SEVENTH FRAMEWORK PROGRAMME

Inventory of existing PA-RIs cooperation

FP7-ICT-248295/INFN/R/RE/D3.1

Project Number: FP7-ICT-2007-4 248295

Project Title: Towards an Open and Sustainable ICT Research Infrastructure Strategy (OSIRIS)

Contractual Date of Deliverable: 31/12/2010

Actual Date of Delivery: **17/01/2011**

Workpackage contributing to the Deliverable: WP3 : Benchmarking

Nature of the Deliverable R

Dissemination level of Deliverable RE

Authors: INFN / Antonio Candiello, Mirco Mazzucato

Contributors:
IMCS / Baiba Kaškina
MFA / Géza Ódor
CSIR / Colin Wright
CSEM / Arno Hoogerwerf
IBBT / Rosette Vandenbroucke
ISERD / Felix Cincarevsky

Abstract:

This deliverable presents a list of the relevant cases of current cooperation between Public Authorities and National Champions and existing ICT RIs.

Keyword list:
Research Infrastructures, ICT, inventory, cooperation.

Clarification

Nature of the Deliverable

R	Report
P	Prototype
D	Demonstrator
O	Other

Dissemination level of Deliverable:

PU	Public
PP	Restricted to other programme participants (including the Commission Services)
RE	Restricted to a group specified by the consortium (including the Commission Services)
CO	Confidential, only for members of the consortium (including the Commission Services)

Disclaimer

The information, documentation and figures available in this deliverable, is written by the OSIRIS ("Towards an Open and Sustainable ICT Research Infrastructure Strategy"*) – project consortium under EC co-financing contract FP7-ICT-248295 and does not necessarily reflect the views of the European Commission.*

Executive summary

This document lists the most relevant cases of current cooperation between Public Authorities (PA) and National Champions (NC) and existing ICT Research Infrastructures (RIs).

A high-level list of relevant ICT Research Infrastructure models is reported, identifying the relevant network RIs, computing RIs, MNT & instrumental-related infrastructures, digital libraries frameworks. The Future Internet vision was also considered in this inventory.

Cases have been selected on the basis of publications like [WGFI, 2008, DMTF, 2009, EC-ESF, 2007, EGI_DS, 2008] and also following indications by the consortium members and stakeholder group. The challenges analyzed and the results of the survey conducted in Work Package (WP) #2 are also taken into consideration in this process, in order to evaluate the specific relevance.

For each domain, the relevant projects, the governance models and challenges experienced are analysed. The consortium members provided the required input in order to better evaluate the cases of PA/NC cooperation with ICT RIs. A detailed list of intra-domain PA/NC-RIs collaborations is also reported.

This inventory will serve as a basis for the planned subsequent benchmark analysis.

Content

1. Introduction

This document lists the most relevant cases of current cooperation between Public Authorities (PA) and National Champions (NC) and existing ICT Research Infrastructures (RIs).

The cases have been selected on the basis of publications like [WGFI, 2008, DMTF, 2009, EC-ESF, 2007, EGI_DS, 2008] and also following indications by the consortium members and stakeholder group.

The challenges analyzed (see §1.1 in the following) and the results of the survey (see §1.2 and §1.3), both conducted in Work Package (WP) #2, were also taken into consideration in this process, in order to evaluate their specific relevance. This Section is mainly devoted to report the WP2 results relevant for the selection and prioritization of PA/NC-RIs cooperation cases.

The consortium members then provided the required input in order to better evaluate the cases of PA/NC cooperation with ICT RIs. This allowed us to frame a set of key ICT RI *models*, presented in Section 2. The detailed analysis of each context brought us to structure an inventory of PA/NC-RIs *collaborations*, which is the subject of Section 3.

Some final considerations will conclude the document.

1.1 ICT RIs Challenges

We recall here the main results of [OSIRIS D2.1, 2010].

Research Infrastructure *is a pillar in developing European Research Area* (ERA). This holds true particularly in the field of ICT Infrastructures. Progress in ICT and its Infrastructures *is a "condicio sine qua non" for most of the other installations*.

ICT Research Infrastructures are identified as *e-Infrastructures* or *Infrastructural Competence Centres* dealing with *Research and Development in the ICT area* and/or *ICT facilities for Research and Development.*

ICT RIs challenges were grouped in the categories:

- *Policies*: public policy, science policy, innovation policy, cooperation policy, technological policy
- *Resources*: financial resources, human resources, users
- *Frameworks*: legal framework, organizational forms of e-Infrastructures, intellectual property rights, infrastructure security and data privacy
- *Scale*: regional dimension and global dimension.

The following challenges were identified:

1) *Infrastructural Competence Centres* (ICCs): creation of European Research Infrastructures or a network of them for the development of *Micro and Nano Technologies* (MNT) *is still a challenge* because of processing constraints between

facilities. The cooperation between the MNT ICCs could be expected to develop on the basis of an interconnection of complementary and not overlapping capabilities.

2) *Data RIs*: the processing of large amounts of data is the *most important and challenging ICT problem* compared, for example, with providing large computing facilities.

3) *Public and science policy*: the wide usage of Research Infrastructures based on the principles of the Fifth freedom is still a challenge for the Member States. The development of Research Infrastructures will challenge changes in the existing science policy of Member States and also the ability to coordinate their national policies in the overall context.

4) *Innovation & Cooperation policy*: the organization of cooperation between scientific institutions and business companies in the framework of Research Infrastructures is rather challenging because the interests of these two groups and their attitude to multinational cooperation may be different.

5) *Technological policy*: creating a common technological policy for e-Infrastructures is a challenge for ICT researchers and for the ICT Research Infrastructures.

6) *Standards*: standardization is necessary for the development of e-Infrastructures and *its implementation is a serious challenge*.

7) *Usability*: “ease of use” principle should be *a primary objective* in the transition of ICT RIs from skilled communities to wider, less skilled communities.

8) *Legal framework*: the work on the creation of the legal framework for Research Infrastructure activities has been started but is still a legal challenge. The law for setting up of the *European Research Infrastructure Consortiums* (ERIC) has been adopted but the rules for functioning are still incomplete.

9) *Intellectual property rights*: the deployment of Research Infrastructures is challenging for experts in *Intellectual Property Rights* (IPR).

10) *Infrastructure security*: ensuring Research Infrastructure security (Solving security issues related to the RI) is rather challenging, and should be considered as an important part of any RI project.

11) *Data privacy*: any projected Research Infrastructure should be audited regarding *personal data protection* issues.

12) *Regional dimension*: the development of *regional level Research Infrastructures* is an interesting and challenging task.

13) *Global dimension*: long-term commitment and continuous financial support required for the e-Infrastructures operating on a global scale will be challenging.

14) *RIs strategy*: the development of Research Infrastructures will continue, but some organizational changes may be expected.

1.2 General characterization of the PA-RIs cooperation

For general insight into problems concerning the cooperation of Public Authorities and Research Infrastructures the results of the OSIRIS Survey may be used. Among other problems this Survey studied the general approach of the main stakeholders to the cooperation with ICT Research Infrastructures and also identified main characteristics of the cooperation. The specific problems of infrastructures, of course, were not investigated. The results of the Survey are overviewed in the deliverable D2.5 presented in two parts, and the discussion here is based on these results.

The sample of respondents to the OSIRIS Survey comprised also representatives from Public Authorities. The respondents were divided into several groups, and the group Policy consisted of the respondents who either belonged to public authorities with competencies in the ICT sector, either had experience in policy, running, planning and/or financing ICT RIs or were engaged in the cooperation with existing ICT Research Infrastructures in other ways. Their opinion concerning the cooperation may reflect to some extent the approach of European Public Authorities. At the same time the opinions of the other respondent groups – Operation and Users may present some understanding of the perception of the PA activities in these groups.

Priorities of Research Infrastructures

The respondents were asked to indicate the main areas they represent, and they could select several Infrastructures of the proposed six types of them. In average the respondents checked about two RIs as the most relevant to their activities. The vast majority of the respondents thought they are closely connected to the Computing Infrastructures (63.6%) and to the Communication Infrastructures (52.7%). The Policy group showed similar results, as it is seen from the following table presenting the percentage of RIs types checked by the respondents of this group. The last column presents the appropriate percentage for all respondents.

Tab. 1 - Percentage of checked RIs by Policy group

Infrastructures	Policy group	Total
Computing Infrastructures	73.1%	63.6%
Communication Infrastructures	53.8%	52.7%
Future Internet	26.9%	25.5%
Data Infrastructures	26.9%	18.2%
Instrumental Infrastructures	19.2%	14.5%
Micro and Nano Technology RIs	15.4%	18.2%

It is quite clear that the Policy group considered the Computing Infrastructures as being the most important Infrastructures – the respondents of this group checked these Infrastructures in 73.10% cases. The results on the table would suggest that the group kept the focus on the care of the Computing and Communication Infrastructures (grid, HPC and similar). The Future Internet, Data Infrastructures and Instrumental Infrastructures appeared to be much less important for the Policy group, and the MNT Research Infrastructures were practically neglected.

It seems that such distribution of priorities does not correspond to the real importance of different Research Infrastructures. It may be agreed that MNT Research Infrastructures will always been less ubiquitous than Computing Infrastructures or especially Communication Infrastructures, but the Data Infrastructures or Future Internet Infrastructures will be able to compete in popularity and importance with Computing Infrastructures.

Additionally the Survey also studied the usage of ICT Research Infrastructures by the respondents, and it was reported that totally more than a half of them used the RIs in their research activity. The usage of these RIs by the Policy group was less active – 39.3% of them used, 35.7% – did not use, and to 25.0% the usage was not applicable. It suggests, however, that, generally speaking, the Public Authorities also use the Research Infrastructures practically and are really engaged in the process of development and collaboration.

Assessment of scientific cooperation

The Survey also studied the opinion of the respondents about the level of scientific cooperation in their field of the ICT sector in their countries. These results can be used to evaluate the general attitude to the cooperation for the development of the Research Infrastructures.

The respondents could choose one of five possibilities to characterize the scientific cooperation (see, the table below) of the ICT sector in their countries. The chosen answers of the Policy group are shown in the following table:

Assessment of cooperation	**Policy group**	**Total**
Satisfactory	28.6%	42.1%
Good in the whole ICT sector	39.3%	29.8%
Good only in the Research Infrastructures	10.7%	12.3%
Poor	17.9%	10.5%
Good in the ICT sector, except Research Infrastructures	0.0%	1.8%
No opinion	3.6%	3.5%
Total	100.0%	100.0%

It may be concluded that the opinions of the Policy group differ from the averaged opinion of all respondents. Most of the respondents thought that the cooperation is satisfactory. The opinions of the Policy group are somehow extreme – more respondents of this group compared with average perceive the cooperation good in the whole ICT sector including RIs and more of them evaluated it as poor. The high percentage of the opinion about poor cooperation in the Policy group may be considered as a kind of self-criticism, because in reality this group could do a lot for improving the scientific cooperation.

It should be mentioned that the respondents in their comments to the answers identified two main obstacles for the development of the scientific cooperation in the ICT sector. The first is the fragmentation of the organisations on the national level i.e. rather large number of small organisations and enterprises with fewer resources for research and also on the European level, where the national communication infrastructures should cooperate. It was proposed to promote networking actions at national level and to coordinate them with European measures. It is quite obvious that the development of national academic networks connected into a pan-European structure (GÉANT) may overcome this threat of fragmentation. It is rather important for the cooperation to define national research programs and national roadmaps for RIs, taking into account well pronounced needs of users.

As the second obstacle to research cooperation the competition between institutes within countries for limited funding was mentioned. It should be added, however, that the EU science policy in Framework programmes always was to support collaborating research institutions of several countries working on sufficiently great projects, which stimulated the research cooperation on the EU level.

Coordination of priority research themes

One of the way the scientific cooperation can be developed is the proper definition of the priority research themes in the country. If the priority research themes are set up without coordination with other countries it is difficult to expect good cooperation with them. The Survey asked respondents if the priority research themes in the ICT sector of the country are coordinated with other Member States.

The appropriate results for the Policy group are presented in the following table:

Tab. 2 – Percentage for coordination of priority research themes

Coordination	**Policy group**	**Total**
Yes	25.0%	38.6%
No, they are set by the science policy of the country	46.4%	38.6%
No opinion	28.6%	22.8%
Total	100.0%	100.0%

It was rather surprising that the respondents from the Policy group were of rather low opinion about the existence of the European coordination of the priority ICT Research themes. The percentage of *No* answers was the greatest in this group as was the percentage of respondents who had no opinion. Only one of four respondents admitted the existence of this type of coordination. The same indicator was about two times larger for other groups.

The comments of the respondents in free form allow to conclude that the coordination mechanisms are rather widespread and well developed, and are mainly based on the participation in the Framework Programmes, eIRG or ESFRI and in various European RI projects like GÉANT, PRACE, EGI, Artemis and others.

Influence of RI on research themes

The influence of existing or planned RIs on the priority research themes in a country can be considered as one of the manifestation and result of the research cooperation relevant to the RIs. Therefore the respondents were asked to assess the observed influence.

The results are presented in the following table:

Tab. 3 – Percentage of influence of RIs on research themes by Main activity groups.

Do RIs influence research themes?	Policy group	Total
Yes	50.0%	52.6%
No	28.6%	21.1%
No opinion	21.4%	26.3%
Total	100.0%	100.0%

Generally speaking, the question was rather unexpected for the respondents because relatively many of them had no definite answer to this question. The opinions of the Policy group were approximately the same as of the averaged answers of all respondents. However, the 50.0% of the positive answers can not be considered as an indicator of great influence of RIs on the development of research themes.

The answers to a previous question concerning the coordination of priority research themes in the respondent countries with other MS were used to find if the coordination of research themes increased the influence of RIs on the research themes or had no effect. The results show with certainty that the research themes that are coordinated with other Member states are influenced by the existing or planned Research Infrastructures. 72.7% of the respondents who agreed that the priority ICT research themes are coordinated with other MS found that RIs influenced the priority research themes and only 13.6% thought that such influence did not exist.

