, but also to numerous research centres worldwide.

- KIT, the Karlsruhe Institute of Technology, was founded in 2009 by the merger of the Karlsruhe Research Centre (Forschungszentrum Karlsruhe, FZK) and the University of Karlsruhe (TH). It is a public corporation according to the legislation of the state of Baden-Württemberg and fulfils the parallel missions of a university and of a national research centre within the Helmholtz Association. With 9,400 employees (including nearly 6,000 employees working in research and education), 370 professors, 3,200 doctoral researchers, and an annual budget of about 790 million (2013), KIT is one of the biggest research and education institutions in Europe. At Steinbuch Centre for Computing (SCC) R&D focuses on computational science and engineering, data life cycle management of scientific data, IT management, web engineering, federated identity management and process integration. Furthermore, SCC is also the IT service provider of KIT. SCC operates very powerful HPC systems, among others the high-performance computer of the State of Baden-Württemberg, being used by regional and federal scientists from academia and industry for computationally intensive projects. In this context SCC's SimLabs form the interfaces between users and providers focusing on joint R&D in HPC and computational science. SCC is a member of the Gauß-Alliance that promotes and facilitates strategic HPC research in Germany. SCC is the location of GridKa – one of twelve WLCG Tier-1 centres worldwide. SCC is leading several national and state-wide projects related to big data management, archival and analysis. With the LSDF SCC has realised an innovative concept for storage, archival and analysis of scientific data. The LSDF supports the complete data life cycle and is in principle available to all scientific disciplines. In several state-funded projects the LSDF is enhanced with additional services for sync&share, block storage and archival available to all other universities in the state of Baden-Württemberg. A federated identity management system enables a seamless usage of the LSDF as well as powerful HPC and data analysis facilities in the state. In SCC's Data Life Cycle Labs (DLCLs), data experts perform joint research and development activities with scientists from several research fields to advance specific data life cycles; this is enhanced by technology research for generic, multi-community tools and services meeting the demands on IT security and trust.

- STFC, the Science and Technology Facilities Council in the UK, is the funding agency with responsibility for providing support for all staff and resources for Particle Physics, Particle Astrophysics, Astronomy, Astrophysics, Cosmology and Nuclear Physics. STFC funds all e Infrastructure for these domains and in particular funds the GridPP project which provides all WLCG computing and the DiRAC project which provides HPC facilities for theoretical particle physics and astronomy. STFC also runs several major UK national laboratories including the

Rutherford Appleton Laboratory, the Daresbury Laboratory, the Astronomy Technology Centre and facilities including neutron and light sources. STFC also hosts the Hartree HPC centre.

- SURFsara is the national HPC and e-Science support centre, and supports research in the Netherlands by developing and offering, together with its partners, advanced and sustainable ICT infrastructure, services and expertise. This support is ensured by providing an integrated ICT research infrastructure with services in the areas of computing (supercomputing, capacity computing, high-throughput computing), data storage and data analysis, visualisation and cloud services, by providing multidisciplinary expertise and support in e-infrastructure use and application, and by doing all required innovation, engineering and development on the ICT infrastructure. SURFsara hosts large national infrastructure services, i.e. the Dutch national Supercomputer service, the National Compute Cluster, large data storage facilities and services, a large Hadoop cluster, HPC Cloud service and all of the centralized national grid services. This includes also the LHC Tier-1 service for CERN wLCG, which is hosted jointly with Nikhef (the Dutch National Institute for Subatomic Physics), and long-term archive for the LOFAR telescope. The facilities and services provided by SURFsara and partners are integrated, via the SURF foundation, with high-performance networking services (SURFnet) and with R&D on digitally-enhanced science (Netherlands eScience Center), implementing a complete and integrated national e-Infrastructure ecosystem. SURFsara serves a number of large research, educational and government institutions, all of the Dutch Universities as well as industries and SMEs. SURFsara also serves a large number of scientific communities including WLCG, CLARIN, LOFAR, BBMRI, Lifewatch, WeNMR, EPOS, and ELIXIR. SURFsara is also representing the Dutch national e-infrastructure in EGI.eu, EUDAT and PRACE AISBL. SURFsara is subsidiary of SURF, the Dutch collaborative ICT organisation for Dutch higher education and research. SURF is responsible for the complete national e-infrastructure (computing, data, network).

EU-T0 Statement of Intent Document (SID)

EXECUTIVE SUMMARY

The EU-T0 vision is to create a hub of knowledge and expertise to coordinate technological development, and integrate data and computing centres into a virtual Tier-0 centre, which can be more easily and transparently used by the different Scientific Communities. It will extend and simplify access to the resources, through a common approach to authorization, authentication and accounting frameworks. It will innovate new services and software tools as required to benefit the Scientific Communities and in particular the means to store, move, archive, preserve, and provide open access for very large scale data sets.

