A worldwide e-Infrastructure for NMR and structural biology

EMBO Global Exchange Course, USP Sao Paolo, Jan. 19-26

Alexandre M.J.J. Bonvin
Project coordinator
Bijvoet Center for Biomolecular Research
Faculty of Science, Utrecht University
the Netherlands
a.m.j.j.bonvin@uu.nl

The Project

A Worldwide e-Infrastructure for NMR and structural biology

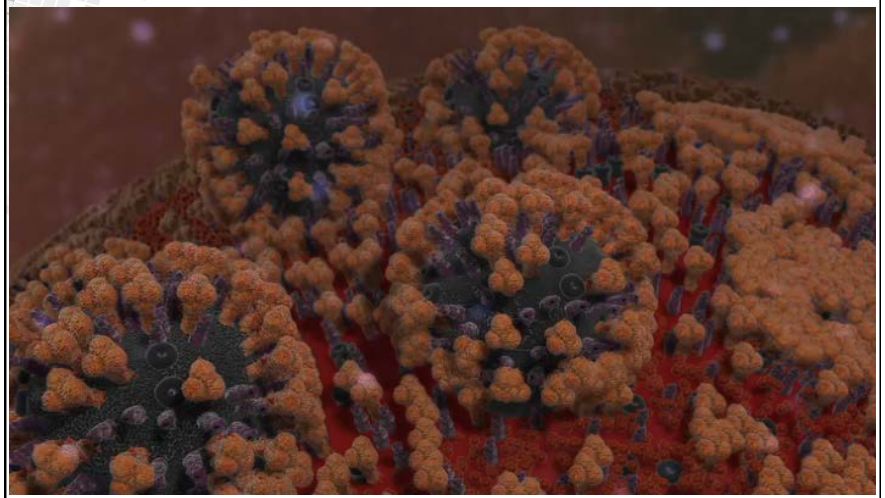
Project Coordinator:
Prof. Alexandre M.J.J. Bonvin, Utrecht University, NL

Contract n°: RI-261572
Project type: CP-CSA
Duration: 36 months (Oct.2013)
Total budget: 2'434'000 €
EC Funding: 2'150'000 €

The team:

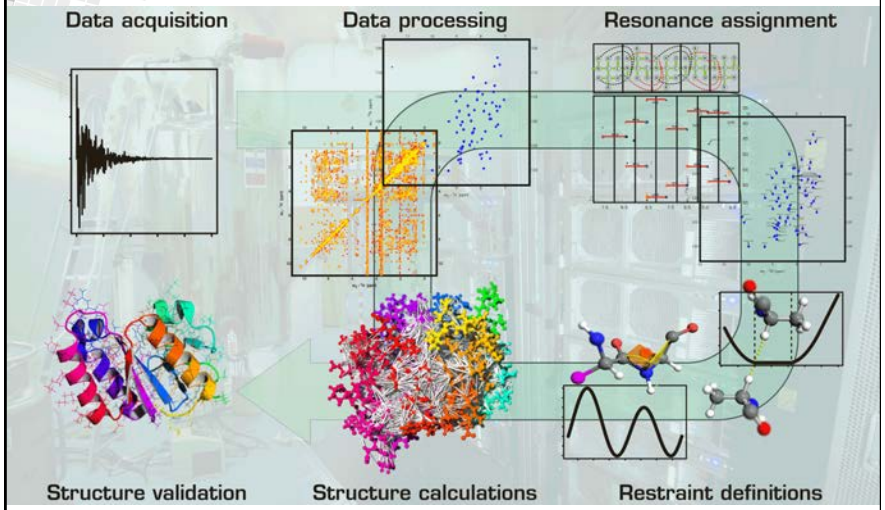
- Utrecht University, Bijvoet Center for Biomolecular Research, NL
- Johann Wolfgang Goethe Universität Frankfurt a.M., Center for Biomolecular Magnetic Resonance DE
- University of Florence, Magnetic Resonance Center, IT
- Istituto Nazionale di Fisica Nucleare , Padova, IT
- Raboud University, Nijmegen, NL
- University of Cambridge UK
- European Molecular Biology Laboratory, Hamburg, DE
- Spronk NMR Consultancy, LT
- Academia Sinica, TW

NMR and structural biology

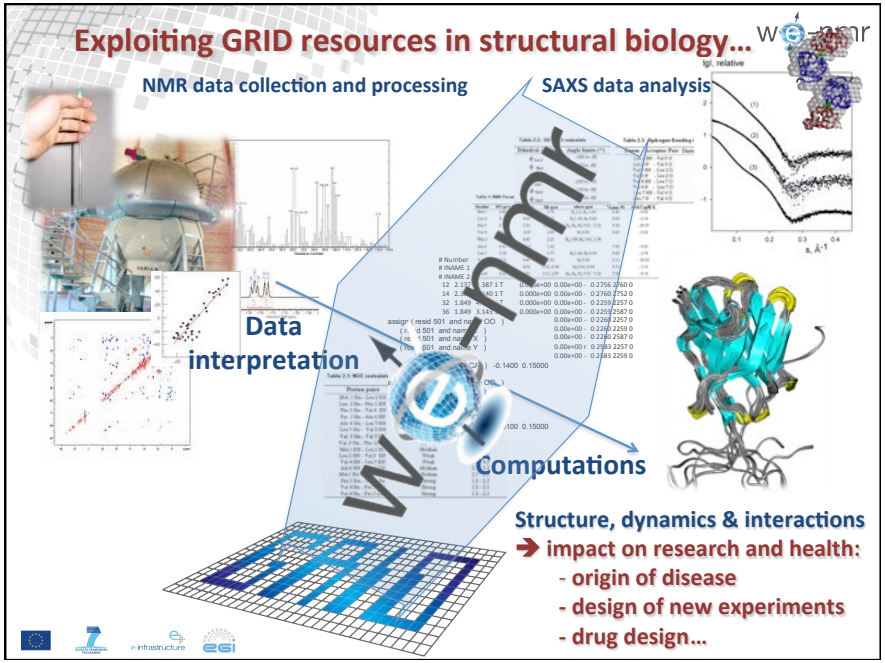


A SpronkNMR production (www.spronknmr.eu)

The NMR structure determination workflow



Exploiting GRID resources in structural biology…



What is a Grid?

