



SEVENTH FRAMEWORK
PROGRAMME

Research Infrastructures

Deliverable 2.1

Sustainability plan from local sites regarding Grid computational infrastructure



Grant Agreement number:	RI-261572
Project acronym:	WeNMR
Project title:	WeNMR, A worldwide e-Infrastructure for NMR and structural biology
Deliverable type and status	Report– Public
Due Date	May 1 st , 2011 (M6)
Delivery Date	May 13 th , 2011
Project coordinator name, title:	Prof. Alexandre M.J.J. Bonvin
Organization:	Utrecht University, The Netherlands
Tel:	+31 30 2533859
Fax:	+31 30 2537623
E-mail:	a.m.j.j.bonvin@uu.nl
Project website address:	www.wenmr.eu

Table of Contents

Table of Contents	2
Summary of Deliverable	3
General project sustainability.....	4
<i>Stakeholders.....</i>	4
<i>Informing, Attracting and involving users.....</i>	5
<i>The WeNMR user community</i>	6
<i>Contact with industry and other projects</i>	6
<i>Grid maintenance, national NGIs and abroad grids</i>	7
Sustainability plans from local Grid sites	9
<i>Sustainability plan – Partner 1 – BCBR.....</i>	9
<i>Sustainability plan – Partner 2 – BMRZ.....</i>	11
<i>Sustainability plan – Partner 3 – CIRMMMP</i>	12
<i>Sustainability plan – Partner 4 – INFN</i>	13
Conclusion	15

Summary of Deliverable

WeNMR will address the long-term sustainability of its services and will seek to establish sustainability measures, both in terms of continuous feedback and quality control as well as in financial form, for future operation and lasting success. This is of great importance to guarantee that users will be able to benefit of the computational infrastructure services and the opportunities offered also beyond the lifetime of the present project.

In the long term we anticipate that the WeNMR project will be further developed in close collaboration with the Associate Centre for Computational Biology currently being defined in the ESFRI INSTRUCT project. Support from EGI will be pursued via interaction with the NGIs. The worldwide extension of the WeNMR infrastructure also contributes to sustainability by increasing the computational resources available to the project. National funding bodies will be involved to develop a support mechanism in which users applying for computational resources can be allocated time on the enmr.eu VO.

The provision of virtual machines for the various Grid services will also simplify the system administration and ensure a smooth continuation of the project. At present, many of the services are running on virtual machines and local site administrators maintain, upgrade and update those machines when middleware updates are released. The provision of up-to-date virtual machines will thus simplify the task of local operators while ensuring uniformity in middleware.

Finally, a model for use of the infrastructure by industry will be developed and commercial training and consultancy activities will be set up by the SME partner that will allow it to provide continuing support of the WeNMR gateway.

General project sustainability

Stakeholders

Five major stakeholder types have been identified with vested interested in WeNMR who will be targeted and are likely to contribute:

- The core partners in the proposal (Partner 1-4) themselves will maintain and update their investment both in manpower (e.g. the local system administrator dealing with hardware and software maintenance) and hardware. Both are standards aspects of a large research facility.
- User groups using WeNMR. Users will be encouraged to lobby with their National Grid Initiatives and funding agencies to convince them to support WeNMR through a fee-based system. The business model of the US SBGrid will be studied, which is working on a fee/research group basis. A model is foreseen in which non-profit users within WeNMR could involve providing VIP support for a fee, on top of free access to the resources.
- National and European grid initiatives. Service Level Agreements (Letters of Support) with several National, European and worldwide grid initiatives and projects have been obtained during the preparation phase of this proposal to ensure computational capacity and operational support.
- Software developers: commercial and academic The aim at developing promotion and licensing agreements with software developers and providers, mainly for the purpose charging for industrial use of the resources since most software is freely available for non-profit users.
- SME and Pharma industries. Setup an industrial user panel to define requirements and help us develop a business model for the provision of services including training to industrial partners. A list of potential candidates has already been compiled.

All stakeholders groups can contribute to the sustainability of WeNMR, thereby creating a synergy of contributions towards the same common result. We expect direct contributions from National Grid Initiatives (contribution of CPUs to the e-infrastructure), commercial software developers (common licensing schemes) and from industries (consultancy and service contracts). The other stakeholders will most likely contribute by influencing, to their possible extent, on allocating funds at the level of funding agencies and/or national governments, as well as, for the core project partners, through their maintenance and upgrade of the own computational infrastructure.