In the comments the respondents mentioned the ways this influence is implemented. They reported about the usefulness of the ESFRI roadmap, various EU projects like FIRE, PRACE, EGI, GÉANT, etc.

At the same time, it was also mentioned that there is interplay of investment into research infrastructures and relative stability of related research teams and laboratories because good and widely orientated projects using European range RIs may be very supportive for survival of research institutions.

Development of EU ICT Research Infrastructures

The respondents also assessed the level of country contribution to the development of the EU ICT Research Infrastructures. This assessment can also serve as an indicator of cooperation activities. The available answers were either – Wide participation, Participation by some institutions only, or No participation. It was proposed also to inform about the contribution to the development of each type of Research Infrastructures. The results are overviewed in the following table for all respondents:

Tab. 4 – Evaluation of the country contribution in development of RIs.

Evaluation	***Computing Infrastructures***	***Communication Infrastructures***	***Data Infrastructures***	***Instrumental Infrastructures***	***MNT Research Infrastructures***
Participate widely	28.1%	49.1%	19.3%	12.3%	8.8%
Only some institutions	54.4%	28.1%	43.9%	26.3%	40.4%
No participation	3.5%	1.8%	8.8%	15.8%	7.0%
No opinion	12.3%	17.5%	24.6%	42.1%	40.4%
No answer	1.8%	3.5%	3.5%	3.5%	3.5%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

This table can be compared with the following table that presents the analogous percentage values for the answers of the Policy group only:

Tab. 5 - Evaluation of the country contribution in development of RIs by Policy group

Evaluation by Policy group	**Computing Infrastructures**	**Communication Infrastructures**	**Data Infrastructures**	**Instrumental Infrastructures**	**MNT Research Infrastructures**
Participate widely	39.3%	50.0%	25.0%	17.9%	10.7%
Only some institutions	53.6%	28.6%	60.7%	35.7%	42.9%
No participation	0.0%	0.0%	3.6%	7.1%	7.1%
No opinion	7.1%	17.9%	10.7%	39.3%	35.7%

No answer	0.0%	3.6%	0.0%	0.0%	3.6%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

An answer of wide participation was highest for the Communication Infrastructures; and second highest for Computing Infrastructures for the total count of answers as well for the Policy group. At the same time, the sum of wide participation and the participation by some institutions only was highest for the Computing Infrastructures, again both for the total values and Policy group.

It is seen also that the Policy group was of higher opinion about participation in all types of Research Infrastructures especially in the development of Data Infrastructures though the level of participation was reported as rather low especially for the Instrumental Infrastructures and MNT Research Infrastructures. Generally the values of the level of participation in the development of various RIs correlate with the importance range of RIs already mentioned above.

1.3 Relevant Survey Results

We report here the main results from the survey submission and subsequent analysis:

R1. The *level of scientific cooperation* in the ICT field in the EU countries is *satisfactory*. It is better in the Operation group than in the Policy group.[1] The development of the cooperation is hindered by fragmentation of the sector and by research competition.

R2. The *priority research themes* in the ICT sector in countries with other Member States are coordinated insufficiently, the level is especially low in the Policy group, though coordination mechanisms exist and are used.

R3. The existing Research Infrastructures have no serious influence on the organisation of research activities, especially in the Policy group. However, the research plans of some countries have identified the need for new RIs, especially in the EU12 group[2]. The Policy group in general does not feel that this is necessary.

R4. The European coordination undertakings, advisory bodies, working groups and organizations from the ICT Research Infrastructure point of view *are considered to be important* though not all their activities are considered necessary.

R5. *Paid research services* could be provided by ICT Research Infrastructures mainly in cases where no other funding sources exist. Generally the respondents think that only open access based on peer review process guarantees the best and most efficient use of RIs.

[1] In WP2, the *Operation Group* was formed by the persons operating national/regional ICT Research Infrastructures and the *Policy Group* was mainly represented by those who have experience in policy, running, planning and/or financing ICT RIs and Public Authority with competences in the ICT sector; the *Usage Group* consisted of ICT RI users.

[2] Country groups were defined in WP2 as: *EU15*, the 15 countries that were members of the EU before the enlargement on 1st May 2004; *EU12*, the 12 countries that joined the EU over the past years; *other countries*, the associated countries and other countries.

R6. To *improve the cooperation between ICT RIs and industry*, it is proposed to intensify PR activities, visibility and information, to introduce appropriate tax incentives, to improve the legal environment, to establish the full innovation chain within the value creation process, to participate in common projects and to make all non-exclusive licenses available royalty free across the entire EU.

R7. Using *intellectual investments* as contributions to the development of the ICT Research Infrastructures is supported by approximately half of the respondents; the other respondents have some doubts about the acceptance of such a model.

R8. *Three out of ten* respondents accepted *Public Private Partnership* (PPP) in the creation of the RIs without any reservations. Another four out of ten thought that *the existing legal framework is not sufficient and should be improved*.

R9. Two out of ten respondents thought that the scientists experienced no difficulties in working with the Research Infrastructures, and more than a half of the respondents saw the necessity for *additional training* orientated to usage of RIs. It may be concluded that *the additional training can and should be provided by the Research Infrastructures themselves*.

R10. The prevailing management systems are *different* for the various Research Infrastructures. The RIs can be divided into *two groups* – the HPC, Grid Computing and Communication Infrastructures with management being primarily at *national level*, and the *Data, Instrumental and MNT Research Infrastructures* where management occurs on the site/laboratory level. The management by an international body was indicated only in some responses, specifically in the Grid Computing and the Communication Infrastructures.

R11. More than half of respondents accept *open access* to the Research Infrastructures for all researchers. One out of five respondents would prefer open access *to academic researchers only*.

R12. One out of five respondents thinks that the existing legal framework of the *Intellectual Property Rights* (IPR) is sufficient, others hope for serious or minor amendments. The necessity for serious elaboration of the existing legal framework is much more pressing in the EU12 group.

R13. The most supported legal form for the RIs is *ERIC*.

R14. Data security *is not sufficient in all RIs*, the best situation is in the *Computing Infrastructures*. Using a federative network is the most desirable way to improve the situation.

R15. The *existing privacy* in the RIs is at a higher level than security although the enforcement of additional privacy measures is considered necessary.

R16. Half of the respondents participate in regional projects. The EU12 group and the Policy group are more active in regional cooperation. The preferred infrastructures for regional development are *Data Infrastructures*. The EU12 group also supports regional development of the Computing Infrastructures, and the EU15 indicates the MNT Research Infrastructures as being the best candidate for regional development

R17. The majority of respondents, but not all of them, support the creation of an organizational structure for *continuous analysis and recommendations* on the development and maintenance of existing and future European ICT RIs. The best legal form for such an organizational structure could be ERIC but Agreement between governments or institutions, Virtual institutes or laboratories could be also considered.

2. Structure and Models of ICT RIs

As an intermediate step for an inventory of the existing ICT RI models of cooperation between Public Authorities and National Champions, a thorough analysis of the different ICT RIs has led us to the conclusion that these have an intrinsic (topological and physical) structure that is built around the following elementary objects:

- *networks* – technological architectures that allow the interconnection of different ICT facilities between each other;
- *facilities* – sites with concentrated resources, like *computing* facilities (as in HPC centres or smaller DCI computing sites/nodes) or *storage* facilities as in scientific repositories, digital libraries and DCI storage sites/nodes;
- *instruments* – scientific equipments capable of creating (in particular, digital) data from real world observation; these provide new scientific data;
- *testbeds* – artificial environments needed to ideate, project and test new kinds of ICT/internet interaction between humans and things;
- *laboratories* – plants for test and production of new ICT physical devices.

Networks are necessary to connect facilities, instruments and laboratories to each other. Because the network grid is a minimum requirement for the cooperation between the other facilities, its need is unquestioned. In general, it is implement with top-of-the-line standard equipment and there is in general only one provider of the research network infrastructure in one country. Competing offers for research network services generally don't exist.

Facilities can be computing facilities or storage facilities. *Computing* facilities work together in the distributed form of *Distributed Computing Infrastructures* (DCI) or in the concentrated form of *High Performance Computing* (HPC). The computing facilities gain from the cooperation in DCI and HPC through load optimization. When needed, some very high performance computing facilities are available and when not needed the computing facility lends its power to another user elsewhere. This cooperation became possible only once the required *grid middleware* was developed to make the contribution of a computing facility transparent to the user. Billing and prioritizing issues appear to have been solved as well. *Storage* facilities store the enormous amounts of scientific data available in this world. These storage facilities gain from cooperation because it allows them to distribute data over different centres and achieve redundancy of all the data. A lot of effort is needed to make the data accessible. This starts by achieving a uniform description of all data. Moreover, a mechanism to define the value of the data is necessary.

Instruments are often expensive facilities that help to improve our understanding of complex problems. They often generate a lot of data for later analysis. Typically, for one type of problem, there is only one, or a very small number of instruments in which many parties participate to share the cost. The results from these instruments contribute in general to basic science and do not have a direct economic impact.

Testbeds are needed to test and verify the newly developed standards. The size of these testbeds will have to be non-negligible to be able to represent a real-life situation, especially for Future Internet.

Laboratories are less related to the other ICT RIs. They mainly *produce* ICT in form of physical devices with new technologies. Laboratories are not currently heavy consumers of ICT RIs – they do not produce much (digital) data, and in general they are not connected to the grid, because they do not have a particular need for access to data or computing power. The results that they produce have in general a direct economic impact in the short to medium term. As a result, laboratories often compete to have the results first. Only recently, the increasing cost and complexity of the laboratory RI has forced them to start some cooperation in order to become more complementary, since no single RI can cover the entire field of ICT technologies anymore.

2.1 Inventory of ICT RIs Environments

The five elementary objects identified helped us to clarify the intrinsic nature of the ICT RIs as follows:

- *networks* are the structural element behind the *Research Network Environment*; they make the structure also for the *Research DCI framework*, as DCIs are based on a hybrid model, that mixes computing, network and storage in order to obtain the best efficiency compromises;
- *facilities* are mainly a characteristic of HPC; at a smaller scale, they are constitutive elements also for DCIs and for *Research Data Infrastructures*;
- *instruments* are the target of the *Remote Instruments access model*, which is an ICT interface to instruments;
- *testbeds* are required in order to outline the *Future Internet vision*, which is first and foremost a standardization effort;
- *laboratories* are the principal objects around which is built the main structure of the *MNT collaboration framework*; the MNT model has however also other interesting features (interconnection, complementarity / no overlapping, integration, aggregation of capabilities).

We complete our analysis by selecting the following **ICT RIs environments** as the main reference points to identify relevant cases of current cooperation between Public Authorities and National Champions and existing and future ICT RIs, namely:

(E1) European and National *Network* RI environment,

(E2) (Grid & Cloud) European and National Research *Distributed Computing Infrastructure* (DCI) framework,

(E3) (HPC) *top parallel computing* RI ecosystem,

(E4) MNT collaboration *facilities interchange* RI framework,

(E5) *Research Data Infrastructure* framework,

(E6) *Remote Instruments* access model,

(E7) *Future Internet* (FI) service-oriented vision.

In the next paragraphs, a brief description of each will be given.

2.2 The European and National Network RI Environment

The large bandwidth European Research Network provides to the research sector the connectivity over which other pan-European RIs can be put into operation. The network service use case is the simplest in the group and its governance model the oldest – GÉANT is currently celebrating ten years of operation.

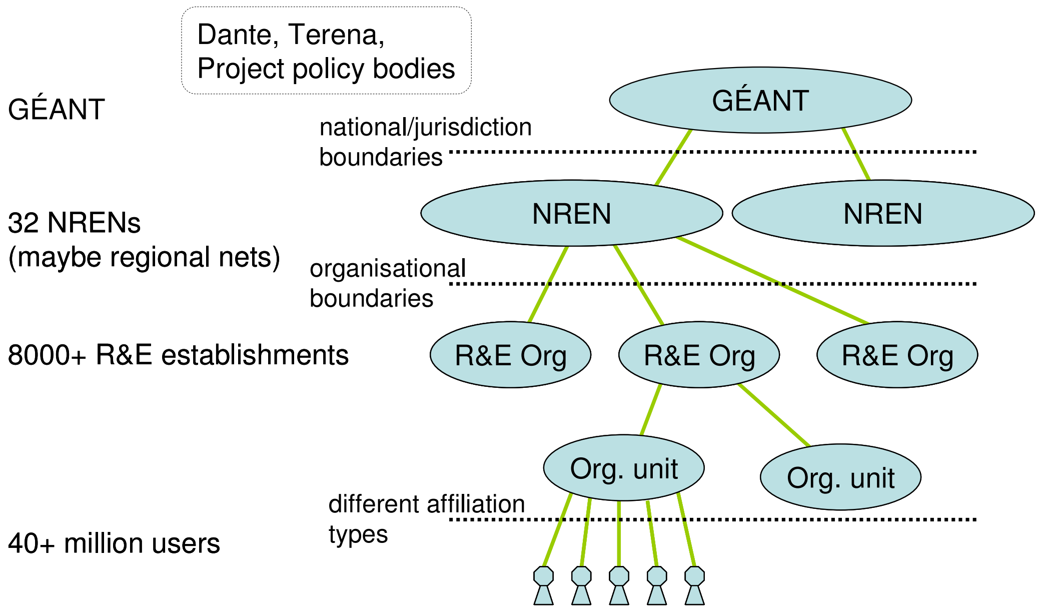

Fig. 1 – The Network RI Governance Model (image modified from a 2007 SWITCH NREN presentation).

Its mature governance model is structured on the per-country specialized internet service providers, namely *National Research and Education Networks* (NRENs), supporting the needs of the research and education communities. The NREN Policy Committee, with one representative per country, provides the general policy and coordination body at European level. The GÉANT series of projects (currently GN3) are co-funded by the NRENs and the EU to guarantee the harmonized necessary innovation and future evolution. GÉANT is coordinated by the managing partner, namely *Delivery of Advanced Network Technology to Europe* (DANTE), which operates the European general backbone, and supported by the dissemination and research partner *Trans-European Research and Education Networking Association* (TERENA), which provides a Forum for the coordination of the R&D efforts between the NRENs.