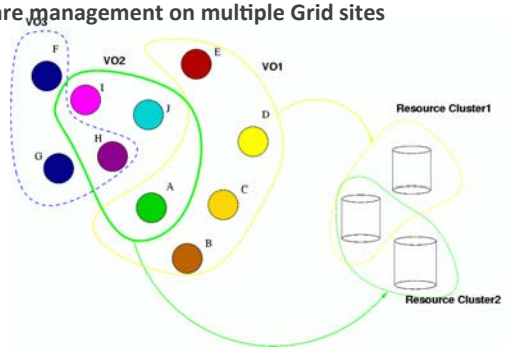
- A distributed and shared computational infrastructure
  - WEB offers a large collection of data
  - GRID offers a large collection of computational power and space
- In a cooperative system with a dedicated fair share, a single user can exploit ALL available unused resources!
- GRID definition
  - *"A flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources what we refer to as virtual organizations. In such settings, we encounter unique authentication, authorization, resource access, resource discovery, and other challenges"*.

The Anatomy of the Grid. Ian Foster, Carl Kesselman - 2001

Virtual Organizations and the Grid

- Virtual Organizations:
  - groups of individuals or institutions who share the computing resources of a "grid" for a common goal, providing:
    - centralized software management on multiple Grid sites
    - users support

Virtual Organizations accessing different and overlapping sets of resources



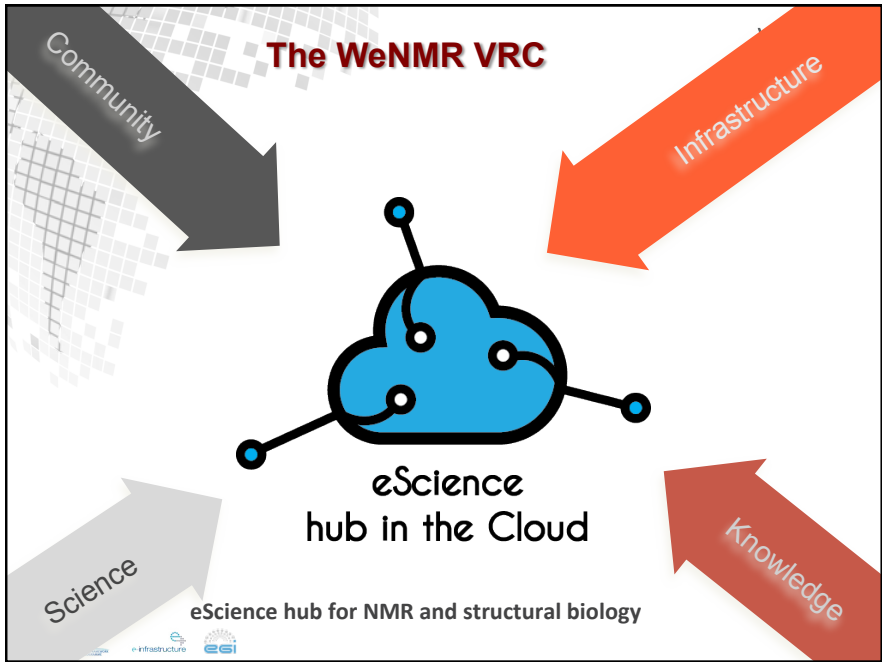
Main objectives

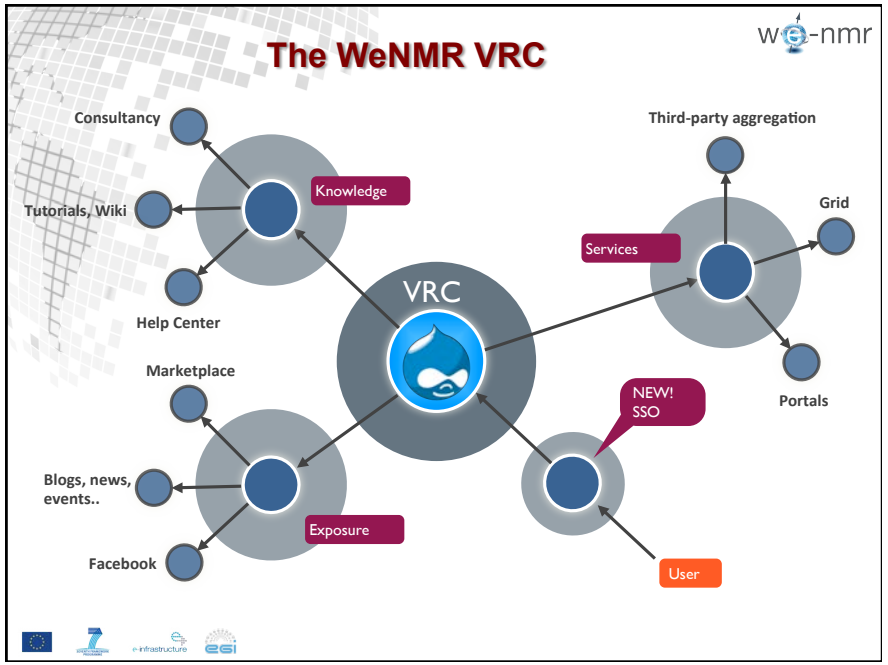
- Operate and further develop a user-friendly **e-Science gateway for the NMR and SAXS communities**
- Establish a **virtual research platform** for (interaction with) the user community
- **Provide support** to software developers, users and other e-Infrastructure projects
- **Foster the adoption and use of e-Infrastructure** in a wide range of flanking disciplines within the life sciences
- **Operate and consolidate the eNMR Grid infrastructure** and to extend it to interoperate with other **worldwide Grid initiatives**
- Develop a model to ensure **sustainability** of the project