Informing, Attracting and involving users

Potential new WeNMR users are attracted and trained by courses and workshops. In the first year of the project a couple of workshops are planned. WeNMR was presented at the EMBO global exchange course in Beijing (April 27 – May 5), with three representatives of the project lecturing in this course. During the same period a one day workshop around the HADDOCK software and its WeNMR webportal was given in Taiwan. This also allowed a formalization of the support to the enmr.eu VO from the Taiwanese Grid (Tier 1 site at the Academia Sinica) and attracted new users. In May 2011, there will be a workshop in Florence for "Computational aspects of the joint use of SAXS and NMR", where the ways of integrating SAXS and NMR in the daily routine of structural biologists are discussed and the computational needs for the broader integration are defined. Emphasis is put on the methodological developments and on advanced computational tools including grid computing and remote experiments. In connection to the Euromar conference in Frankfurt am Main in August 2011, a "NMR applications on the Grid" workshop is planned. The workshop will give an introduction, overview and practical training on the usage of some of the widely used and state of the art biomolecular NMR applications that are available on the grid. The covered software packages are all integrated in WeNMR webportals.

To involve the WeNMR users within the project, the www.wenmr.eu website offers a discussion support forum in which the users can comment on different topics; e.g. the services, webportals and registration. A Help Centre on the website gives the users the opportunity to solve issues directly with the developers and a Wiki and tutorial section on the website give detailed background information and explanation of the software and NMR, SAXS and grid services. WeNMR users can submit grid support requests, feedback as well as tickets via the Global Grid User Support (GGUS) enmr.eu VO support unit by the connected feedback mailinglist (eg. enmr-feedback@dlist.server.uni-frankfurt.de). A Virtual Research Community (VRC) is being built and WeNMR has become officially the 1st VRC recognized by the EGI. The Memorandum of Understanding has been signed in Vilnius during the 2011 EGI User Forum.

WeNMR is also promoted with demonstrations, seminars and lectures at various Bio-NMR institutions, laboratories, workshops and conferences. This way, WeNMR activities are widely promoted and receive feedback. In addition, the exchange of knowledge, new computational strategies and tools with the communities inspires the project staff with new ideas for further innovation. An overview of past and upcoming events where WeNMR is presented are shown on the WeNMR website:

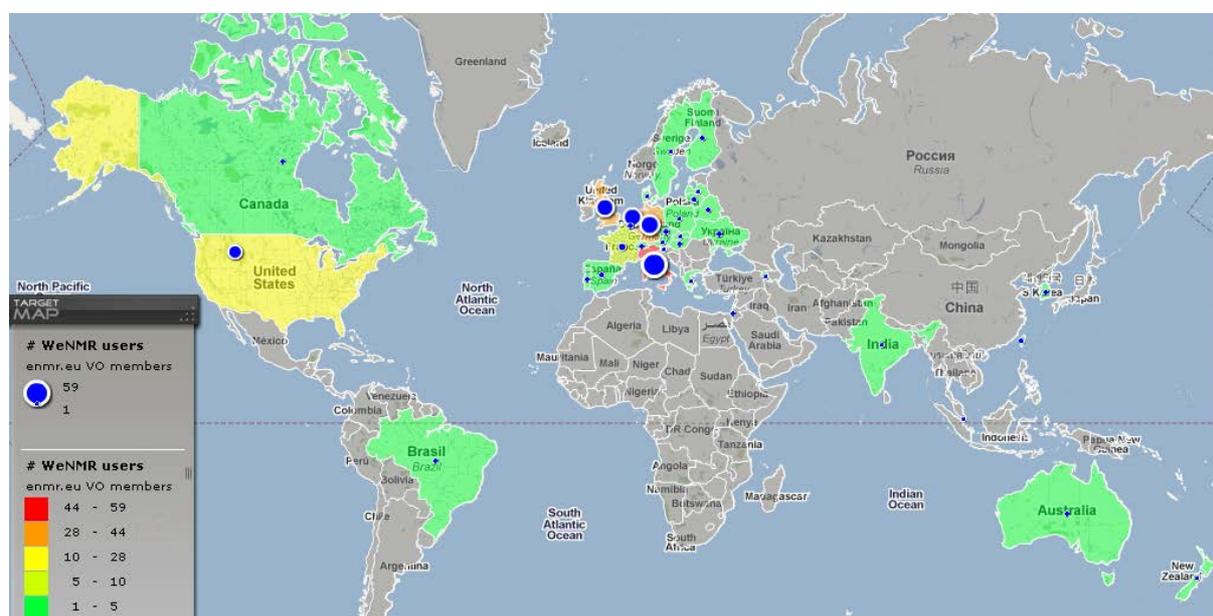
<http://www.wenmr.eu/wenmr/activities/events>

The WeNMR user community

At the end of April 2011 the enmr.eu VO counted 283 registered members, an increase of more than 25% since the start of the WeNMR project after e-NMR. About 15% of the users are from countries outside Europe. The enmr.eu VO is the second largest VO in the Life Sciences area and the first global VO in the Life Sciences.