Other pan-European RIs have taken the NRENs' governance model as a reference.

Around the globe, other multi-domain networking research networks use a diverse range of technologies, operational approaches and procedures within their own individual networking domain. There is a need to deliver services and applications that can traverse these domains and provide a seamless networking experience to the user. GÉANT is also involved in finding networking solutions to facilitate this concept of "multi-domain" networking services both across Europe and on a global scale.

International network collaborations are active (see the following Fig.1) with North America (DICE), South America (ALICE2/RedCLARA), Africa (EUMEDCONNECT, UbuntuNet Alliance), India, South Est Asia (TEIN3, CAREN, BSI).

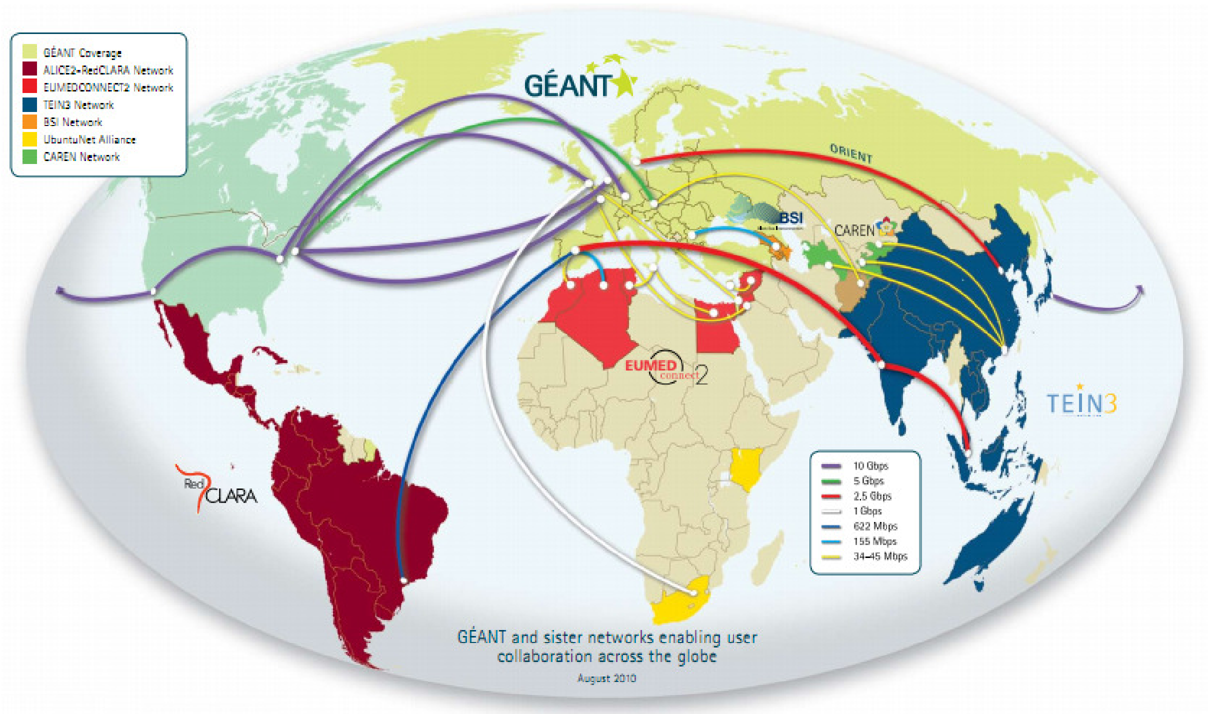

Fig. 2 – The GÉANT Network Environment and its worldwide interconnections (image from www.geant.net).

2.3 The European and National Research DCI Framework

Research Distributed Compute Infrastructures (Grids&Cloud)

The pan-European *Distributed Computing (& Storage) Infrastructure* (DCI) provides general access and sharing services to more than 300 Compute and Storage Resource Centres in Europe; the *European Grid Initiative* (EGI) consolidates ten years of research and developments, but also of operation and production-level delivery of service, following the *Enabling Grids for E-sciencE* (EGEE) series of European projects.

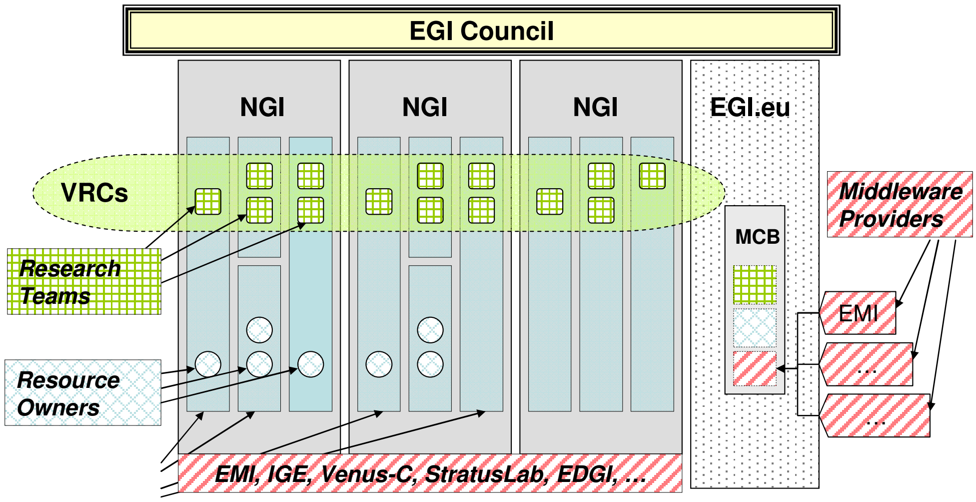

Fig. 3 – The EGI Governance Model (image modified from EGI_DS documents).

The EGI infrastructure is connected world-wide through *regional collaborations* with Latin America (current and past projects ACTION-GRID, EELA-2, EELA), Asia (current and past projects EUAsiaGrid, EUIndiaGrid, EUChinaGrid), Mediterranean and South East Europe (EUMedGrid, BALTICGRID-II, Baltic Grid, SEE-GRID-SCI, SEE-GRID2, SEE-GRID).

The DCI governance is put into effect by the EGI Council following the model of the *European Grid Initiative* (EGI) which was created as an evolution of the NRENs' model that could accommodate the added complexity of middleware R&D and of user community management. The EGI model, designed according to the framework of the EGI Design Study project and then later refined, is grounded on the per-country *National Grid Initiatives* (NGIs), responsible for the national grid e-Infrastructure and for maintaining relationships with customers and by the EGI.eu organization which provides global services to all NGIs and by the EGI Council which provides the general coordination and policy decision framework. .

The EGI *Integrated Sustainable Pan-European Infrastructure for Researchers in Europe* (EGI-InSPIRE) project has as its goal to guarantee the overall operation of the European DCI together with its innovation and harmonized evolution and is co-funded by the NGIs and the EU.

The *middleware* required was not available from commercial providers at the time of the creation of scientific DCIs, so several years of research and development were needed to provide a production-class middleware layer. Contrary therefore to the Networking RI which is based on commercial products, the DCIs are based on software services provided by middleware providers coming in general also from the Research sector – like those behind the current European projects such as the *European Middleware Initiative* (EMI) for ARC, gLite, UNICORE & dCache, the *Initiative for Globus in Europe* (IGE), EDGI for desktop computing and Venus-C and Stratus Lab for the cloud environment. All these are working to support and evolve the EGI model with activities of research and development in the software domain.

As key stakeholders, the *Virtual Research Communities* (VRCs), often involving a big *European Intergovernmental Research Organisation* (EIRO), are managed in EGI via specific support processes caring for them and for new user communities development.

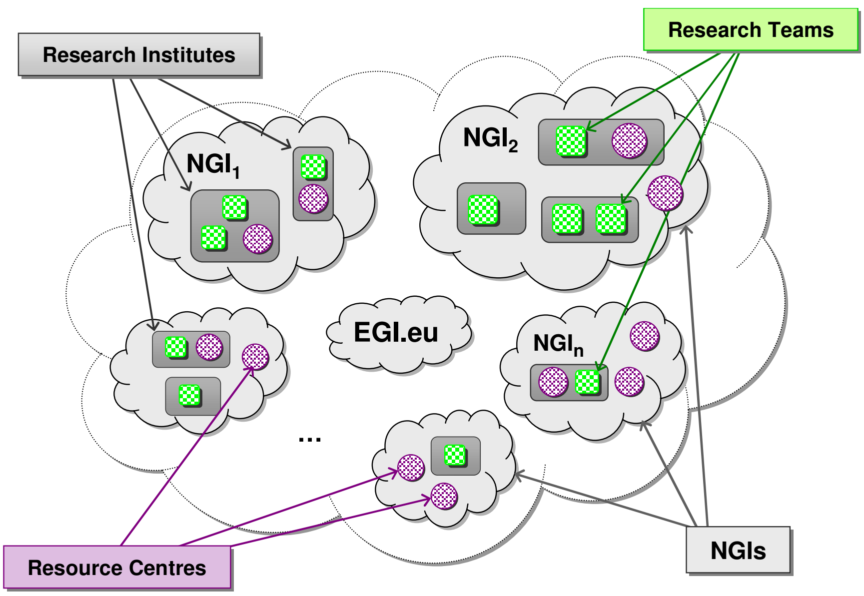

Fig. 4 – EGI Stakeholders. Research Teams are organized in thematic VRCs (image modified from EGI_DS documents).

The service model of grid & cloud DCIs

Scientific computing grids generally adopt a *shared use* model: user communities share distributed computing and storage resources, obtaining shared access to grid resources. The implicit extensibility of this model to new user communities and its intrinsic sustainability explains the popularity of grid computing in the research community. The counterpart is that users (and site managers) have to manage the added complexity of the middleware layer.

The *cloud paradigm* is introducing further flexibility and elasticity in the DCIs offering and has favoured the decoupling at the same time of the physical layer from the software environments available to the users. On the other hand, it favours a rapidly growing service model that is warmly accepted also in the private industry sector thanks to its simplicity. The introduction of cloud services in DCIs will allow research teams to acquire the resources, or better the software environments they need, in a much easier way together with the possibility to pay for what they use. Clouds offer to users scalable and elastic access to computing resources leaving at the same time complete freedom for the software environment choice.

International policy, dissemination, technical standardization and grid-cloud interoperability are coordinated via specific initiatives, notably the *Standards and Interoperability for eInfrastructure Implementation Initiative* (SIENA), that aims to “define scenarios, identify trends, investigate the innovation and impact sparked by cloud and grid computing, and deliver insight into how standards and the policy framework is defining and shaping current and future development and deployment in Europe and globally”.

2.4 The (HPC) Top Parallel Computing RI Ecosystem

The European *High Performance Computing* (HPC) infrastructure is targeted towards providing the highest sustained parallel peak computing capacities to scientific research communities. The Partnership for Advanced Computing in Europe (PRACE), which has already entered into operation will maintain as a single European entity, a pan-European HPC service computing infrastructure consisting of up to six *top of the line* leadership systems (Tier-0) which are well integrated into the European HPC ecosystem. Each system will provide computing power of several Petaflop/s, in the longer term targeting the Exaflop/s computing power range. Four nations (France, Germany, Italy and Spain) have agreed to provide 400 million Euro to implement supercomputers with a combined computing power in the multi Petaflop/s range over the next five years. This funding is complemented by up to 70 million Euros from the European Commission which is supporting the preparation and implementation of this infrastructure.

As is common in the supercomputing field, the PRACE service model requires that researchers from across Europe apply for HPC time from a series of hosting nations via a central peer review process; calls for proposals for computer time on PRACE RI machines are issued regularly and have strict closing dates. Computing resources, in the HPC use case, are scarce, so user communities have to compete against each other in order to have access to them. The pan-European peer review system for PRACE is laid out in the statutes. It is based on the principle “the best systems for the best science” and takes into account the results of the HET group and an analysis of national HPC peer review systems. Throughout the [PRACE D2.4.2, 2008] deliverable, the eight principles for PRACE peer review (*Transparency*, *Expert Assessment*, *Confidentiality*, *Prioritisation*, *Right to Reply*, *Managing Interests*, *No Parallel Assessment*, *Ensure Fairness to the Science Proposed*) have been considered.

PRACE is cooperating with and following the *Distributed European Infrastructure for Supercomputing Applications* (DEISA) series of projects, targeted at the consolidation of a persistent European HPC ecosystem. The current DEISA HPC infrastructure is based on the tight coupling of eleven leading national (Tier-1) supercomputing centres, connected via the GÉANT/NRENs high speed network infrastructure. The DEISA consortium has deployed middleware that enables the transparent access to distributed resources, high performance data sharing at European scale, and transparent job migration across similar platforms.

On June 9th 2010, the non-profit international association of nineteen members named "Partnership, for Advanced Computing in Europe AISBL"[3] was set up in Brussels to represent PRACE; the PRACE Council has held its constituting meeting, and elected the Council Chair and the *Board of Directors* (BoD) Chair; an SSC Establishment Workshop, organised by six independent representatives of the European scientific communities then nominated the candidates for the PRACE *Scientific Steering Committee* (SSC). With the Council, the BoD, and the SSC in place, PRACE has implemented the main bodies of the governance structure that was developed in the project.