The WeNMR philosophy for interacting with the Grid

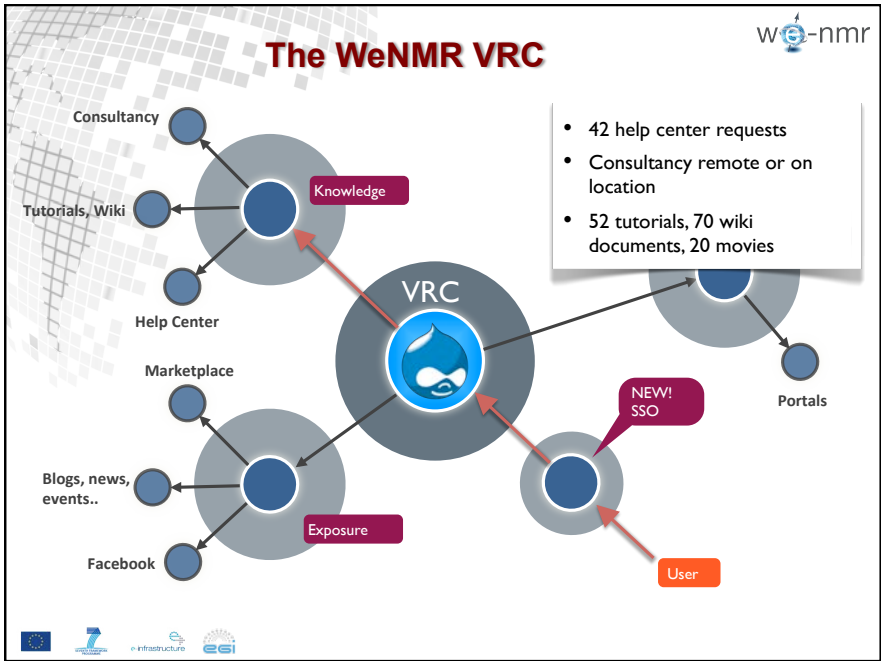
- Our policy is to **shield as much as possible the end user from the Grid** and all middleware related issues and commands
- For this, we chose to develop mainly application web portals providing **"protocolized" access to the Grid**
- To facilitate operation, we **use when possible robot certificates**
- Experienced users can still interact directly with the Grid via UI and we **provide "ready-to-go" customized UI distributions** (MILU) for download