A distribution per country is shown in the following map (statistics of April 30th, 2011).

<http://www.wenmr.eu/wenmr/wenmr-grid-statistics>



Contact with industry and other projects

WeNMR will very likely participate in the PSDI (Protein Structure Determination in Industry) meeting taking place in Gothenburg by a round table discussion with industrial scientists.

WeNMR has also very close contacts and collaborations with FP7 funded infrastructure projects such as East-NMR and Bio-NMR. Their user communities and researchers in Europe form a coherent research community prepared to tackle scientific and biomedical challenges of increasing complexity at the forefront of research worldwide.

Finally, collaboration with several other FP7 projects has been initiated (e.g. OpenAIRE, e-ScienceTalk CHAIN, Gisela, ...), aimed at a dissemination of results obtained in WeNMR to a broader, global audience and at an extension of the e-Infrastructure supporting the project. Memoranda of Understanding have been signed during the EGI user forum in Vilnius with

EGI-InSPIRE, EMI, e-ScienceTalk and EUMedGrid. Finally, the WeNMR project coordinator, Prof. Bonvin from the BCBR partner is a member of both the User Community Board (UCB) and External Advisory Committee (EAC) of EGI-InSPIRE, providing a direct link to EGI.

Grid maintenance, national NGIs and abroad grids

The grid infrastructures at the core sites have been maintained and kept up-to-date with the latest Middleware software and Operating System upgrades and updates. The sites in Frankfurt am Main (BMRZ) and Florence (CIRMMP) are certified European Grid Infrastructure (EGI) sites and supported by their National Grid Initiatives (NGI) in Germany and Italy. The Utrecht site (BCBR) has candidate certification status and is supported in The Netherlands with the National BigGrid project. The grid services at BCBR, BMRZ and CIRMMP are fully dedicated to the enmr.eu VO and the services hosted by the INFN partner are shared with the Italian Grid Infrastructure (IGI). The services provided span all areas from job management, information system, data management and storage, monitoring and accounting. The WeNMR project is furthermore supported by the national grid initiatives of Belgium, France, Portugal, UK / Scotland, South Africa and Latin America via the Gisela project (Brazil, Argentina and Venezuela).

<http://gstat-prod.cern.ch/gstat/summary/VO/enmr.eu/>
<http://gstat-prod.cern.ch/gstat/geo/openlayers#/VO/enmr.eu>
<http://operations-portal.in2p3.fr/vo/rd/vo/enmr.eu>

The additional support from several NGIs, sharing their production resources with enmr.eu VO allowed a quick growth of the computing capacity, which is also very important for the reliability and sustainability of the project. All of the resources available to the VO are monitored, controlled and accounted by the official EGI operation tools and procedures, with the help of the support teams of the corresponding NGIs. Three operation tools were fully regionalized: the helpdesk based on GGUS (the Global Grid User Support web-service interface ticketing system operated by the NGIs), the Nagios-based monitoring and GStat 2.0. Other tools based on a central deployment model will be regionalized in the EGI-InSPIRE program. Furthermore, the Accounting Processor for Event Logs (APEL) service is used within several European NGIs and the Distributed Grid Accounting System (DGAS) with Home Location Registers (HLR), initially adopted in Italy has been extended to the grid sites of other NGIs supporting the enmr.eu VO.

Currently many services and grid elements are being upgraded at all sites from the gLite 3.1 Middleware on a Scientific Linux 4 Operating System to gLite 3.2 on Scientific Linux 5. In addition CREAM-CE services will be provided. All the core sites provide the essential grid elements for WeNMR and additional tools for job submission and security. These include a Computing Element (CE) with several workernodes (WNS), a Storage Element (SE), a Workload Management system (WMS) with Logging and Bookkeeping (LB), a Security

Encryption Service (HYDRA), a Berkeley Database Information Index (BDII) and a User Interface (UI). Some NGIs require supplemental services such as APEL, which are provided as well. Additional tools are provided by the INFN and CIRMMP such as a Data management LCG File Catalog (LFC), WeNMR specific Monitoring services (Gstat, Nagios, WMSMonitor), the enmr.eu VOMS server and the before mentioned DGAS/HLR accounting service.