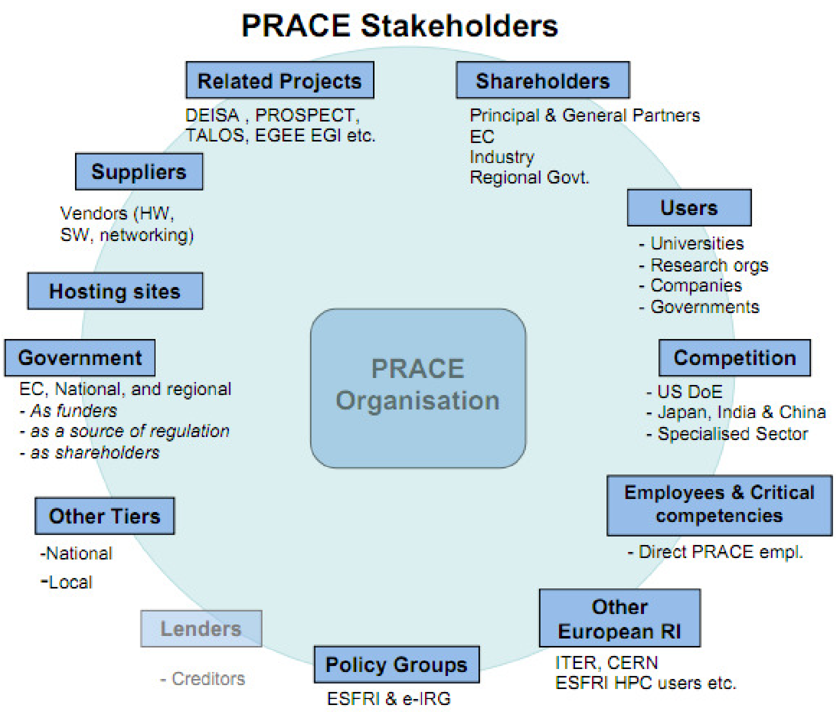

Fig. 5 – PRACE Stakeholders (image from PRACE D2.2.1 Deliverable).

The four Principal Partners (called hosting members in the PRACE AISBL) countries, namely Germany, France, Spain and Italy, have already now made binding commitments to contribute Tier-0 resources each with a value of € 100 Million in the next five years; the Netherlands followed later in 2010. The AISBL currently has 20 members representing the 14 countries that participated in the PRACE project plus six additional countries (Bulgaria, Cyprus, Czech Republic, Ireland, Serbia, and Turkey) that have joined the PRACE Initiative since the project started in 2008.

The first production system, already in operation, is a one Petaflop/s IBM Blue Gene/P (JUGENE) installed at Forschungszentrum Jülich, Germany, a Gauss Centre for Supercomputing member site. The second production system, a 1.6 Petaflop/s BULL system called CURIE will be located near Paris and operated in a new computing centre, the *Très Grand Centre de Calcul* (TGCC), funded by *Commissariat à l'Énergie Atomique et aux*

[3] In [PRACE D2.1.1, 2008] a detailed benchmark between different legal entities was done.

énergies alternatives (CEA). This BULL system, having a general purpose architecture, will extend the PRACE HPC services so as to significantly increase the high performance computing capability span across the scientific needs of European researchers. The third world-class tier-0 supercomputer announced by PRACE is SuperMUC, that will start operation in mid 2012 and will be one of the fastest general purpose supercomputers in the world with a 3 Petaflop/s peak performance, 320 TeraBytes of main memory and 12 PetaBytes of permanent storage. SuperMUC is to be located at LRZ (*Leibniz Supercomputing Centre*, also a Gauss centre for supercomputing member site). SuperMUC being based on current general purpose architecture will allow scientists to use established programming models.

It has to be noted that *PRACE facilitated the transition from competition to cooperation mode among the European HPC centres, which is an achievement per se*. It has lead to *an exchange of best practices and expertise* in all aspects of HPC service provisioning among leading European centres, ranging from *procurements*, *infrastructure*, *operation* and *management*, *user support*, and *peer review* to *algorithms*, *code development* and HPC *technology*. These have been documented and not only serve as a model for the PRACE Tier-0 centres, but also for new national and topical centres – especially in the new member states.

Regarding industry collaborations, PRACE has raised significant interest amongst European and international HPC vendors and technology providers. This has been leveraged to create a framework for information exchange and cooperation between PRACE and the HPC industry through the *PRACE advisory group for STRAtegic TechnolOgieS* (STRATOS). This structure will eventually foster the further development of a European HPC industry.

2.5 The MNT Collaboration Facilities Interchange RI Framework

Considered as a research infrastructure, *Micro & Nano Technology* (MNT) is built around a limited number of costly facilities and laboratories, each requiring heavy investments and highly specialized personnel to maintain them. Due to its large fabrication facilities and clean rooms requirements, *microelectronics* is one of the most capital intensive production and research activities; *nanotechnology* has more moderate requirements for its facilities, but still needs costly instrumentations and laboratories. Both will be considered in the following, but it should be clear that the dimensions are different.

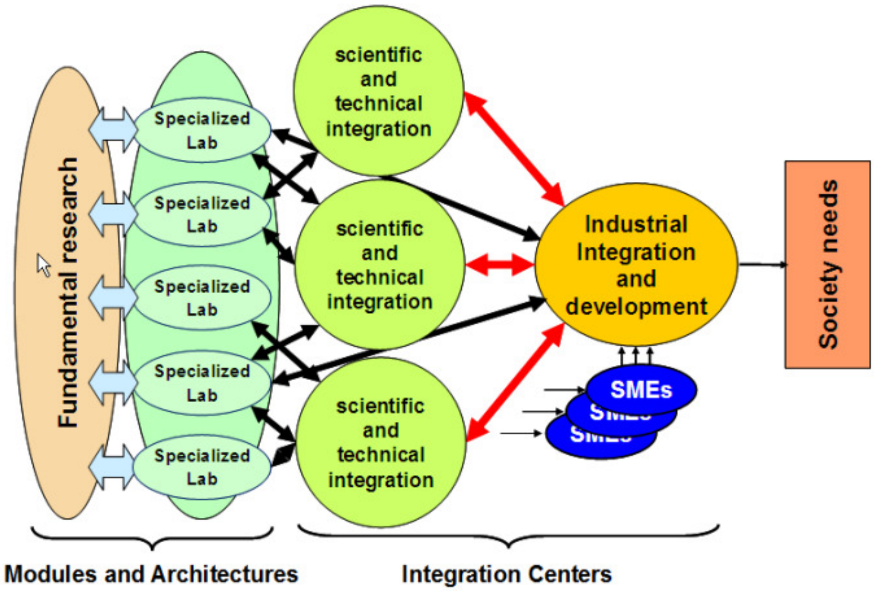

Fig. 6 – Links between Industry and Academia in MNT (image from PRINS documents).

MNT Domain Classification

More specifically, MNT ICT technologies span five major domains:

a) *Micro/Nano-electronics*. This technology enables the fabrication of microprocessors, memories, and RF communication devices. The investments for a state-of-the-art facility have become extremely heavy (multi-billion €), so only a few players are left in this field. In Europe, CEA-LETI, IMEC and some Fraunhofers are the only RIs left in this field.

b) *MEMS*. This domain originally started by using the microelectronics technology for making devices that interacted with a non-electrical domain. Examples of MEMS devices are accelerometers microphones, and fibre-optical switches. Many RIs exist in this domain, while the number of new applications of the technology is rapidly diminishing.

c) *Electro-optics*. This technology is similar to microelectronics, but uses III-V semiconductors to fabricate solid-state lasers, LEDs and some high-frequency circuits. This technology uses very old micro-electronics technology, combined with complex machines for the growth of the semiconductors.

d) *Top-down nano-technology*. This technology uses some very complex machines (e-beam writers) to define extremely small structures that interact at the molecular level. One example of such technology is a filter that consists of a membrane in which very small pores have been defined using an e-beam.

e) *Bottom-up nanotechnology*. This technology also defines structures that interact at the molecular level, but uses combinatorial chemistry to define them. An example is a filter which has an alumina membrane that has been made nanoporous by anodization.

MNT RI Collaborations

The *Pan-European Research Infrastructure for Nano-Structures* (PRINS) project aims at enabling European innovative research for the ultimate scaling of electronics component and circuits. The PRINS platform is searching for the convergence of "top-down" microelectronics technology with "bottom-up" methods derived from fundamental disciplines such as material physics, chemistry, biotechnology and particle electronics. This research infrastructure should enable the cross-disciplinary fertilization of academic and industrial competences in the areas of nanoelectronics, nanosystems, nanobiology, nanophotonics, etc.

The *Micro and Nano Technology Europe* (MNTEurope) project has led to the creation of a unique distributed platform for research and development in the field of Micro and Nano Technologies (MNT), based on the national facilities of its five partners, Léti (France), CSEM (Switzerland), IZM-M and IISB Fraunhofer (Germany), IMEC (Belgium) and Tyndall (Ireland) research institutes; the *MNT Europe Alliance* currently offers a list of 227 MNT *technology capabilities* available in their respective national facilities.

The *European Integrated Activity of excellence and Networking for Nano and micro-electronics Analysis* (ANNA) offers analytical support to Research Institutions and also production facilities of SMEs and Industries. The ANNA consortium comprises 12 partners from seven European countries. The strength of this network is based on the combination of complementary backgrounds of the partners. The consortium covers research institutions, universities and industry.

According to the project analysis, the MNT RI should be distributed and created as an extension of current Research Organizations (RO) facilities, as it appears improbable a strong, public-supported, effort to create a centralized MNT RI. The collaboration efforts are mainly directed towards the definition of common agreement for the access to the processing equipments and to consolidate a common view of all offered capabilities. Intellectual Property (IP) and legal aspects are of importance in this sector and must be clarified via contract agreements and definite procedures.

MNT research infrastructures have several points of contact and share several characteristics with nanobiotechnology, so it could be useful to learn from parallel nanobio experiences. The EuroNanoBio study has shown that a distributed infrastructure based on a coordinated pan-European alliance of research driven clusters is the most efficient instrument for fast and efficient translation of nanotechnological discoveries into biological and medical applications. We highlight in the following (with italics) the relevant facts emerging from the fifteen recommendations reported in the study:

R1. To cover the large range of scientific disciplines involved in nanobiotechnology and the diversity of application areas, a European infrastructure has to be built on *regional nanobio clusters*, which have *world-class facilities and expertise* with high levels of engagement between industry and academia.

R2. The nanobio clusters *need to be connected and coordinated to share knowledge and equipment* and to cover the whole value chain in specific application areas of nanobiotechnology such as environment or medicine, for example.

R3. A *dedicated infrastructure management* should improve the engagement between academic disciplines, research centres and companies inside and between the involved clusters.

R4. *Clear technical roadmaps* for each of the application areas within nanobio should be defined *to provide a catalyst for collaboration between industry and academia* within the infrastructure.

R5. Experts should be encouraged to work collaboratively with science departments, research institutes and industry to help explore *Ethical, Legal and Social Aspects* (ELSA) of developing nanobiotechnology thereby enabling early decision making about the probability of commercialisation in a socially and ethically responsible manner.

R6. Set-up and upgrading of clusters will require *local, national and European political support and funding* supplemented by *private investments at a later more mature stage*.

R7. A *European reference centre* is needed for *characterization and toxicology studies of nano-objects*, which can be accessed by all nano-object producers and users from academia and companies similar to the Nanotechnology Characterisation Lab at NCI/USA.

R8. A *European centre for Risk and Safety Management* should be established, which provides information and advice about handling of nano-objects and protection measures to SMEs and universities that cannot afford expensive risk assessment.

R9. *Clusters should help especially SMEs* to articulate their needs and interests to regulatory and standardisation bodies.

R10. The infrastructure *should provide pools of experts* and *professional communication tools* necessary for engagement with the public.

R11. Promotion of the capabilities of nanobiotechnology to SMEs and clinicians should be facilitated by *showcasing examples of successful exploitation of nanobiotechnology*.

R12. Engagement of the European infrastructure with nanobio clusters and research centres *outside Europe should be encouraged*.

R13. The *highly interdisciplinary nature of nanobiotechnology* requires the integration of dedicated nanobio modules *preferably at the MSc or PhD level*.

R14. Because nanobiotechnology touches on many important wider issues, *teaching an understanding of ethical and social aspects* together with *training in science communication and public engagement* should be *included* at the MSc and PhD level.

R15. Due to the rapid development in nanobiotechnology *targeted education and training programmes for in-career training need to be developed*.

2.6 The Research Data Infrastructure Framework

Research Data Infrastructures (RDI) are represented by various initiatives involving one or more domains but are currently missing a general coordination framework. The most popular components included in RDI are “Scientific Repositories” and “Digital Libraries” and variations.

As reported by the EC (see e-Infrastructure: “Scientific Data, Accessing Knowledge” leaflet):

“Scientific data have come to the forefront of modern science; it is the new way to express knowledge. Many disciplines are developing into highly data-intensive areas of science such as high-energy physics, astronomy, bio-informatics, genomics and medical imaging. Scientific experiments, observations, theories, models and simulations generate unprecedented volumes of data which will quickly reach the Exabyte scale It is stored in complex databases consisting of numbers, text, images, diagrams and formulas.”

“The European Commission has been actively promoting and funding digital libraries policies, initiatives and research over the last decade. The i2010 Digital Libraries initiative has been integrated and continued in the new policy framework “Digital Agenda for Europe”, one of the seven flagship initiatives of the Europe 2020 strategy for economic growth. In November 2007 Europe’s cultural institutions created a foundation that provides for the organisational underpinning of the service. A demonstration site of Europeana was presented in February 2008. A full prototype was launched in November 2008.”

We recall here (the italics are ours) the conclusions of the *Data Management Task Force* (DMTF) reported in the e-IRG Report on Data Management [e-IRG DMTF, 2009]:

- A very large number of *individual* and *domain-specific* data initiatives and scientific databases exist in Europe.
- *Only a small number* of the data initiatives are *federated*.
- Interoperability is de facto limited to a particular scientific domain which applied standard formats at an early stage of the data initiative.
- *Long-term sustainability is a major issue for all data initiatives*, not only is the problem limited to the underlying hardware infrastructure *but also to software* accessing and exploiting the data.
- *Open access* to data has *not yet become a reality* in *all* scientific domains. It is often hampered technically by the absence of search engines, and institutionally by the absence of *clear policy guidelines*.
- Organisations are moving away from a centralised stand-alone model towards *distributed networks of federated data repositories*.
- New requirements for *cross-disciplinary* research will require interoperability between different disciplines and different types of data.
- *Metadata* is recognised as being paramount for long-term data access and usability (including documenting the research process, not just the data itself).
- The suite of data analysis tools is growing and becoming more complex (i.e. the use of GIS in many fields).
- The focus is on curating data for reuse, not necessarily for *long-term preservation*,
- *Open access* to data is becoming more common.