The Virtual Research Community



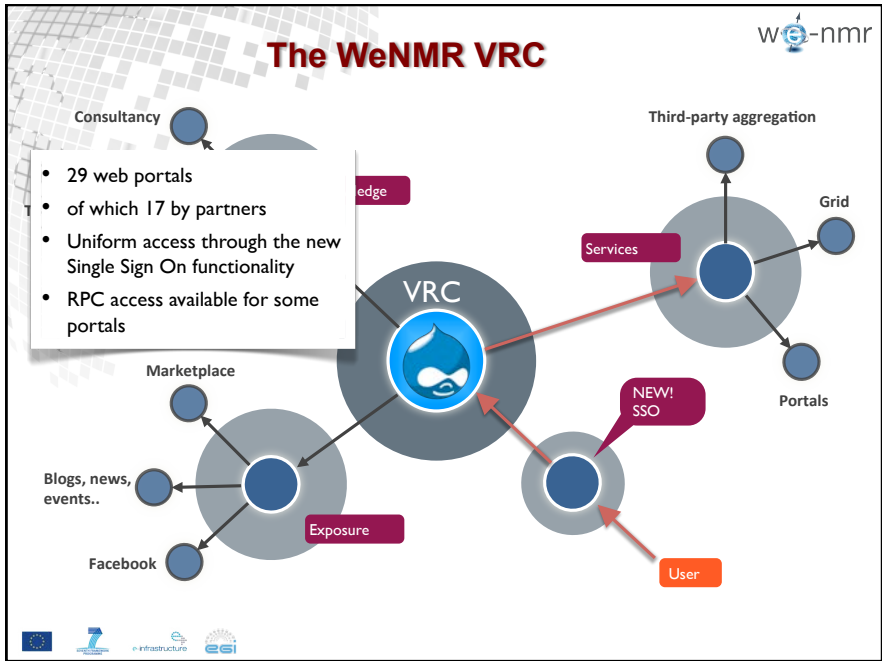


The WeNMR VRC

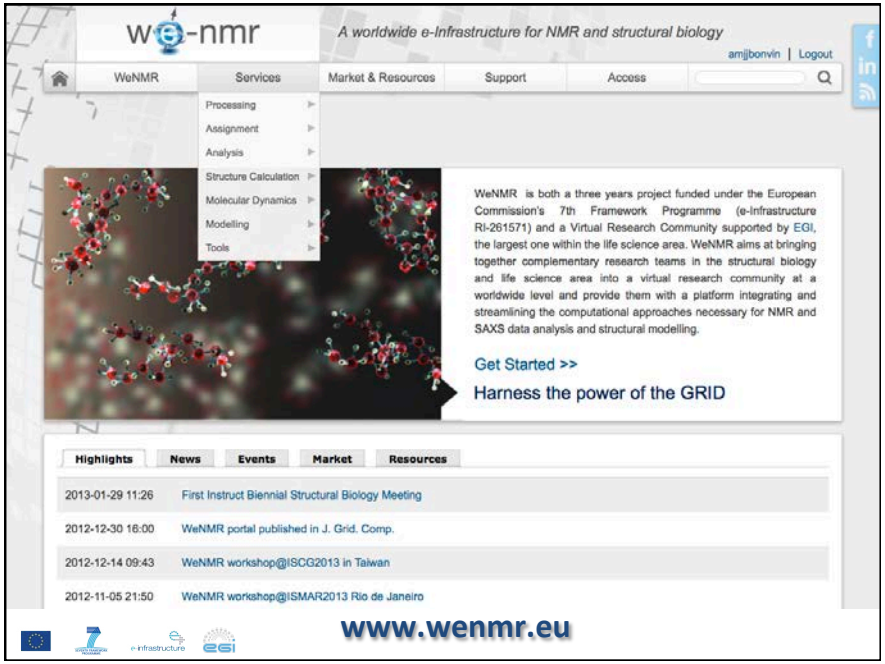


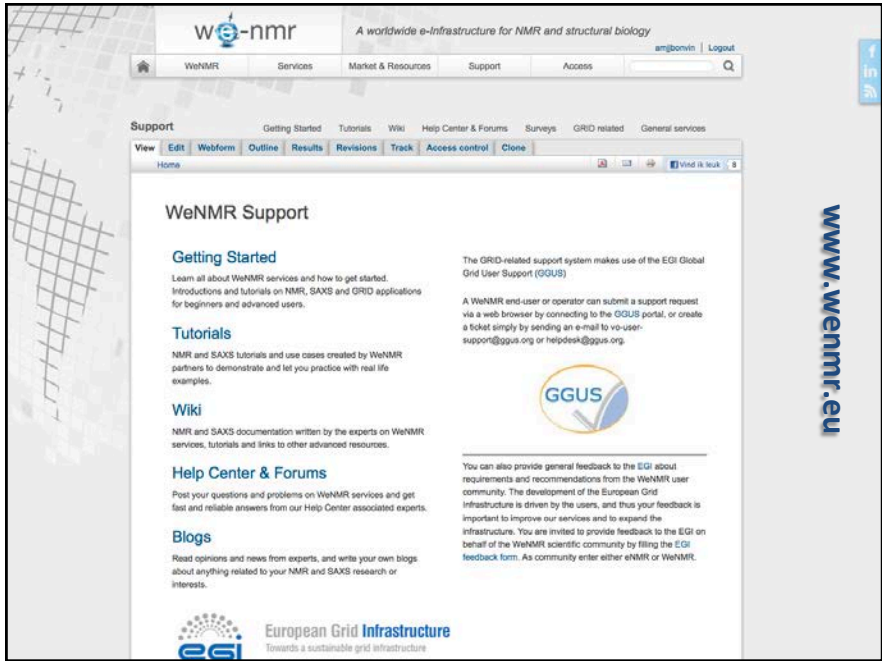
- 42 help center requests
- Consultancy remote or on location
- 52 tutorials, 70 wiki documents, 20 movies