To ease the use of grid applications, many webportals have been developed by the WeNMR consortium and more (among others for SAXS data analysis) are planned for the coming years. Currently the following NMR services have been ported as grid webportals, covering many computational aspects of Biomolecular NMR in the streamline to get from Data to Structure:

Processing: MDDNMR
Assignment: AutoAssign, MARS
Analysis: TALOS+, AnisoFit
Structure Calculation: Xplor-NIH, CYANA, CS-Rosetta
Molecular Dynamics: AMBER
Modelling: 3D-DART, HADDOCK
Tools: Format Converter

More software applications have been deployed without a webportal. Those are directly available via the standard gLite User Interface (e.g. NMRPipe, PROSA, GARANT, INFIT, MAPPER, CNS, GROMACS, NUSSAMPLER). Tutorials and use-case examples are provided for all the (web-based and UI-based) applications, which explain how the software can be used on the WeNMR grid.

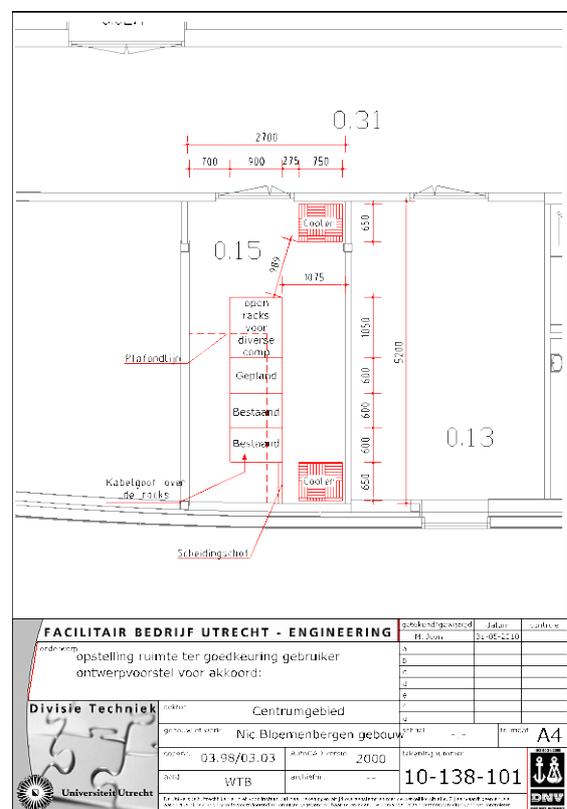
Sustainability plans from local Grid sites

Sustainability plan – Partner 1 – BCBR

The computational infrastructure at the BCBR site at Utrecht University consists of cluster with 23 nodes for production tasks and several other nodes hosting the various grid services under virtual machines. This cluster can be monitored from the following link:

<http://haddock.science.uu.nl/ganglia/?c=HADDOCK%20cluster>

Until recently, the computational infrastructure was located within an experimental NMR hall. This situation was far from ideal because of the noise generated and the space available. This has prevented further expansion of the infrastructure in the past. We have been working together with the campus management of Utrecht University to design a new and dedicated computer room to host our computational infrastructure and allow future growth. This dedicated room was delivered at the end of March 2011. It was designed to host three full clusters and some additional servers (total power of +/- 160A max. (32 kW) and corresponding cooling capacity). This capacity should allow to more than double our hardware infrastructure over the coming years. This represents a major investment from our university in the computational infrastructure provided by WeNMR, estimated to be over 80'000 euro.



With the move of our clusters and other servers to this new computer room, there is now room for expansion. We have already purchased four new nodes to expand the current cluster, each with two XEON QUAD CORE E5620 2,40GHZ processors (an investment of 7050.75 euro from funding distinct from the EU funds. This will bring the total number of production nodes to 26 for a total of 208 cores.