The EU-T0 consortium was initiated on 11th February 2014 by CERN, CIEMAT, DESY, IFAE, IN2P3, INFN, KIT and STFC. INAF, CEA(IRFU) and SURFsara joined in 2015. These are funding agencies and organisations which between them fund and provide very large scale computing and data processing resources for a set of scientific communities which already have Exabyte scale requirements. Participation in the consortium by similar funding agencies and organisations is encouraged.

The data centres of the consortium, and the communities they support, have been highly successful scientifically; they have innovated the development of distributed computing on a global scale, and represent an important part of the European e-Infrastructure for research. The forecast growth in both the rates and complexity of the data arising from the scientific communities is enormous, and will rise from Exabytes to Zettabytes within a decade. This gives rise to many technical and financial challenges in the future in order to meet these demands.

The EU-T0 partners already work closely together and now see the opportunity to cooperate even further in order to address the challenges and opportunities, with the expectation that this will yield scientific, technological and cost-efficiency benefits.

EU-T0 is multi-domain, its Scientific Communities include astronomy, astrophysics, astro-particle physics, cosmology, nuclear physics and particle physics and in some cases national and international photon and particle beam facilities as well as engineering and life sciences. EU-T0 seeks synergies with other research domains, and will collaborate with other projects and initiatives. EU-T0 will make an important contribution to structuring the European Research Area (ERA) through cooperation with ESFRI RIs.

EU-T0 will also seek to integrate with other e-Infrastructures and broaden existing collaborations beyond Europe.

EU-T0 will explore and propose a new model of international coordination for joint procurement and joint services and resources provision.

1. The EU-T0 Motivation and Vision

The EU-T0 initiative was launched on the 11th February 2014 by some major research funding agencies and organisations: CERN, CIEMAT, DESY, IFAE, IN2P3, INFN, KIT and STFC. INAF, CEA(IRFU) and SURFsara joined in 2015. The participation of similar agencies and organisations is encouraged. The above listed agencies and organisations are hereafter referred to as the *Partners*.

The Partners have considerable overlap in their roles and responsibilities in support of scientific research. Those *Scientific Communities* which are supported by each Partner are listed in Appendix A. In most cases these include astronomy, astrophysics, astro-particle physics, cosmology, nuclear physics and particle physics. In some cases photon and particle beam facilities as well as engineering and life sciences interests are also included.

Between them, the Partners own or provide a very large amount of the physical resources used by their Scientific Communities. Their data centres, which rank as some of the largest globally, act as the primary international repositories for communities already dealing with data volumes on the scale of Exabytes. They have amassed a considerable amount of expertise in their operation. EU-T0 is therefore already de-facto a major, successful, and existing component of the European e-Infrastructure.

The forecast growth in both the rates and complexity of the data arising from the Scientific Communities is enormous, i.e. from Exabytes to Zettabytes. A commensurate increase in the computing power is required and the need to move, manage and access data globally will be even greater. In the light of such growing demands, the Partners see several common opportunities and challenges for the future, which include:

a. How to innovate shared computing and data processing services, which can be more easily and transparently used by their Scientific Communities.
b. How to provide for the increasing resource demands within challenging financial constraints, including realising economies of scale.
c. How to provide greater and simpler accessibility, through a common approach to integrating Authentication, Authorization and Accounting frameworks.
d. How to efficiently utilise the next generation of computer architectures.
e. How to store, move, archive, preserve, and provide open access for Zettabyte scale data sets.
f. How to engender information exchange in respect of technology, services and standards between scientific communities, and thereby enable new communities to capitalise on the existing investment.
g. How to provide stability on the scale of decades which is required for the largest and longest lived scientific endeavours, and the ensuing requirements of data stewardship.
h. How to share benefit and experiences from technological development programmes.

It is therefore natural and appropriate that the Partners continue to work together to address the future opportunities and challenges presented above, with the expectation that this will yield scientific, technological and financial benefits. The Partners will work towards integrating their data and computing centres, into a virtual Tier-0[1] centre accessible to a wide set of communities, and coordinate the development of future computing and data processing services. This is the rationale for the formation of EU-T0 Consortium and its vision:

"The EU-T0 vision is to create a hub of knowledge and expertise to coordinate technological development, and optimise the investment of the funding agencies in their existing data centres by broadening, simplifying, and harmonising access, driven by well-defined user requirements."