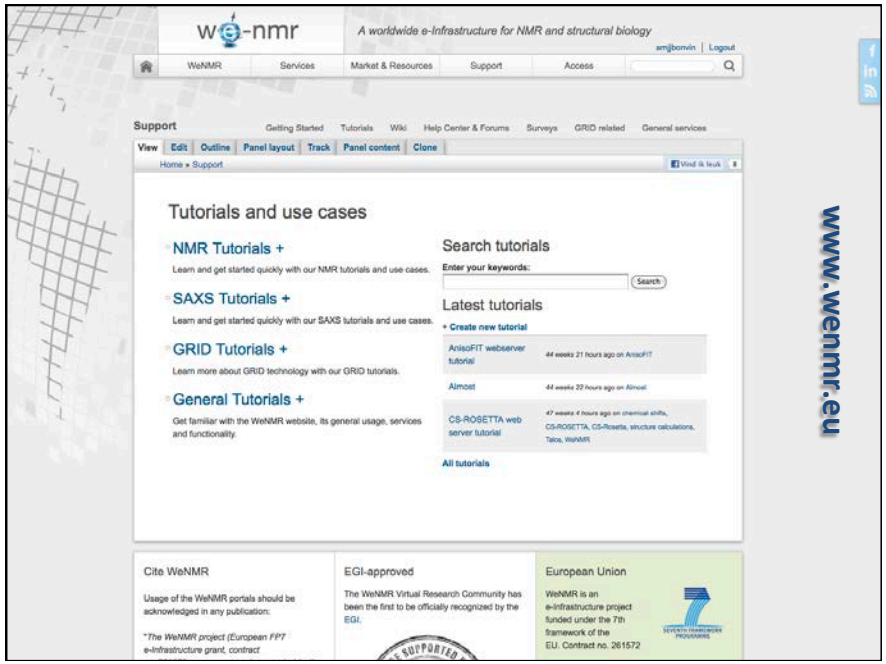
The WeNMR VRC

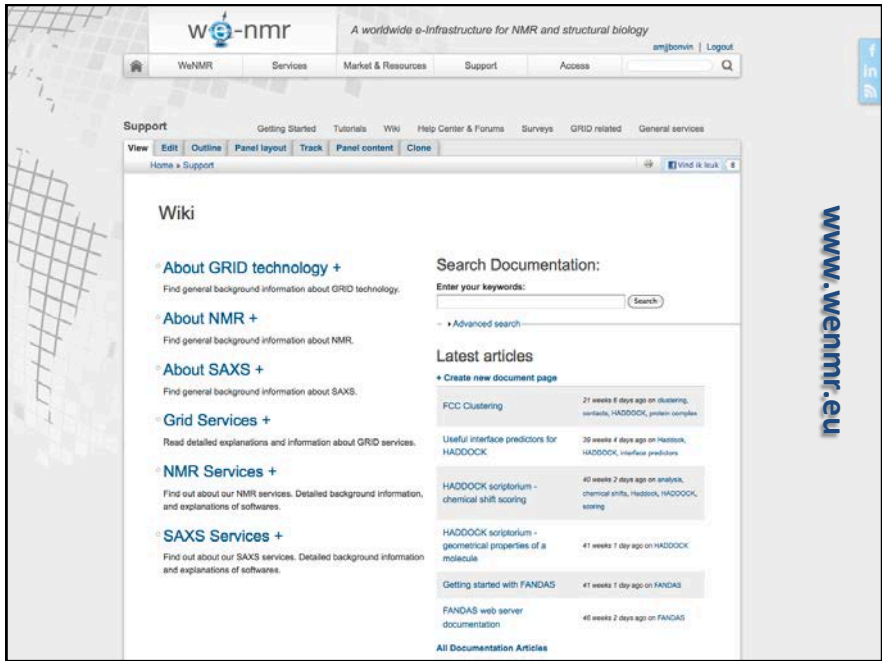


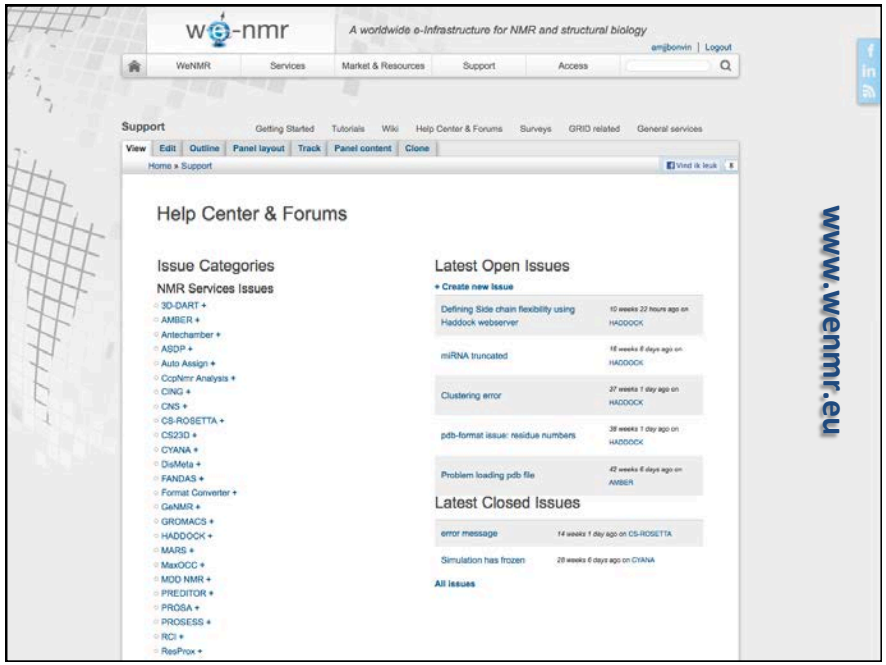
- 29 web portals
- of which 17 by partners
- Uniform access through the new Single Sign On functionality
- RPC access available for some portals