In order to relieve the load on the current cluster and dedicate it for our web portals and Grid computing, we have recently ordered a new cluster, which will be dedicated in first instance to research and development within the research groups of the BCBR site but will also benefit the WeNMR portal development. The cluster will consists of 6 nodes with each

48 cores (AMD Opteron Socket G34 Twelve-Core 6172, 2.1 GHz, Magny Cours), for a total of 288 CPU cores, next to a master node with over 8 TB storage space. This cluster represents an investment of over 66.000 euro, from funding distinct from the EU funds.

The cluster, which will be delivered in May 2011, will be configured to have a local copy of all web servers running in order for it to provide a backup system in case of failure of our Grid-based cluster.

We also have 50.000 euro available from a Dutch National Science Foundation VICI grant to Prof. Bonvin, which is reserved for an update of the computational infrastructure over the period of two years.

With these recent significant investments, both from Utrecht University and from the BCBR partner directly, and with the budget available for future upgrades, our computational infrastructure is up-to-date and its sustainability ensured for the coming years, which should allow a smooth running of our services beyond the duration of the WeNMR project.

Finally it should be mentioned that that the important role of WeNMR in structural biology is well recognized at the national level by the Dutch BigGrid project (<http://www.biggrid.nl/en/big-grid-infrastructure/communities>), which provides a substantial share (>80%) of computational resources to the enmr.eu VO and also by the recent e-BioGrid project (<http://www.ebiogrid.nl>), which might provide support in software development.

Sustainability plan – Partner 2 – BMRZ

The various grid services at the BMRZ-FRANKFURT site are maintained on a cluster dedicated to the Center for Biomolecular Magnetic Resonance (BMRZ) and an additional web-server for the webportals. The cluster was purchased with funding from the Goethe University / BMRZ; representing a local investment of about 82.000 euro for the computational e-Infrastructure. The cluster consists of 20 nodes, each with state-of-the-art 2.8 GHz 2x Xeon Quadcore E5462 processors (8 cores per node), 16 GB memory and 250 GB disk space. The gLite Storage Element (SE) has additional 1 TB disk space. The cluster is maintained by local scientists, site administrators, and located in a suitable computer room at the University Campus Riedberg; with sufficient space for expansion when needed. The current CPU power and capacity is presently more than sufficient. Nevertheless, to keep up-to-date the coming years and for sustainability of the e-Infrastructure during and after the WeNMR project, local investments (of around 50.000 euro estimated in 2012) in hardware additions and upgrades of the cluster are expected and available.

The BMRZ-FRANKFURT is a certified grid production site and associated to the German Regional Operation Centre (ROC): the German National Grid Initiative (NGI-DE). The following gLite Middleware services have been deployed and are maintained: a Computing Element (CE) with workernodes (WNS), a Storage Element (SE), a User Interface (UI), a Berkeley Database Information Index (BDII), an Accounting Processor for Event Logs (APEL), an Encryption service (Hydra) and a Workload Management System (WMS) with Logging and Bookkeeping (LB). The site services are permanently being monitored by Nagios and have an outstanding availability and reliability ranking; easily fulfilling the performance requirements with minimum target values of 70% availability and 75% reliability. The latest site reports for the five months December 2010 to April 2011 show 100%, 99%, 96%, 99%, 99% availability and 100%, 99%, 100%, 99%, 99% reliability per month.

The site administrators of BMRZ are involved in NGI-DE Operation Meetings in which the latest grid topics are being discussed. The NGI-DE and associated grid sites in Germany have been requested to support the enmr.eu VO, which initiated an internal discussion. The GridKa will very likely support and enable the VO. In future such requests can possibly be made within the NGI-DE using a questionnaire. Efforts are ongoing to involve more NGI-DE sites.

Sustainability plan – Partner 3 – CIRMMMP

The computational infrastructure dedicated to WeNMR at the CIRMMMP site consists of a 10-node cluster, providing 32 CPU's as computational power for the Worker nodes and 24 CPU's to host the various Grid services. The cost of the cluster was about 50.000 € and was entirely funded by internal financial resources. In 2010, two additional servers were purchased to host the web services; at the same time the RAM of all the nodes in the cluster was expanded, for a total investment of nearly 20.000 €. In the same year, the disk space for internal lab users was separated by that dedicated to the WeNMR activities, by acquiring approx. 3.0 Tb of additional space.

Two 48000 BTU air conditioners have been installed to cool the room hosting the cluster. The air conditioning system is organized in such a way as to ensure that each conditioner is operational each other week and can automatically replace the other one in case of failure. In addition, during extreme heat periods (such as August) both conditioners will automatically switch on.