- New projects, many of them financed by the EU, will profoundly change the current data landscape in Europe and set standards for the rest of the world.
- Additional efforts should be put into gathering more detailed information about project-specific needs on data management and what kind of activities are required.
- *Communication and cooperation between the data initiatives/projects should be stimulated to achieve better interoperability and reuse of solutions and infrastructures.*

In the Final Report [e-SciDR, 2008] twelve mutually reinforcing recommendation sets for policy measures were summarized towards driving a European e-infrastructure for and of e-Science digital repositories.

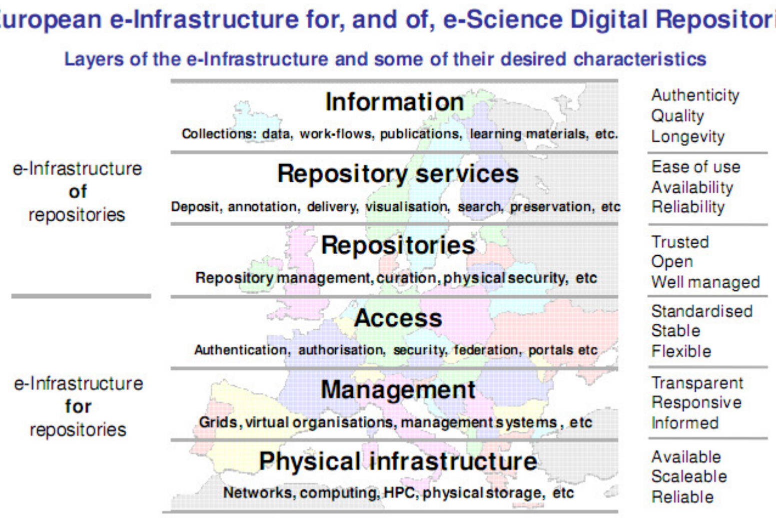

Fig. 7 – The Vision of an European e-Infrastructure for/of e-Science Digital Repositories.

The recommendations were grouped according to perspective and level:

A. *Building an e-infrastructure* for research continuity: funding reform, European e-Infrastructure of and for European e-Science digital repositories, governance and management;
B. Engaging *users* and *service providers*: support for data producers, trust and recognition, training and awareness;
C. *Providing access* to researchers, educators, students, unaffiliated stakeholders and interest groups: discovery and navigation, *open access* to publicly funded data, Legal issues, International;
D. *Maintaining and preserving information*: collections management, selection and appraisal for sustainability, preservation of digital information.

2.7 The Remote Instruments Access Model

As noted in [ESFRI, 2010], European RIs' *instruments and facilities* – such as *synchrotrons*, databases, *telescopes*, *sensor networks*, and *biomedical facilities* – are an unprecedented asset. (In Europe) there are *more than 500* of them, of which at least 300 have strong international visibility, attracting world class researchers. These RIs represent an aggregate European investment of more than €100 billion. Some 50,000 researchers a year use them to produce 3,000 to 6,000 high-impact research papers annually – as well as a chain of patents, spin-off companies and industrial contracts. […] RIs provide the means and impetus to develop a truly sustainable e-infrastructure to store, share and protect digital data. This permits Europe to lead the development of e-science. […] Advanced computer and communications technologies are changing not just the tools of science, but also the methods. *Scientific e-infrastructure permits researchers in one place to undertake experiments on RIs remotely, in real time*; to model, simulate or infer conclusions from vast data sets; and to collaborate with researchers of widely different backgrounds and disciplines. Some [Hey, Tansley, Tolle, 2009] see this creating a "fourth paradigm" of science – beyond observation, theory and simulation, and into a new realm of correlation to mine new insights from vast, diverse data sets.

The increasing availability of instruments and facilities (*physical* infrastructures) for researchers, combined with the flexibility and versatility of the deployed e-infrastructures (*digital* infrastructures) is opening a completely new paradigm for the conduction of science experiments, where scientists will have the possibility of interacting remotely with their instruments, reducing the geographical distances thanks to the *Remote Instruments access model*.

The EU FP7 *Deployment Of Remote Instrumentation Infrastructure* (DORII) Project (http://www.dorii.eu) complemented the Grid middleware with the Instrument Element (IE) in order to enable remote access to instrument resources like:

- seismic sensors and actuators for *Earthquake observation*,
- conductivity, temperature, depth (CTD), pressure, oxygen and turbidity *sensors*, digital cameras, Autonomous Underwater Vehicle (AUV) and others for *Environmental sciences*,
- X-ray scattering, detectors, beamline, synchrotron radiation in medial application, for *Experimental science*.

IT technology, provided by partners who have recognised the initial requirements of the scientific communities, worked to deliver the functionality developed in projects which had a focus on accessing remote instrumentation (GRIDCC, RINGrid), on interactivity (int.eu.grid), on software frameworks for application developers (g-Eclipse) and advanced networking technologies (GN2) with EGEE based middleware.

2.8 The Future Internet Service-Oriented RI Vision

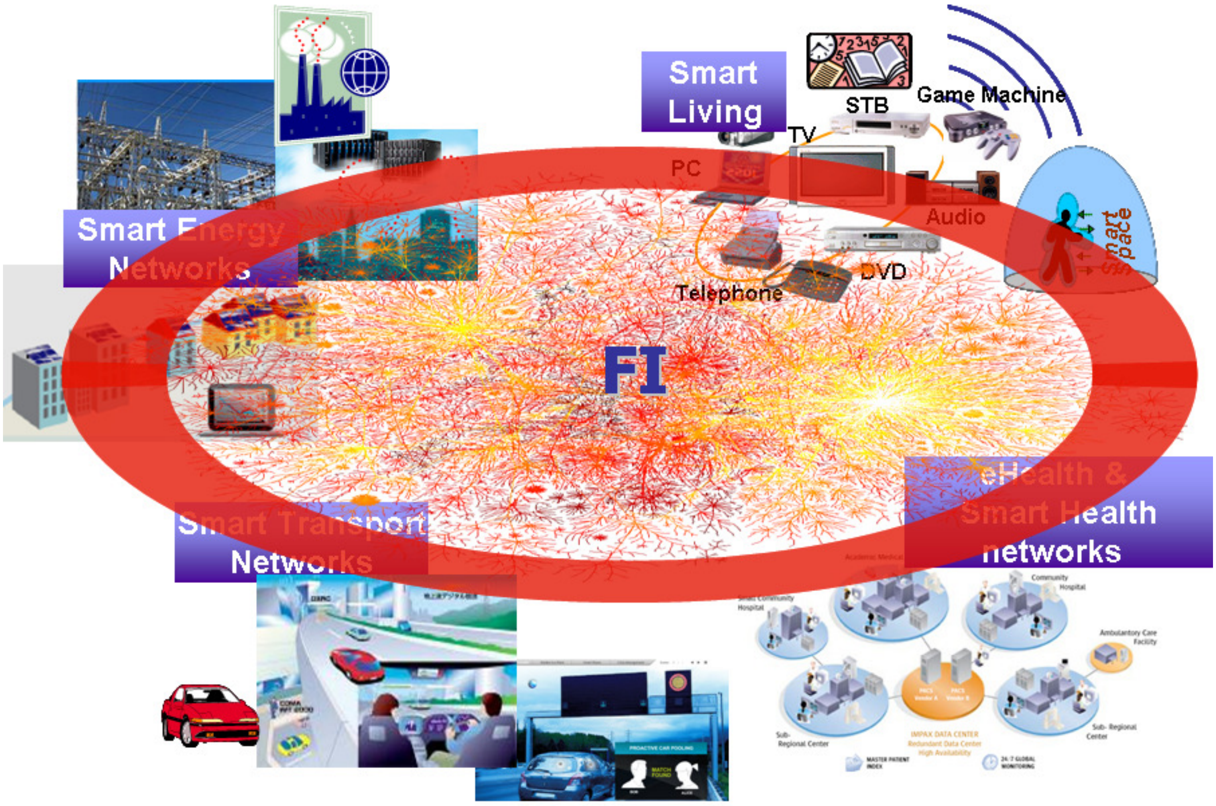

Fig. 8 - The Future Internet Vision.

The following *usage trends* for the Future Internet (FI) were suggested in the [WGFI, 2008] report:

1. *Mobility*, that needs a *native support* in an infrastructure conceived for fixed usages;
2. End to End *very high rate throughput*, reaching the limits of current Internet protocols, not designed for ultra broadband scenarios;
3. *Security and Trust, Privacy*, to be supported in FI directly in the service and network infrastructures (Privacy by design);
4. The "internet of things" – device connectivity, coupling of virtual world data with physical world information (RFID, sensors), needing an FI with high network architecture scalability and new protocols and service architectures to support device generated traffic and service discovery;
5. *User generated services*, as a follow up to user generated content, via new FI service architectures enabling dynamic, secure and trusted service compositions and mash ups;
6. *3D* becoming mainstream, requiring a FI capable of resource intensive usage of computing, networking platform and new standards;

7. *Negotiated management and control of resources*, negotiated SLA's, to be enabled via FI dynamic and predictive network management, infrastructure observability and controllability;
8. *User controlled infrastructure*, requiring FI ad hoc network and service composition.

New, FI-related infrastructures, still to be defined, should satisfy these new requirements.

Fig. 9 - EC-funded Future Internet projects (image from www.future-internet.eu).

In view of possible FI Research Infrastructures, the same document enumerates the main issues to be addressed for an EU wide approach to FI:

- *Networking* – networking of actors, within and across the national boundaries, as in Eureka and Framework Programme; networking of knowledge, for a multi-disciplinary approach, such as in the "Living Lab" and open innovation models; networking and development of skills, considering the experience of the *Knowledge and Innovation Communities* (KICs) of the *European Institute for innovation and Technology* (EIC); networking of users, that should be involved early in research stages, in a bottom-up process; networking also outside Europe, in a global approach;
- Changing the main target, *from technology to applications* – the use of experimental testbeds and facilities is suggested as an effective mechanism to account for application requirements into technological developments. Tighter links between FI and GÉANT (and FEDERICA) could help to federate interconnection of testbeds and sustain a virtuous circuit of innovation for the development and testing of new technologies and solutions. From the deployment perspective, focus should be on applications of public interest whose implementation can be supported by public

administrations. Health care, education, e-science, e-administration are typical application areas to consider;

- Addressing *the right mix of actors* – smaller users, SME's and innovative actors are called upon to actively contribute to the process where significant industry driven work is currently on-going to define a cross domain perspective of the FI;
- The identified *common topics* – rethinking the *network management* to fit the widest possible application requirements, *mobility* and end-to-end very high speed rates, the "Internet of things", the "Internet of services", new model of *massive 3D usages* for entertainment and public applications like health, information networking (saving, sharing and delivering), security/trust/identity, advancing federated testbeds experimental infrastructure.

3. Inventory of PA/NC-RIs Collaborations

In the following, different kinds of PA/NC-RIs collaborations are enumerated.

3.1 The European and National Network RI environment

(C1) National Research & Education Networks (NRENs) Collaboration

A *National Research and Education Network* (NREN) is a specialised internet service provider dedicated to supporting the needs of the research and education communities within a country. It is usually distinguished by support for a high-speed backbone network, often offering dedicated channels for individual research projects.

(C2) Network Backbone (GÉANT) Project Collaboration

The GÉANT project is a collaboration between 34 project partners comprising 32 European NRENs, DANTE and TERENA; and four Associate NRENs. Together, GÉANT and the national networks create a common pan-European service area (known as the *GÉANT Service Area*) enabling advanced network services and applications harmonised across GÉANT to be offered by NRENs at local level to institutions, projects and researchers.

(C3) Managing partner (DANTE) Collaboration

Delivery of Advanced Network Technology to Europe (DANTE) plans, builds and operates advanced networks for research and education. It is owned by European NRENs, and works in partnership with them and in cooperation with the European Commission. DANTE provides the data communications infrastructure essential to the development of the global research community.

Currently fifteen shareholding organisations, either NRENs or National government funding bodies own DANTE. The shareholders meet twice per year to discuss issues related to DANTE company policy, financing and management. The shareholders are represented by a *Board of Directors* (BoD) to whom responsibility is delegated for overseeing the running of DANTE. The Directors represent the fifteen shareholders and hold meetings four to six times a year. They oversee the running of the company, and liaise with the General Managers of DANTE on issues of company management and policy.

(C4) Dissemination & research partner (TERENA) Collaboration

Trans-European Research and Education Networking Association (TERENA) offers a forum to collaborate, innovate and share knowledge in order to foster the development of Internet technology, infrastructure and services to be used by the research and education community. Its core business is to bring together managers, technical specialists and other people in the research networking community with their counterparts from other countries in Europe, mobilising the expertise and experience of hundreds of professionals in the research and education networking area.

The *General Assembly* (GA) is the highest authority in the TERENA organisation. It is comprised of representatives of the TERENA members and meets twice a year to discuss policy issues concerning the association. The GA elects the members of the *TERENA Executive Committee* (TEC) who are responsible for managing the organisation. The TEC manages and administers the TERENA organisation. The *TERENA Technical Committee* (TTC) is a group of experienced and knowledgeable networking professionals whose task is to coordinate and supervise the TERENA Technical Programme. The *TERENA Advisory Council* (TAC) usually meets once a year at the annual TERENA Conference. It provides the TERENA members with an opportunity to advise on the direction of the Technical Programme, suggest new initiatives, exchange information with each other about national networking activities, and propose TTC members.