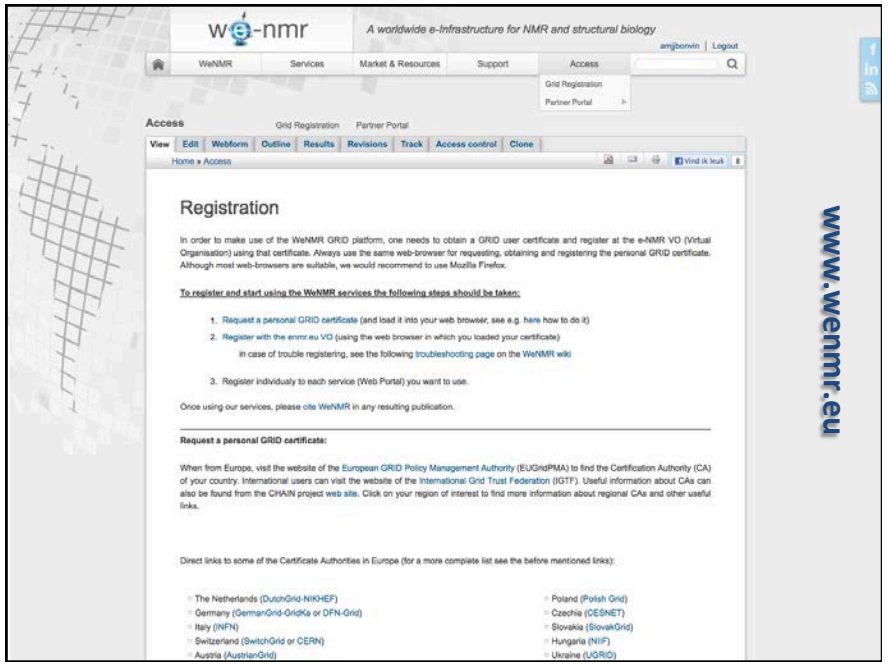


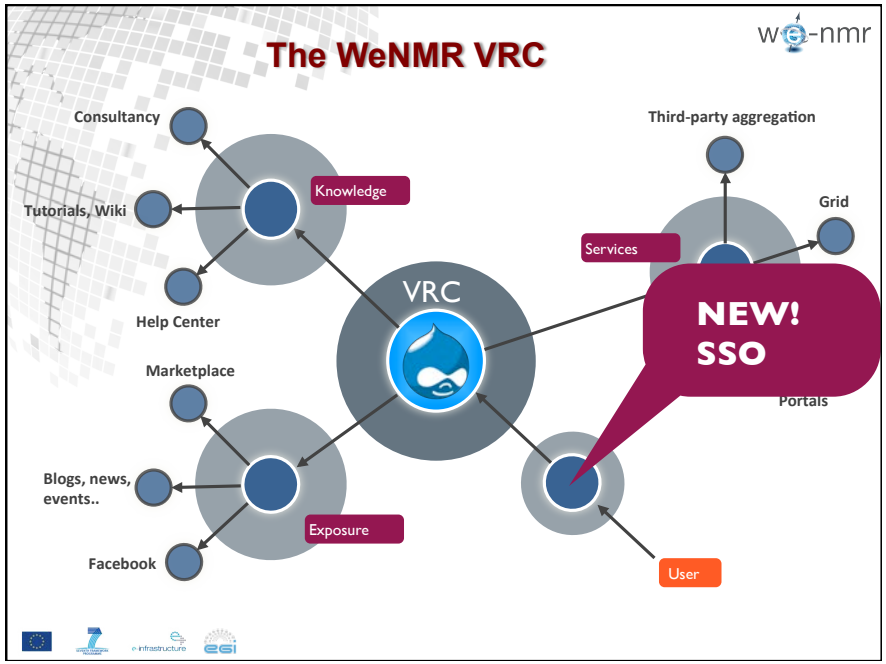


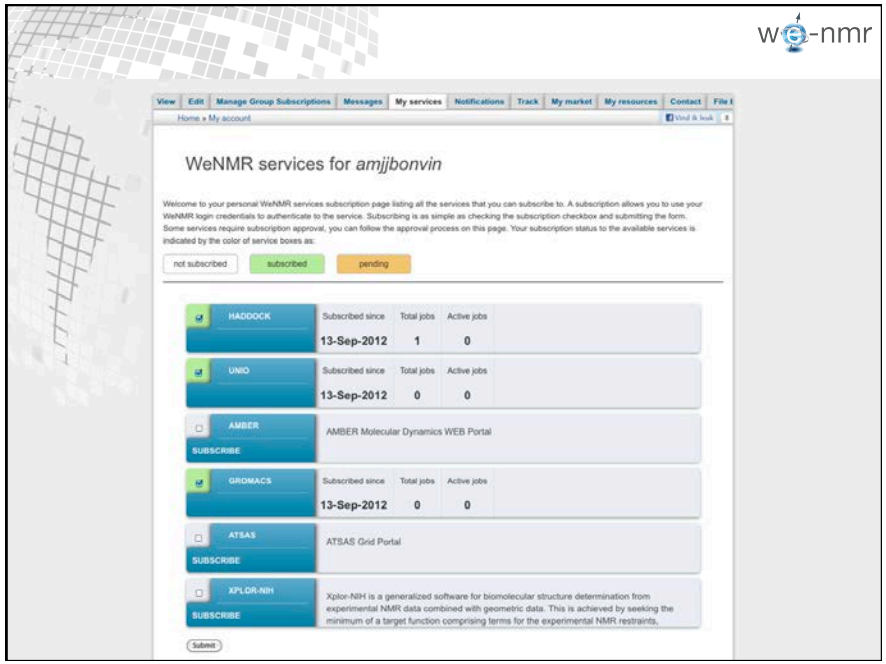

w-nmr
A worldwide e-Infrastructure for NMR and structural biology
WeNMR
Services
Market & Resources
Support
Access
Support
Tutorials and use cases
NMR Tutorials +
SAXS Tutorials +
GRID Tutorials +
General Tutorials +
Search tutorials
Latest tutorials
Cite WeNMR
EGI-approved
European Union
www.wenmr.eu

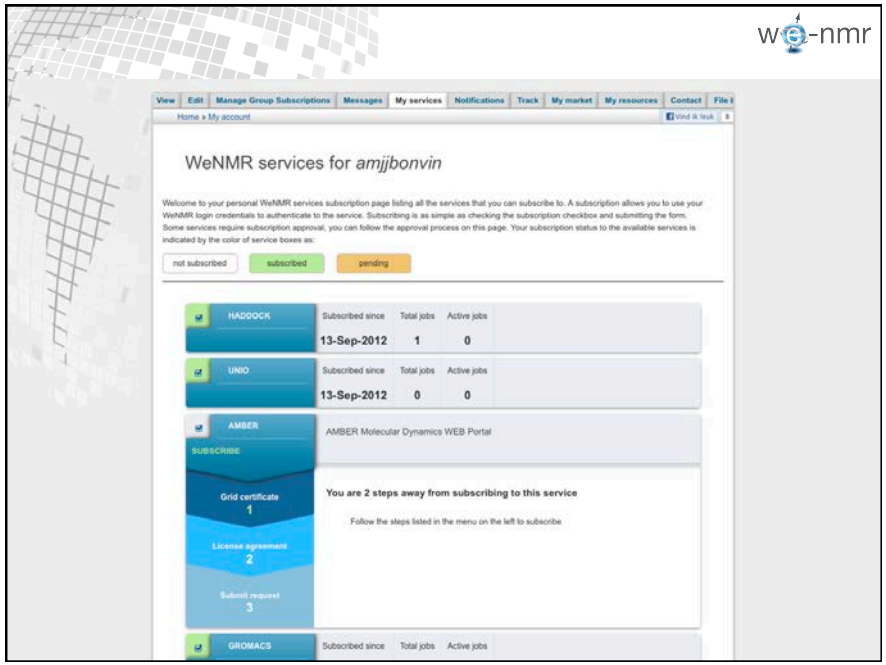

w-nmr
A worldwide e-Infrastructure for NMR and structural biology
Support
Wiki
About GRID technology +
About NMR +
About SAXS +
Grid Services +
NMR Services +
SAXS Services +
Search Documentation:
Latest articles
www.wenmr.eu