The EU-T0 vision embodies the idea of a hub focussed on science domains, which will primarily undertake activities that follow directly, or indirectly, from the requirements of its Scientific Communities. Many years of experience have demonstrated that the most successful e-Infrastructures are those that are

[1] The term Tier-0 is used to mean a top-Tier data processing and storage system capable of archiving, processing and providing access to data arising from experimental and simulation facilities. Referring to the top-Tier as Tier-0 is consistent with the terminology used in the Worldwide LHC Computing Grid (WLCG) and the Partnership for Advanced Computing in Europe (PRACE).

closely coupled to, and driven by, the Scientific Communities. Therefore EU-T0 recognises that its Scientific Communities must be paramount in defining its strategy and activities.
In the longer term EU-T0 will be open to incorporating additional scientific communities where this leads to benefits to all communities concerned, and furthers the long term strategic objectives set out in this document.

EU-T0 seeks the most effective means to achieve its vision. Therefore wherever possible, the consortium will seek to leverage existing solutions, frameworks, and infrastructures, and to collaborate with other existing and new projects, in order to ensure mutual benefit and minimise duplication.

EU-T0 is consistent with the recommendations presented by the IT working group of the EIROforum in a recent document[2]. EU-T0 is an example of a "research accelerator hub - Reach" focussed upon its Scientific Communities.

EU-T0 addresses the main recommendations provided by the e-Infrastructure Reflection Group (e-IRG). EU-T0 will be a contributor to structuring the European Research Area (ERA) through cooperation with ESFRI RIs and other domain focussed infrastructures.

2. Objectives

The objectives of EU-T0 are:

I. To build upon the existing, successful and well proven, computing and data processing centres that are owned or funded by the EU-T0 Partners, by enhancing cooperation and harmonisation in respect of all aspects of the infrastructure.
II. To integrate the computing and data processing centres into a European Tier-0 centre by creating shared services which can be more easily and transparently used by the different Scientific Communities.
III. To extend and simplify access to the EU-T0 data centres and their services, through a common approach to integrating Authentication, Authorization and Accounting frameworks.
IV. To achieve a more cost effective provision arising from economies of scale, minimisation of duplication, and adoption of common solutions.
V. To innovate new services and software tools as required for the benefit of the Scientific Communities. In particular to develop the means to store, move, archive, preserve, and provide open access for Zettabyte scale data sets.
VI. To integrate with other e-Infrastructures and to work towards convergence of HPC and HTC services.
VII. To engage with commercial infrastructure providers to develop a workable mixed model of provision.
VIII. To broaden existing collaborations beyond Europe.

The EU-T0 consortium will act as a focus to address future opportunities targeted at supporting these objectives. It will support collaboration with other projects where appropriate and will also support the initiation of new projects and proposals where appropriate.

[2] EIROforum IT working group (2013). e-Infrastructure for the 21st century. ZENODO DOI: 10.5281/zenodo.7592

3. Initial Activities

A likely set of initial activities is described below. The nature and extent of activities undertaken will be determined by the resources that can be made available by partners or external funds which can be won through competition. In the immediate future EU-T0 members expect to present proposals to the Horizon 2020 framework programme for some of the activities identified below.

I. Coordination and collaboration to develop and implement solutions which will allow ingestion, storage, movement, archival, preservation and access (including open-access when relevant) of increasingly large and complex datasets, with a horizon of reaching in a decade aggregate levels of Zettabytes with Exascale number of data objects. Improving the ease of data exchange with HPC installations. This will allow services to fulfil the increasing requirements of the Scientific Communities and the corresponding large-scale facilities.

II. Coordination of the work to develop and deploy virtualised services, collaboration with other projects working in this area, and piloting of the incorporation of commercial cloud resources. This will bring the benefit of providing a simpler underlying computing fabric which can more easily be (virtually) tailored to different Scientific Communities as required, including those of long tail science.

III. Explore joint procurement opportunities where these would be cost effective.

IV. Coordination of deployment of common AAA systems to simplify access for diverse Scientific Communities. This may include international policy and trust agreements, and endorsing and supporting projects to develop European or global AAA infrastructures. This will bring the benefit of "single sign on" to multiple diverse resources.

V. Coordination and collaboration in the development, provision and deployment of a portfolio of services, including new services and software tools, as required. In particular those which allow the building of customizable Virtual Research Environments, composed of services for data workflow management, user access to software, archives, and scientific analysis systems; digital platform relating publications, data and analysis software repositories.

VI. Working closely as a single point of contact with initiatives from scientific domains, such as ESFRI RIs, to develop agreements to provide resources and deploy domain based services across the EU-T0 data centres.

VII. Reduction of cost through the sharing of resources and expertise between various scientific communities and data centres.

VIII. Further harmonisation of security coordination, policy development and incident handling.

IX. Promotion of a training network to create and develop data scientists.