3.2 The European and National DCI Framework

(C5) National Grid Initiatives (NGIs) Collaboration

The NGIs are legal national organisations responsible for:

- establishing, managing and operating the *national Grid infrastructure* to an agreed level of service and ensuring its integration with the European e-infrastructure;
- maintaining relationships with its *national stakeholders: Research Institutes* (RIs) and Research Projects carried out by *Virtual Organizations* (VO) constituted by *Research Teams* (RTs); *Resource Providers* (or centres) which offer resources to support the computing needs of the RTs in the country. The NGI and its Resource Providers form a national "business alliance" to jointly develop and "sell" a specific national marketplace solution (the national grid infrastructure) to their national researchers, each with its specific responsibilities;
- mobilizing the *national funding and resources* to guarantee the sustainable availability and operation of the national grid infrastructure as required by national users and to contribute to EGI.org for the common tasks;
- *representing* all its national stakeholders in the EGI Council and in the relations with EGI.eu; have the capacity to sign the Statutes of EGI.org – either directly or through a legal entity representing it;
- contributing to the decisions of the EGI Council and the EGI.eu technical bodies regarding international standards and to the EGI policies and quality criteria, and ensuring adherence at the national level to the agreed criteria;

- supporting *user communities* (application independent, and open to new user communities and resource providers).

(C6) EGI.eu Organization Collaboration

EGI.eu is a foundation established under Dutch law to create and maintain a pan-European Grid Infrastructure in collaboration with the NGIs and *European Intergovernmental scientific Research Organisations* (EIROs), to guarantee the long-term availability of a generic e-infrastructure for all European research communities and their international collaborators. EGI.eu is governed by a *Council*, which has representatives from all of its participants and is responsible for providing the long-term direction of the organisation, and the *Executive Board* which provides frequent guidance to the Director, who leads the organisation on a day-to-day basis.

(C7) European Grid (EGI InSPIRE) Project Collaboration

The ultimate goal of EGI *Integrated Sustainable Pan-European Infrastructure for Researchers in Europe* (EGI-InSPIRE) is to provide European scientists and their international partners with a sustainable, reliable e-Infrastructure that can support their needs for large-scale data analysis. EGI-InSPIRE will coordinate the transition from the EGEE project-based system to a sustainable pan-European e-Infrastructure. The four-year project will support grids of *High Performance Computing* (HPC) and *High Throughput Computing* (HTC) resources. The project is ideally placed to integrate new Distributed Computing Infrastructures (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit the user communities within the European Research Area. EGI InSPIRE follows the projects DataGrid (2002-2004), EGEE-I, -II and -III (2004-2010).

(C8) Middleware Provider Collaboration (EMI as example)

The middleware providers' main goals are to improve the reliability, usability and stability of middleware services, closely listening to the requirements of users and infrastructure providers. Different providers (EMI, IGE, Venus-C, StratusLab, EDGI) support several middleware technologies (Arc, gLite, Unicore, Globus, Cloud Middleware, Desktop Grids Middleware, …), consolidating and improving the existing middleware services.

(C9) Virtual Research Community (VRC) Collaboration

A *Virtual Research Community* (VRC) is defined as an organisational grouping that brings together transient *Virtual Organisations* (VOs) within a persistent and sustainable structure. A VRC must be a self-organising group that collects and represents the interests of a focussed collection of researchers across a clear and well-defined field. Named contacts are agreed upon by the VRC to perform specific roles and these then form the communication channel between the VRC and EGI. (*See: MoU VRC / EGI.eu template*)

(C10) Standardization OGF Europe Collaboration

The *Open Grid Forum* (OGF) & OGF-Europe are committed to driving the rapid evolution and adoption of applied distributed computing, which is critical to developing new, innovative and scalable applications and infrastructures essential to productivity in enterprise and within the science community. These goals are accomplished through open fora that build the community, explore trends, share best practices and consolidate these best practices into standards.

(C11) Standards and Interoperability Collaborations (SIENA)

Within the DCI community, there is the need to accelerate and co-ordinate the adoption and evolution of interoperable DCIs. The *Standards and Interoperability for Einfrastructure implemeNtation Initiative* (SIENA) collaboration is targeting this goal through engagement with other *Standards Development Organisations* (SDOs) and major stakeholders to forge community agreements on best practices and standards for distributed computing.

SDOs relevant to RI DCIs are: *Cloud Security Alliance* (CSA), *Distributed Management Task Force* (DMTF), *Open Cloud Consortium* (OCC), *Open Cloud Computing Interface Work Group* (OCCI-WG), *Object Management Group* (OMG), *Storage Networking Industry Association* (SNIA); the Grid Technical Committee in the *European Telecommunications Standards Institute* (ETSI), the *Open Geospatial Consortium* (OGC), the *International Organization for Standardization* (ISO), the *Organization for the Advancement of Structured Information Standards* (OASIS).

(C12) Outreach and Dissemination (e-Science Talk, SIENA)

As part of its International outreach programme, SIENA's ambition is to engage major stakeholders to forge community agreement on best practices and standards for distributed computing. This will be operationally sought through the *Industry Expert Group* (IEG), the *Special Liaison Group* (SLG) and the *Roadmap Editorial Board* (REB) made up of international experts.

e-ScienceTalk brings the success stories of Europe's e-infrastructure to a wider audience. The project coordinates the dissemination outputs of EGI and other European e-Infrastructure projects, ensuring their results and influence are reported in print and online. *GridCafe*, *The Digital Scientist – formerly international Science Grid This Week* (iSGTW), *GridGuide* and *e-Infrastructure Concertation Meetings* are its differentiated audience dissemination channels .

3.3 The (HPC) Top Parallel Computing RI Ecosystem

The following collaboration use cases are extracted from the PRACE project documents (D1.5, D2.1.1, D2.2.1, D2.5.1, D2.4.2).

(C13) HPC Ecosystem

The definition of an HPC Ecosystem can be visualised as a wide consortium including all aspects which stimulate the efficient usage of various kinds of computational resources producing high-class scientific results. The HPC service is often described by the Performance Pyramid, which consists of multiple layers:

- The European level *capability computing centres* (tier-0), which represent the highest available computing power, providing computing services to the top research groups across national borders and scientific disciplines;

- The *national and regional computing centres* (tier-1) with sufficient computing services for the HPC users and to facilitate the access ramp to the resources of the European level centres;
- The *local computing centres* (tier-2) in the university environments, research labs or in other organisations;
- The *personal computer or terminal resources* (tier-3) available to individual researchers.

The focus of PRACE is on the high end of computing and the enabling issues for it, i.e. the tier-0 and the HPC parts of tier-1. PRACE will coordinate and collaborate with other projects in those domains, most notably the DEISA 2 (tier-1) and EGI/EGEE (tier-1, tier-2 and tier-3).

(C14) PRACE Collaboration

The *PRACE organisation* is responsible for providing access to a suite of Tier-0 systems for use by European research communities. PRACE will be able to:

- define and implement a strategy for providing a world class HPC infrastructure in Europe,
- manage the formation of a suite of complementary Tier-0 systems in Europe,
- manage the contributions of partners,
- operate an open and fair access system based on peer review to the Tier-0 services,
- manage the interaction with industrial organisations wishing to access PRACE systems,
- perform training and computational science R&D activities,
- interact with multiple stakeholders in order to promote HPC in Europe with a long term and sustainable approach,
- deliver appropriate accounting, administration, human resources, marketing and communication activities,
- provide a secretariat to PRACE governance bodies.

(C15) Tier-0 HPC Centres Collaboration

The Principal Partners Committee (PPC, called hosting members in the PRACE AISBL) decides on tasks that are essential for the definition, hosting, construction, funding or financing of the Tier-0 HPC infrastructure. The MB takes all other decisions of strategic importance. PPC and MB have face-to-face meetings every 2 months. These project meetings are organised jointly with PPC and MB meetings of the PRACE Initiative.

(C16) Tier-1 HPC Centres Collaboration

Taking into account the PRACE Tier-1 General Partners is essential to achieving the goals of PRACE of developing HPC in Europe. The two Tiers are both an essential components of the constituency, and are extremely important contributors both with respect to HPC operation and procurement and in the area of leading HPC application to science and technology. The contributions of the General Partners may vary, depending on the funding model chosen, with some general partners paying more into the budget than others and some consideration of the financial contribution given should be reflected in voting rights.

(C17) HPC Hardware Vendors Collaborations

Hardware vendors are key stakeholders for PRACE since their solutions will eventually provide the high end computing platforms deployed in Europe. PRACE and vendors collaborate in various areas including prototyping activity, application benchmarking and other technical development. In addition, PRACE partners individually and vendors work together in commercial basis.

PRACE expects the vendors to be able to share technical and other relevant information, which might be helpful in developing the PRACE collaboration or services in the current computing environment. The actions shareholders need to take include active participation in various European HPC forums, proactive interest towards PRACE work and open attitude towards new ideas how to develop HPC activities in Europe.

The permanent research platform *Advanced HPC Technology Platform* (AHTP) is being implemented in PRACE. This includes the designation of the partners of AHTP from PRACE partners and collaborating industrial partners. AHTP will actively seek cooperation with European projects with the potential to contribute to petascale computer and communication components and will support similar activities in order to take benefit from all experiences and innovative ideas available throughout Europe. PRACE will also collaborate with recently established HPC collaborations targeting European industry, such as PROSPECT and TAL.

(C18) Industry Collaborations for Innovation

As a complement to the work that is aimed at the procurement, construction and operation phases of HPC services in 2009/2010, an equally important activity has been devoted to the evaluation and active initiation of research on future multi-petascale technologies beyond 2010. STRATOS, the PRACE advisory group for strategic technologies, has been created as the main vehicle to foster the development of components and technologies for future multi-petascale systems. Within STRATOS, partners from PRACE and industrial consortia, including more than eighty organisations, co-operate on the specification and developments of such components. The STRATOS MoU was signed in December 2008 by twelve PRACE partners and the consortia PROSPECT and Ter@Tec. Its long-term work plan comprises joint activities on Exascale Software, and Green-IT and HPC Leadership resources.

3.4 The MNT Collaboration Facilities Interchange RI Framework

(C19) MNT Research Organizations (ROs) Collaboration

(*see: PRINS Final Report*). The combined Research Infrastructure of the ROs will offer various access mechanisms for academia:

- hosting research teams in the RI and/or building up common R&D laboratories can enhance the long term collaboration between Academia and the hosting Research Organization;

- the mobility of researchers, which will be essential in order to leverage the impact of the research community and enhance the industry / research centre / academia linkages;
- the cooperation between Academia and the combined RI can also be achieved in a fast and cost-effective way by cooperating in scientific and technical projects between the RI and its associated laboratories;
- providing academic research teams with basic materials/data for their own research programs that are complementary and relevant to the combined RI will leverage the whole investment.

The access given by the combined RI must enable the academic community to perform advanced research in several domains and to remain at the leading edge of the research world. The use of state-of-the-art equipment in an industry-like environment will increase the industrial relevance of the research, thereby lowering the gap between academic research and industrial R&D.

Different kinds of Advisory Committees were identified in the PRINS project: for Research Organizations, a *Scientific Advisory Committee* (SAC); for Industrial Partners Collaboration, an *Industrial Advisory Committee* (IAC). Possibly it could activate a *Public Authorities Advisory Committee* (PAC).

(C20) MNT Strategic Alliances

(*see: MNTEurope DoW*) MNTEurope worked to integrate the Micro and Nano Technology (MNT) research capabilities in Europe into the form of a distributed platform for research and development. The collaboration model was developed through a strategic Alliance between the five major European Microelectronics and MNT Research Centres and through the creation of a common technology portfolio.

In order to make operational an Alliance that could implement organisational and strategic agreements, exchange of people / knowledge and the definition of common roadmaps the following networking activities were identified:

- creation of the environment and legal framework to allow the Integration of Activities which would significantly lower the barriers to collaboration between the participant organisations, in particular legal issues and those related to the lack of personal relationships will be targeted for improvement,
- preparation of the integration to build a common view on future developments;
- organization of the operational phase of the Alliance: global management, external collaboration, dissemination of the Alliance concept.

(C21) Transnational Access collaboration model

The objective of the (ANNA I3[4]) scheme is to sponsor new opportunities for research teams (including individual researchers) to obtain access to the facilities of the ANNA consortium. In a competitive access model similar to the HPC one, access to a given infrastructure will be granted following a selection of the proposals by “peer review”. Costs for access are covered by the EC.

[4] *See:* http://www.i3-anna.org/context.jsp?ID_LINK=2&area=8.

- European Community funded Transnational Access to infrastructure (laboratories, clean room, meteorology) at locations,
- access to instrumentation and analytical services is either in person ("hands-on") or remotely by suitable (electronic) communications,
- potential users of the infrastructure are researchers or groups of researchers from small & medium enterprises, large scale industry, research centres, or universities,
- interested users apply for research Access by submitting a short project proposal,
- selection of user proposals is by "peer review" on the basis of scientific/technical merit. Priority is given to first-time users and to users in countries without a similar infrastructure.

Each installation/facility is listed, described, and a temporal unit of access is specified (e.g., week, day). The installations are managed by the cooperating organizations that own them.

(C22) Nanobio Clusters

Concentration in geographical space of all the partners (R&D centres, universities / departments, hospitals, industry, entrepreneurs/spin off companies, tech transfer advisers, technological foundries…) whose contribution is required for innovation to happen in a particular, interacting and effective fashion to bring new products/services into the market/hospital. These interactions can be led and/or managed by an agency or office, with dedicated staff.

A nanobio *innovation cluster* is built around a nanobio *research cluster*. This is mostly due to its cross-disciplinary profile, characterized by the effective interaction of scientists and engineers from very diverse expertise (materials science, microelectronics, photonics, cell biology, molecular biology, inorganic and organic chemistry, medicine…) among each other and with high-tech SMEs and Spin-Offs. Geographical size varies largely from one to another. Roughly, in most cases, it will imply one large city and a suburban radius of 100 km.

We list in our inventory the nanobio clusters mainly as a useful case regarding MNT collaboration models. This case is also, directly, correlated to ICT RIs only when such clusters act:

(a) *as producers*, when nanobio collaborations that use biology or chemistry to develop completely new ICT devices such as new nano scale switches that could replace transistors,

(b) *as consumers*, (case perhaps less relevant), when nanobio collaborations use ICT tools to solve biological or chemistry problems. For example, biochips that use electronic or other components for detection biological processes.