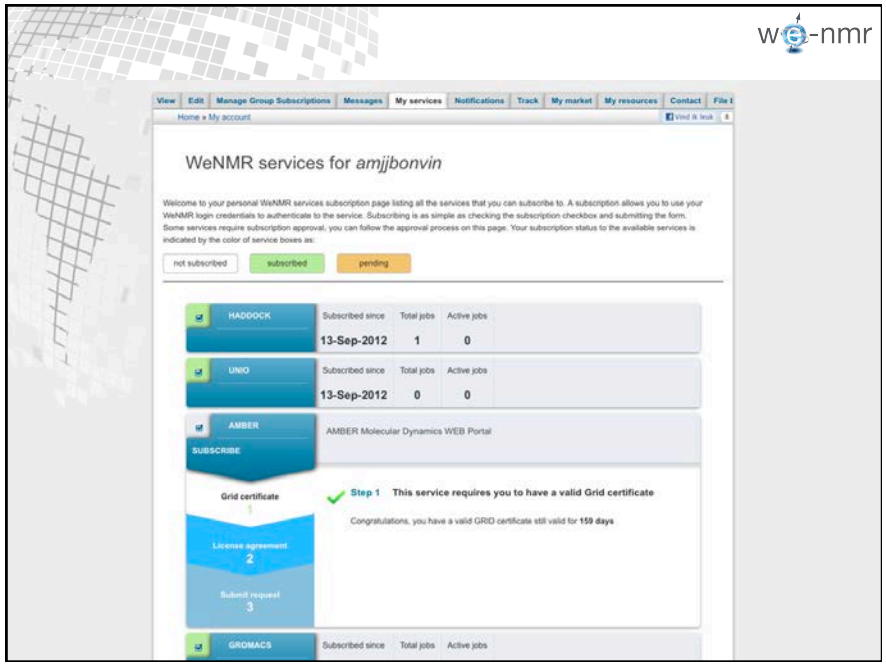

w-nmr
A worldwide e-Infrastructure for NMR and structural biology
Support
Help Center & Forums
Issue Categories
NMR Services Issues
Latest Open Issues
Latest Closed Issues
www.wenmr.eu

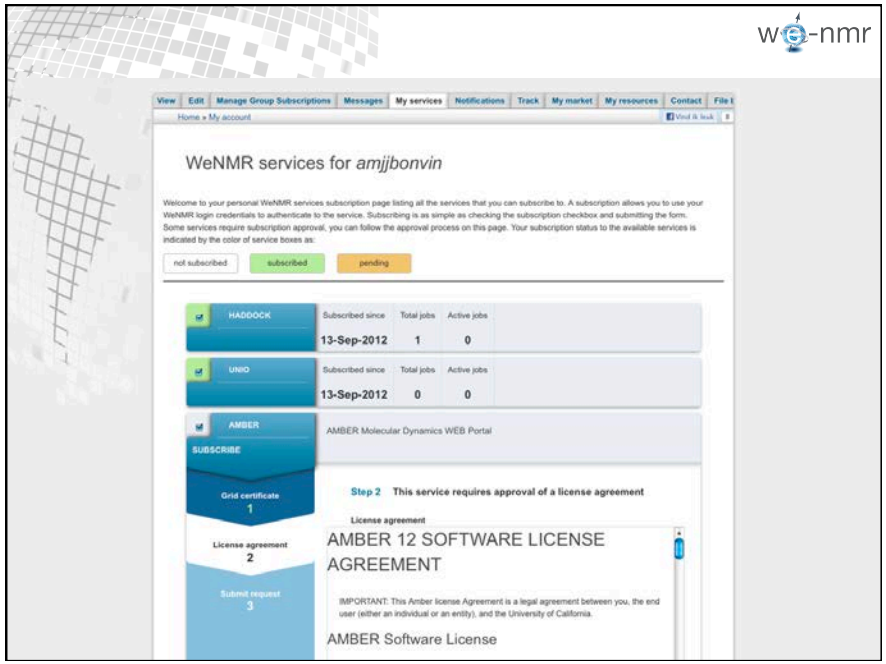

w-nmr
A worldwide e-Infrastructure for NMR and structural biology
Access
Registration
www.wenmr.eu

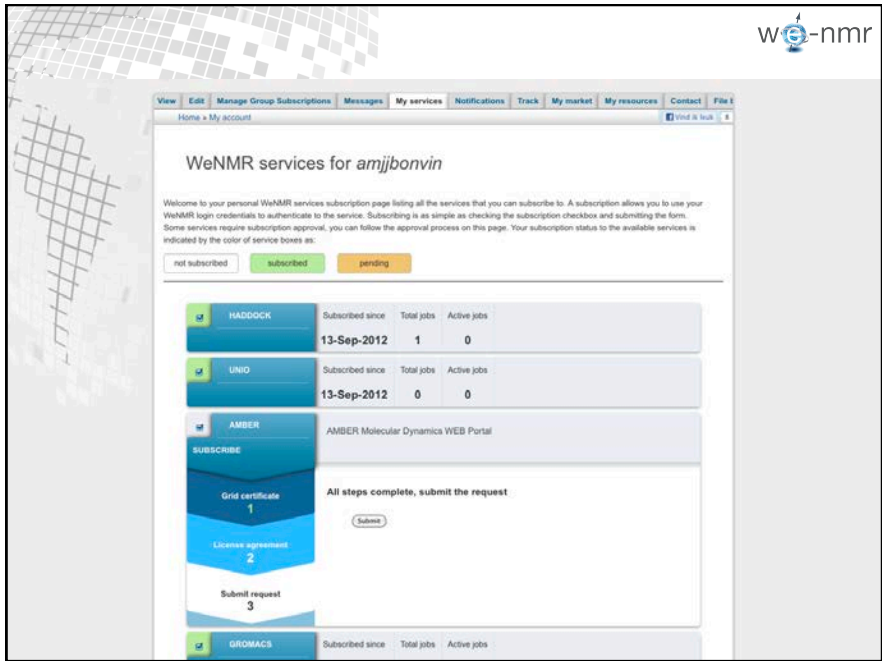

The WeNMR VRC
Consultancy
Tutorials, Wiki
Help Center
Knowledge
Third-party aggregation
Grid
Services
VRC
NEW!
SSO
Marketplace
Portals
Blogs, news, events..
Facebook
Exposure
User