CIRMMMP is a certified grid production site and is involved in the Italian Grid Initiative (http://www.italiangrid.org/grid_operations/infrastructure/sites_and_resources). The per-month availability and reliability rates since the beginning of the year have always been greater than 90%, with peaks at 100%.

In summary the total investment by CIRMMMP in computer hardware so far has been of about 80.000 €, which does not include the expenses for the newly installed air conditioners. The hardware has been periodically upgraded. The strategy of CIRMMMP to guarantee future upgrades, substitutions, and expansions is mainly focused on exploiting public-private partnerships at the regional and national levels. Among the various initiatives, within the Operational Regional Programme (POR) CIRMMMP is involved in a partnership with Novartis Vaccines and Diagnostics in the field of structural vaccinology. CIRMMMP is also involved in the "Poli di Innovazione" initiative to foster technology transfer to private stakeholders in the fields of ICT and Life Sciences. Finally, collaborations to develop and provide grid-based services of use to biobanks are being sought.

Sustainability plan – Partner 4 – INFN

INFN is the main contributor (85%) in term of resources to the Italian Grid Initiative (IGI). IGI is the Italian NGI (National Grid Initiative) currently formed by 17 Italian Research Institutions and Universities and is part of EGI. With more than 30.000 CPU-cores and 30 PB of storage resources distributed across 58 computing centres all over Italy, IGI operates one of the largest infrastructures in Europe, which represents around 15% of the current EGI production grid accessible through the gLite middleware.

IGI is currently an EU Joint Research Unit (JRU), based on a Memorandum of Understanding signed between the members in December 2007, and formally supported by the Italian Ministry for University and Research (MIUR) and the European Commission. It provides a unique representation at national, European and international levels for all the different components of the Italian Grid Infrastructure: Regional Grids, single Compute and Storage centres, and facilities for end users of the Research Institutions related to the national grid infrastructure.

During 2010 INFN has funded its grid activity with 4.9 MEuro of contribution for the procurement of computing and storage resources at the sites and for the central grid services. At the end of 2010 the Italian MIUR has funded INFN with a 2 MEuro contribution for starting the establishment of IGI as a new legal organization (a.k.a. Consortium), while 5 MEuro are foreseen yearly from 2011 onwards. A two phase strategy is planned. In the phase 1, expected to last approximately one year, IGI will be an internal INFN project (“Progetto Speciale”) with the goal of: to transfer and consolidate inside the project the current technical activities of general interest carried out by the partners; to guarantee as IGI the commitments with the ongoing EU grid projects EMI and EGI-InSPIRE, and with the Virtual Research Communities (including WeNMR); to organize its personnel autonomously but in collaboration with the partners; to prepare the charter and create the Consortium as soon as possible in order to be able to participate to the upcoming national and international funding calls; to develop a sustainability strategy and the administrative instruments for the new Consortium. The phase 2 will start after the creation of the Consortium: IGI will become an autonomous entity managed by the partners (including INFN); its duration will be defined in the charter; it will engage its own personnel with long-term contracts (around 60 full time positions are estimated).

Within the context described above, the IGI special project within INFN will keep running and made available to WeNMR project the grid central services currently serving the enmr.eu VO (WMS, LB, BDII, LFC, VOMS, Hydra, DGAS, HLRMon, GStat, WMSMonitor, Nagios) and hosted at INFN centres of Padua and CNAF (Bologna). Part of these services are shared with other VOs, while some others are fully dedicated to WeNMR, as described in details in the project deliverable D4.1.1 issued also at Month 6. Furthermore, the 5% of INFN overall grid resources (i.e. around 1500 CPU-cores) will be made available to enmr.eu VO, which is currently already enabled on several sites of the Italian grid since the previous e-NMR project.

The IGI special project will take over from INFN its leading role in the middleware development currently carried out in the context of the EMI project and will also ensure the availability of the latest versions of the most relevant grid services included in the future EMI releases.

Conclusion

The current status of the WeNMR grid e-Infrastructure is being kept up-to-date and extended within Europe and worldwide. New software, webportals and developments are incorporated. The user community is growing and will become more involved in the project by the discussion support forums and within the Virtual Research Community (VRC). Contacts have been established and will be further extended with Industrial partners and related projects. WeNMR, as first recognized VRC by the EGI, has become an important partner in e-Science-related activities representing of a growing and vital research community in structural biology and life sciences, both at European and global level. The support of EGI and other FP7 projects should boost our sustainability efforts.