3.5 The Research Data Infrastructure Framework

As reported[5] in the Communication from the Commission of 30 September 2005 to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – i2010: digital libraries [COM(2005) 465 final – Official Journal C 49 of 28.2.2008], "Digital libraries are *organised collections of digital content made available to the public*." The content is material that *has been digitised*, such as digital copies of books and other 'physical' material from libraries and archives. Alternatively, they can be based on information *originally produced in digital format*. This is increasingly the case in the area of scientific information, where digital publications and enormous quantities of raw data and software are stored in digital repositories.

(C23) Digital Libraries initiative

The i2010 *Digital Libraries initiative* covers both *digitised* and *born digital* material. Three areas have been identified:

1) Digitisation of Analogue Collections
2) Online Accessibility of Digital Content
3) Digital Preservation.

The identified key issues for the European Digital Library are:

- **Choice of materials:** the selection process should be 'bottom up'. Libraries and archives which own the materials should decide what is included as part of their own strategic development plans and available resources.
- **Meeting user needs:** the needs of the user should be central. Development will be demand-driven,but it is also important to take a longer term and visionary view of what the user will get from the library in the way of services.
- **Leveraging public/private partnerships:** at present, digitisation efforts in the Member States are progressing rather slowly. Public/private partnerships or private sponsorship can be a useful means to complement public funding and help accelerate digitisation.
- **European value:** financing digitisation is mainly a responsibility of the Member States. EU programmes can contribute where there is most European added value, for example by helping to aggregate digitised material across borders.
- **Intellectual property:** under current legislation only public domain works (where there is no longer copyright) can be made freely available to the public online. Hence, only works with specific rights clearances will be accessible through a digital library.

[5] See: http://europa.eu/legislation_summaries/information_society/l24226i_en.htm.

(C24) Scientific Repositories collaborations

Projects relevant to Scientific e-Infrastructures are:

- DRIVER II (e-Infrastructures and info-structures) – delivers a pan-European infrastructure *federating scientific repositories*. It uses open standards and supports complex information objects. Digital repositories will make both published and experimental data widely available for use and thus support new paradigms for research.
- PARSE.Insight aims to *help define the infrastructure needed to preserve and use the digitally encoded information* on which our society increasingly depends and which future generations will inherit.

Specific domain projects are also relevant e-Science use cases:

- EURO-VO-AIDA (astronomy): moving the astronomical European Virtual Observatory into a fully functioning operational phase,
- GENESI-DR (Earth observation): open and seamless access to Earth science repositories (space, airborne and in-situ sensors data),
- IMPACT (bio-informatics): improving protein annotation through coordination and integration of databases,
- METAFOR (climatology): common information model and tools for using climate data and models,
- NMDB (space physics): digital repository for cosmic ray data, pooling archives and collecting observations real-time,
- PESI (biodiversity): taxonomically validated standardised nomenclatures for biological and biodiversity management.

(C25) The Europeana Digital Library model

Europeana is a simple but powerful tool for finding resources from all over Europe. Books, journals, films, maps, photos, music etc. will be available for everyone to consult – and to use, copyright permitting. For example, the library will be a rich source of materials for the creative and information industries in developing new products and services, for tourism and for teaching.

(C26) The Open Access Infrastructure for Research in Europe

(*see*: www.openaire.eu) Some 2.5 million research articles are published in 25,000 peer-reviewed journals and conference proceedings worldwide every year. Currently, just 15%-20% of these articles are available in Open Access repositories or Open Access journals. The rest are only accessible through pay per read schemes or by paying for a subscription to the publication. The *Open Access Infrastructure for Research in Europe* (OpenAIRE), launched December, 2nd 2010 by the EC at the University of Ghent in Belgium, could eventually *open up access* to all scientific papers and data produced by researchers funded by the EU's FP7, including scientists receiving grants through the *European Research Council* (ERC), and beyond. OpenAIRE will provide *a network of open repositories providing free online access to knowledge produced by scientists receiving grants from the FP7 and ERC*, especially in the fields of health, energy, environment, parts of Information & Communication Technology and research infrastructures, social sciences, humanities and science in society. Since FP7 started in 2007, some 10,000 projects have been funded. The OpenAIRE project could also lead to new ways of *indexing*, *annotating*, *ordering* and *linking* research results – and new methods to automate all this. This could trigger the development of new services on top of the information infrastructure which OpenAIRE provides. The project is running a helpdesk in 27 European countries, consisting of a network of experts and a portal of tools helping researchers to make their articles available online.

In the following, we recall the results of the survey of data initiatives conducted in [DMTF, 2008], 1.3, pp 16-50.

(C27) Arts and humanities, social sciences

In social sciences, important datasets are often collected not by research teams but by government departments to inform policy makers. These data are usually of high quality, are national samples or census data, and are often under-utilised with high secondary value as a source of information for social science research. Social science archives recognised this potential early and some have been able to negotiate access to these data for the wider research community.

Data Initiatives: CASPAR – *Cultural, Artistic and Scientific Knowledge for Preservation, Access and Retrieval*, CLARIN – *Common Language Resources and Technology Infrastructure*, TextGrid, IASSIST – *International Association for Social Science Information Service and Technology*, DOBES – *Dokumentation Bedrohter Sprachen, HRELP-Hans Raising Endangered Languages Project*, DARIAH – *Digital Research Infrastructure for the Arts and Humanities*, DC-NET – *Digital Cultural Heritage NETwork,* CESSDA – *Council of European Social Science Data Archives, DANS -Data Archiving and Networked Services*, UKDA – *UK Data Archive, University of Essex*, ESDS – *Economic and Social Data Service*, RELU-DSS – *Rural Economy and Land Use Data Support Service*, SHARE – *Survey of Health, Ageing and Retirement in Europe*, COMPARE, DCC – *Digital Curation Centre*, DPC – *Digital Preservation Coalition*, ICPSR – *Inter-University Consortium for Political and Social Research*, NCeSS – *National Centre for e-Social Science*.

(C28) Health Sciences

The European Union's Member States are committed to sharing their best practices and experiences to create a European e-Health Area, thereby improving access to and quality of health care at the same time as stimulating growth in this industrial sector. The *European e-Health Action Plan* plays a fundamental role in the European Union's strategy. Work on this initiative involves a collaborative approach among several parts of the Commission services. The European Institute for Health Records is involved in the promotion of high quality electronic health record systems in the European Union.

Data Initiatives: ELIXIR – *European Life-Science Infrastructure for Biological Information*, EATRIS – *European Advanced Translational Research Infrastructure in Medicine*, ECRIN –

European Clinical Research Infrastructures Network, EU-OPENSCREEN – *European Infrastructure of Open Screening Platforms for Chemical Biology*, INFRAFRONTIER – *European Infrastructure for Phenotyping and Archiving of Model Mammalian Genomes*, BBMRI – *Biobanking and Biomolecular Resources Research Infrastructure*, EBI – *European Bioinformatics Institute*, BIOSAPIENS, EMBRACE – *European Model for Bioinformatics Research and Community Education*, EMMA – *European Mouse Mutant Archive*, EUMODIC – *European Mouse Disease Clinic*, HEALTH-E-CHILD.

(C29) Natural Sciences and Engineering

There are quite different domain-specific practices in the natural sciences. Astronomy and earth and environmental sciences are well-organised communities with well-established best practices for the preservation of data. This is not the case for the High Energy Physics community despite on-going efforts around the upcoming LHC. A nascent initiative around neutron and photon sources in Europe has initiated discussions on *how to preserve the data* originating from the increasing number of these large facilities for multidisciplinary research. This initiative will seek to address the need to *preserve and structure access to the data* generated by a community of more than 25,000 scientists in Europe.

Data Initiatives: BODC – *British Oceanographic Data Centre*, OpenDOAR – *The Directory of Open Access Repositories*, DRIVER-II – *Digital Repository Infrastructure Vision for European Research*, METAFOR – *Common Metadata for Climate Modelling Digital Repositories*, D4SCIENCE – *Distributed Collaboratories Infrastructure on Grid Enabled Technology for Science*, GMES – *Global Monitoring for Environment and Security*, HMA – *Heterogeneous Missions Accessibility*, GEOLAND – *Integrated GMES Project on Land Cover and Vegetation*, MyOcean – *Ocean Monitoring and Forecasting*, INSEA – *Data Integration System for Eutrophication Assessment in Coastal Waters*, MERSEA – *Marine EnviRonment and Security for the European Area*, PARSE.Insight – *Permanent Access to the Records of Science in Europe*, SOSI – *Spatial Observation Services and Infrastructure*, CLIMATE-G, DEGREE – *Dissemination and Exploitation of GRids in Earth sciencE*, SeaDataNet, EMODNET – *European Marine Observation and Data Network*, HIDDRA – *Highly Independent Data Distribution and Retrieval Architecture*, GENESI-DR – *Ground European Network for Earth Science Interoperations*, LIFEWATCH, BIOCASE, GBIF – *Global Biodiversity Information Facility*, EDIT – *European Distributed Institute of Taxonomy*, ENBI – *European Network for Biodiversity Information*, MARBEF – *Marine Biodiversity and Ecosystem Functioning*, MARINE GENOMICS EUROPE, SYNTHESYS – *Synthesis of Systematic Resources*, APA – *Alliance for Permanent Access*, DELOS, PDB – *Protein Data Bank*, EuroVO-AIDA – *European Virtual Observatory-Astronomical Infrastructure for Data Access*, HEP – *High Energy Physics*, ICOS – *Integrated Carbon Observation System.*

3.6 The Remote Instruments access model

(C30) The Instrument Element infrastructure model

(*see:* www.dorii.eu) The *Deployment of Remote Instrumentation Infrastructure* (DORII) project aimed to deploy an e-Infrastructure for remote instrumentation. DORII applications require the integration of scientific instruments with computational and storage resources to facilitate data acquisition, storage and processing. Coordinated and secure access to instruments, data and computational resources is an important requirement for the effective remote usage of these instruments by the application and their users. To fulfil the above requirements the DORII project *utilized the capabilities of Grid infrastructures*.

The main principle is *the use of existing e-Infrastructures in Europe adding the necessary components and services to facilitate remote instrumentation*. The DORII eInfrastructure is mainly based on the EGI/EGEE infrastructure and its middleware of choice gLite (see: EMI gLite). The middleware service dealing with the management of remote instrumentation is the *Instrument Element* (IE) that was built by the DORII project. To deal with the interactivity requirements of the applications the DORII eInfrastructure deployed a selection of services built by the Interactive European Grid Project (int.eu.grid).

The first version of the DORII e-Infrastructure was comprised of *resource centres* (sites) distributed among the partners of the project in several countries such as Germany, Greece, Italy, Poland and Spain. Ten resource centres were already available in the DORII infrastructure in its first version, providing more than 2300 non-dedicated CPUs and several terabytes of storage.

(C31) The ESFRI instrument access model

As noted in the ESFRI Implementation Report [ESFRI, 2009], Research Infrastructures produce and distribute huge volumes of data. The open access to the data generated, as well as the best use of the infrastructures themselves, require the development and continuous improvement of the underlying e-infrastructure, allowing *data acquisition, transfer and analysis*, as well as *data conservation and administration*, to make both data and infrastructure easily accessible to scientists. The attention of ESFRI has however been more focused on *Digital Repositories* [DMTF, 2009] than on an *Instrument Access model*, which is still in a defining phase.

(C32) The Observatory access model

(*see*: www.augeraccess.net) The AUGERACCESS project has been designed to improve the communications link between the Auger Observatory and the CLARA network in Argentina and hence to international networks. Over seven Terabytes of data per year are created at the Observatory site and the desired transmission rate greatly exceeds the available transmission capacity between the Observatory and the city of Mendoza, about 400 km to the north.

The *Virtual Laboratory for Earth and Planetary Materials* VLab, funded by the National Science Foundation and hosted by the Supercomputing Institute for Digital Simulations and Advanced Computation at the University of Minnesota, is an interdisciplinary consortium dedicated to the development and promotion of the theory of planetary materials. Computational determination of geophysically important materials properties at extreme conditions provides today, and maybe for a long time to come, the most accurate information to a) interpret seismic data in the context of likely geophysical processes and b) be used as input for more sophisticated and reliable modelling of planets.

(C33) The Global Monitoring access model

(*See: CYCLOPS project*) CYCLOPS brings together two important Communities: *Global Monitoring for Environment and Security* (GMES) and GRID, focusing on the operative sector and needs of European Civil Protection (CP). The main objectives of CYCLOPS are:

- To disseminate EGI/EGEE results to the CP Community, assessing EGEE infrastructure for CP applications. A variety of activities will focus on dissemination and outreach, training, workshops, possibly in close relation with EGEE events and on promoting a close collaboration between the two communities.
- To provide the EGI/EGEE Community with knowledge and requirements that characterise the CP services. These requirements will also be used to assess the possibility for the development of an advanced grid platform enabling Real Time and near-Real Time services and implementing a security infrastructure very close to the defence systems standards.
- To evaluate the possibility to utilise the present EGEE services for CP applications, developing the research strategies to enhance EGEE platform.
- To develop the research strategies to enhance EGEE platform, especially for Earth sciences resources. CYCLOPS contributed to the EU policy developments establishing liaisons and synergies with other existing projects and initiatives dealing with GMES, GRID and complementary sectors, among them: PREVIEW, Risk EOS, RISK-AWARE, BOSS4GMES, EGEE Networking Activities and Application Support, e-IRG and INSPIRE.

3.7 The Future Internet service-oriented vision

As reported in [WGFI, 2008], beyond national initiatives, multiple frameworks currently exist to address FI issues. Examples are the following.

(C34) The FP FIRE Initiative[6]

The *Future Internet Research and Experimentation* (FIRE) Initiative is creating a *multidisciplinary research environment for investigating and experimentally validating highly innovative and revolutionary ideas for new networking and service paradigms*. FIRE is promoting the concept of experimentally-driven research, combining visionary academic research with the wide-scale testing and experimentation that is required for industry. FIRE works to create a *dynamic, sustainable, large scale European Experimental Facility*, which is constructed by *gradually connecting and federating existing and upcoming testbeds* for Future Internet technologies.

Ultimately, FIRE aims to provide *a framework in which European research on Future Internet can flourish* and establish Europe as a key player in defining Future Internet concepts globally. With a strong network focus, the first wave of FIRE projects was launched in

[6] See: http://cordis.europa.eu/fp7/ict/fire.

summer 2008, with a budget of 40 million Euros and each year sees this increasing along with the scope of the projects.

(C35) Cross-European Technology Platforms (X-ETPs)

The *Strategic Research Agenda* (SRA) of the *cross-European Technology Platforms* (X-ETPs) Group reflects a comprehensive collection of active and upcoming developments in the Future Internet research world, while clearly aiming towards the realisation of the Future Internet vision. The main objective of this SRA is to provide a well structured and consistent publication that reflects and covers a broad set of FI aspects. The document is up-to-date with regards to new developments and evolutions of research topics addressed by the FI community. The aim is to define short, medium and long-term research challenges and identify future important trends before they have started entering hype mainstream, in order to help the European ICT community to gain a leadership position. (see: http://www.future-internet.eu).

(C36) EIT ICT Labs

The mission of EIT ICT Labs, an EIT KIC for the future information and communication society, is to turn Europe into a global leader in ICT innovation. It aims at the radical transformation of Europe towards a knowledge-based society and in turning ICT innovation into quality of life. EIT ICT Labs will equip students, researchers, academics and business people with skills for creativity, risk taking spirit and entrepreneurial capacity, by realigning the existing activities thus leveraging existing regional, national and EU-level funding instruments. EIT ICT Labs will build upon five nodes and their co-location sites, namely Berlin, Eindhoven, Helsinki, Paris, and Stockholm – and turn these already excellent regional clusters into world-class innovation hotspots. The nodes represent five leading European countries in ICT including global companies, leading research centres, and top universities. The cohesion across these nodes will be achieved by strong, CEO-type, management with clear IPR policies for open innovation.

Thematic areas	B	E	H	P	S
• **Smart Spaces** (Juha-Pekka Soininen - VTT) including service-centered home	P	P	D	P	P
• **Smart Energy Systems** (Udo Bub - DT) smart energy management, Green ICT	D	P	P	P	P
• **Health & Well-Being** (Stan Smits) including ambient assisted living, digital medicine	P	D	P	P	P
• **Digital Cities of the Future** (Khaldoun Al Agha) towards intelligent and sustainable digital cities	P	P	P	D	P
• **Future Media and Content Delivery** (Carl Gustaf Jansson, KTH) entertainment, education, accessing media	P	P	P	P	D
• **Intelligent Transportation Systems** (Tanja Kessel, EICT) novel forms of safer & sustainable traffic and transportation systems	D	P	P	P	P

Fig. 10 - Thematic areas for the EIC KIC.

(C37) CELTIC - EUREKA

Celtic is a EUREKA cluster and the only European R&D programme in ICT fully dedicated to *end-to-end telecommunication solutions*. Celtic initiates and runs privately and public funded ICT / telecommunications R&D projects. It is supported by most of the major European players in communication technologies. Celtic projects are focusing on telecoms networks, applications, and services looking at a complete end-to-end system approach. The size of the Celtic budget is in the range of 1 billion euro. Celtic projects are open to any kind of company and organization from all EUREKA countries.

3.8 Joint Efforts

(C38) E-IRG – European Reflection Group

Web site: http://www.e-irg.eu/

The *e-Infrastructure Reflection Group* was founded to define and recommend best practices for the pan-European electronic infrastructure efforts. It consists of official government delegates from all the EU countries. The e-IRG produces white papers, roadmaps and recommendations, and analyses the future foundations of the European Knowledge Society.

(C39) EEF – European E-infrastructure Forum

Web site: http://www.einfrastructure-forum.eu/

The *European e-Infrastructure Forum* is a forum for the discussion of principles and practices to create synergies for distributed Infrastructures. The goal of the European e-Infrastructure Forum is the achievement of seamless interoperation of leading e-Infrastructures serving the European Research Area. The focus of the forum is the needs of the user communities that require services which can only be achieved by collaborating Infrastructures.

The membership of the forum is limited to representatives of large-scale, multi-national, multi-disciplinary Infrastructures. New members of the forum will be invited to join subject to agreement by a majority of the existing members. The initial membership is drawn from the following Infrastructures: EGEE, EGI, DEISA, PRACE, TERENA, GÉANT.

The forum will operate by consensus and will meet approximately once per quarter. Meetings will be hosted by members on a rotating basis, preferably associated with other major events.

(C40) EIRO – European Intergovernmental Research Organization Forum

Web site: http://www.eiroforum.org/

EIROforum is a collaboration between eight European intergovernmental scientific research organisations that are responsible for infrastructures and laboratories: CERN, EFDA-JET, EMBL, ESA, ESO, ESRF, ILL and European XFEL. It is the mission of EIROforum to combine the resources, facilities and expertise of its member organisations to support European science in reaching its full potential.

Final Conclusion

The field of ICT Research Infrastructures is ample and diversified, with widely varying collaboration models. There are in some cases functional *similarities* (for instance between network & DCIs), but there are also significant *differences* (for instance between the network integration model and the multi-facilities MNT model of collaboration).

Also, the *maturity* of collaboration models (integration & coordination, funding, users, industry involvement) is not comparable, as in some domains there are already production-level infrastructures operative and well-defined governance (e.g., networks, DCIs), whilst in some other domains the picture is more fragmented (e.g., data infrastructures) or even still not defined (Future Internet).

A high-level list of seven relevant ICT Research Infrastructure models has been surveyed in this document, outlining for each domain the relevant projects, the governance models and the experienced challenges. A finer granularity inventory of per-domain PA/NC-RIs collaboration models has also been attempted, listing currently forty use cases.

Identified models and use cases will serve as a reference basis for the subsequent benchmark analysis which will be the next WP3 task.

Abbreviations and acronyms

RISGE-RG – Remote Instrumentation Services in Grid Environment Research Group (OGF)

ARI-WG – Access to Remote Instrumentation in a distributed environment Working Group (OGF)

ANNA – European Integrated Activity of Excellence and Networking for Nano and Micro-Electronics Analysis

ASCR – Academy of Sciences of the Czech Republic

BoD – Board of Directors

CERN – Conseil Européen pour la Recherche Nucléaire

CESNET – Czech academic network operator

CSEM – Centre Suisse d'Electronique et de Microtechnique

DEISA – Distributed European Infrastructure for Supercomputing Applications

DMTF – Distributed Management Task Force

EC – European Commission

ECDL – European Conference on Digital Libraries

ECRI – European Conferences on Research Infrastructures

EDGI – European Desktop Grid Initiative (DCI middleware provider / project)

EGEE – Enabling Grids for E-sciencE

EGI – European Grid Infrastructure

EGI_DS – EGI Design Study, project for the conceptual setup and operation of EGI

EEF – European E-Infrastructure Forum

EFII – European Future Internet Initiative

EIB – European Investment Bank

EIC – European Institute for innovation and Technology

EIRO – European Intergovernmental Research Organization (e.g., CERN, EMBL, ESA, …)

e-IRG – e-Infrastructure Reflection Group

EMI – European Middleware Initiative (DCI middleware provider / project)

ERA – European Research Area

ERC – European Research Council

ERIC – European Research Infrastructure Consortium

ESFRI – European Strategy Forum on Research Infrastructures

EU – European Union

EURAB – European Research Advisory Board

IaaS – Infrastructure as a Service

IE – Instrument Element

IPR – Intellectual Property Rights

IGE – Initiative for Globus in Europe (DCI middleware provider / project)

ISC – International Supercomputing Conference; European equivalent to the US based SC0x conference. Held annually in Germany.

FEDERICA – Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures

FIRE – Future Internet Research & Experimentation

FP6, FP7 – Framework Programme 6, 7

GÉANT – Gigabit European Academic Network

HPC – High Performance Computing. Computing at a high performance level at any given time; often used synonym with Supercomputing.

HET – High Performance Computing in Europe Taskforce. Taskforce by representatives from European HPC community to shape the European HPC Research Infrastructure

ICC – Infrastructural Competence Centre

ICT – Information and Communication Technology

IE – Instrument Element

IMCS – Institute of Mathematics and Computer Science, University of Latvia

INFN – Istituto Nazionale di Fisica Nucleare

ISTAG – IST Advisory Group

KM3NeT – Kilometre Cube Neutrino Telescope

KIC – Knowledge and Innovation Community (see EIC)

LHC – Large Hadron Collider

MNT – Micro and Nano Technologies in this deliverable includes also photonics, organic electronics, etc. and similar activities in the hardware development

MNT-Europe – Staircase Towards European MNT Infrastructure Integration

MS – Member States

NGI – National Grid Initiative

NREN – National Research and Education Network

OLWG – Open Living Working Group

OSIRIS – towards an Open and Sustainable ICT Research Infrastructure Strategy

PPP – Private Public Partnership

PPC – Principal Partners Committee (for PRACE), also called hosting members

PRACE – Partnership for Advanced Computing in Europe

PRINS – Pan-European Research Infrastructure for Nano-Structures

RI – Research Infrastructure

RDI – Research Data Infrastructure

RO – Research Organization

RPF – Regional Partner Facilities

SDO – Standards Development Organisation

SME – Small and Medium sized Enterprise

SKA – Square Kilometre Array

Tier-0 (for EGI/LCG) – Denotes the first level of Grid systems (located at Cern). NGIs host the Tier-1 and Tier-2 systems.

Tier-0 (for HPC) – Denotes the apex of a conceptual pyramid of HPC systems. In this context the Supercomputing Research Infrastructure would host the Tier-0 systems. National or topical HPC centres would constitute Tier-1.

UNICORE – Uniform Interface to Computing Resources. Grid software for seamless access to distributed resources.

VRC – Virtual Research Community

VO – Virtual Organization

References:

[1] European Grid Initiative (EGI) Design Study Knowledge Base, http://knowledge.eu-egi.eu/knowledge

[2] European Middleware Initiative (EMI) Project wiki, https://twiki.cern.ch/twiki/bin/view/EMI

[3] TERENA wiki, https://wiki.terena.org

[4] RISGE-RG: http://www.ogf.org/gf/group_info/charter.php?review&group=risge-rg

[5] ARI-WG: http://www.ggf.org/gf/group_info/view.php?group=ari-wg

[6] EEF: http://www.einfrastructure-forum.eu/

[7] Future Internet Portal: www.future-internet.eu

[8] ESFRI, "European Roadmap for Research Infrastructures – Implementation Report 2009", 2009;

[9] European Portal On Research Infrastructures' Services (online database): http://www.riportal.eu/public/index.cfm?fuseaction=ri.search

[10] DMTF, e-IRG Report on Data Management, 2009

[11] Working Group on Future Internet (WGFI), "The Future of Internet, Report from the National ICT Research Directors WGFI", European Commission Information Society & Media, November 2008

[12] European Commission European Science Foundation, "Trends in European Research Infrastructures Analysis of data from the 2006/07 survey", European Commission, July 2007

[13] e-SciDR – Towards an e-Infrastructure for e-Science digital repositories - Final Report, 2008 for DG Information Society and Media Unit F – GÉANT and e-Infrastructure

[14] Council conclusions of 20 November 2008 on the European digital library EUROPEANA (2008/C 319/07)

[15] COM(2008) 467, final 2008/0148 (CNS), Proposal for a COUNCIL REGULATION on the Community legal framework for a European Research Infrastructure (ERI) {SEC(2008) 2278} {SEC(2008) 2279} (presented by the Commission)

[16] High level Expert Group on Scientific Data (HlEGSD) - Riding the wave - How Europe can gain from the rising tide of scientific data - Final report of the High Level Expert Group on Scientific Data – European Union, October 2010

[17] e-IRG, "Blue Paper" 2010

[18] EGI_DS, "EGI Blueprint", EGI_DS D5.3, 2008

[19] EGI.eu, Template for MoU between EGI.eu and VRCs, September 2010

[20] European Parliament, Directorate General For Internal Policies, "Evaluation of the European Research Area (ERA): Governance Aspects Study", IP/A/ITRE/ST/2008-13 PE 429.977, EN, November 2009

[21] e-SciDR Project Team, "Towards a European e-Infrastructure for e-Science Digital Repositories Final Report", DG Information Society and Media, Unit F – GÉANT and e-Infrastructure 2008

[22] PRACE Collaboration: Eickermann, T. (FZJ), "PRACE Final Report", D1.5, 2010;

[23] PRACE Collaboration: Berberich, F. (GCS / JSC), Griffiths, E (BSC), "Report on options for a legal entity", D2.1.1, 2008;

[24] PRACE Collaboration: Griffiths, E., Subirada, F. (BSC), Lichnewsky, A. (GENCI), "Report on analysis of adequate governance structure", D2.2.1, 2008;

[25] PRACE Collaboration: Dias, A.B. (NCF), Douglas, S. (EPSRC), "Final Report on the Peer Review Process", PRACE, D2.4.2, 2009;

[26] PRACE Collaboration: Laaksonen, L., Järvinen, J., Koski, K., Rauhansalo, K. (CSC), Polak, M. (GUP), Kunszt, P. (CSCS), Griffiths, E. (BSC), "Analysis of HPC Ecosystem", PRACE, D2.5.1, 2008;

[27] ESFRI: Rizzuto, C., "Research Infrastructures and the Europe 2020 Strategy", Science Business, 2010;

[28] Hey, T, Tansley, S. and Tolle, K, (Eds.) "The Fourth Paradigm: Data-Intensive Scientific Discovery." Microsoft Research. Redmond, Wash: 2009.;

[29] EuroBioNano Collaboration: Weltring, Klaus-Michael (Ed), EuroBioNano Final Report "Concept for a European Infrastructure in Nanobiotechnology", January 2010.