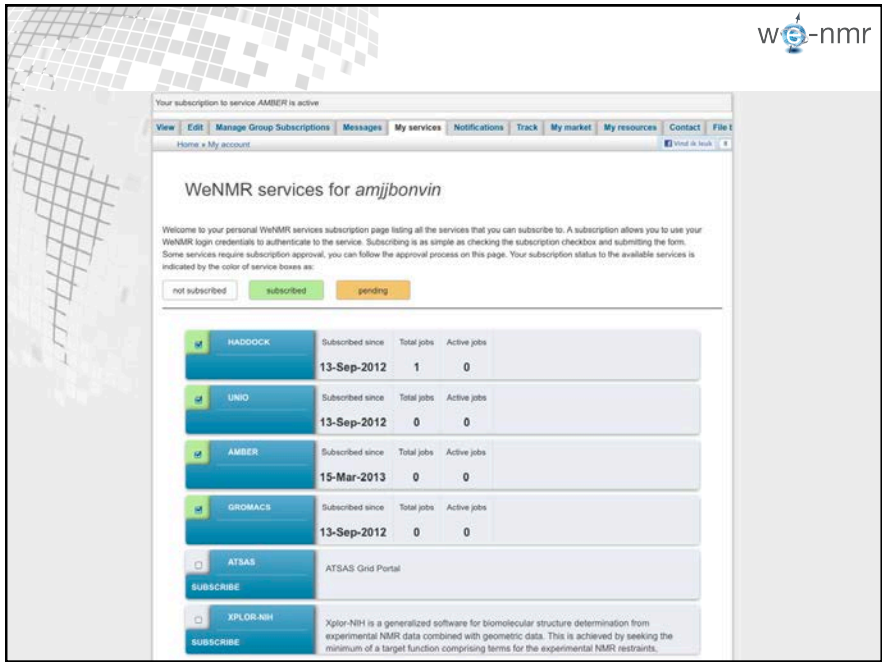

WeNMR services for amjjbonvin
HADDOCK
UNIO
AMBER
AMBER Molecular Dynamics WEB Portal
GROMACS
ATSAS
ATSAS Grid Portal
XPLOR-NIH
SUBSCRIBE
Submit

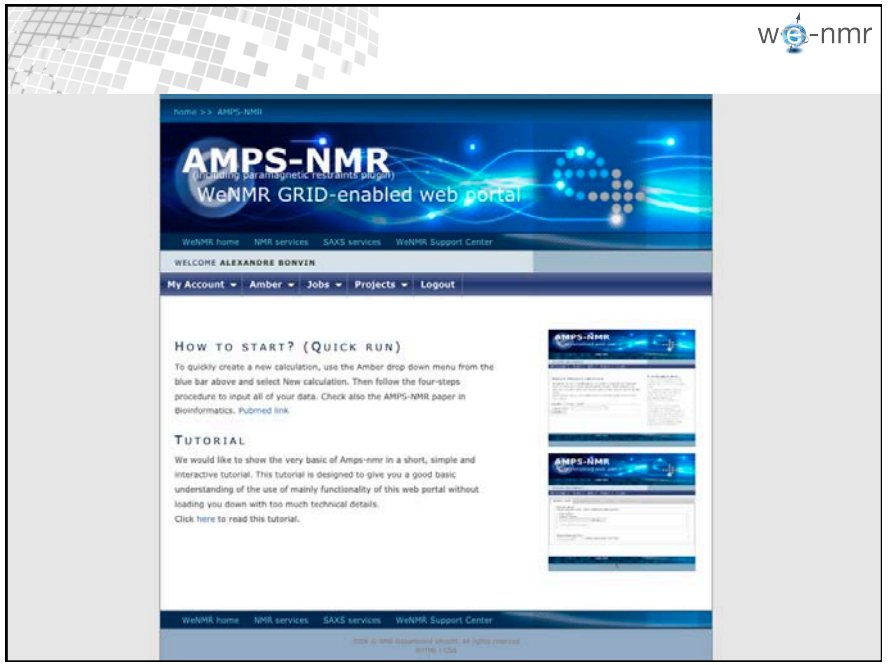

WeNMR services for amjjbonvin
HADDOCK
UNIO
AMBER
AMBER Molecular Dynamics WEB Portal
SUBSCRIBE
Grid certificate
License agreement
Submit request
You are 2 steps away from subscribing to this service
Follow the steps listed in the menu on the left to subscribe
GROMACS


WeNMR services for amjjbonvin
HADDOCK
UNIO
AMBER
AMBER Molecular Dynamics WEB Portal
SUBSCRIBE
Grid certificate
Step 1
This service requires you to have a valid Grid certificate
Congratulations, you have a valid GRID certificate still valid for 159 days
License agreement
Submit request
GROMACS


WeNMR services for amjjbonvin
HADDOCK
UNIO
AMBER
SUBSCRIBE
AMBER Molecular Dynamics WEB Portal
Step 2 This service requires approval of a license agreement
AMBER 12 SOFTWARE LICENSE AGREEMENT
AMBER Software License

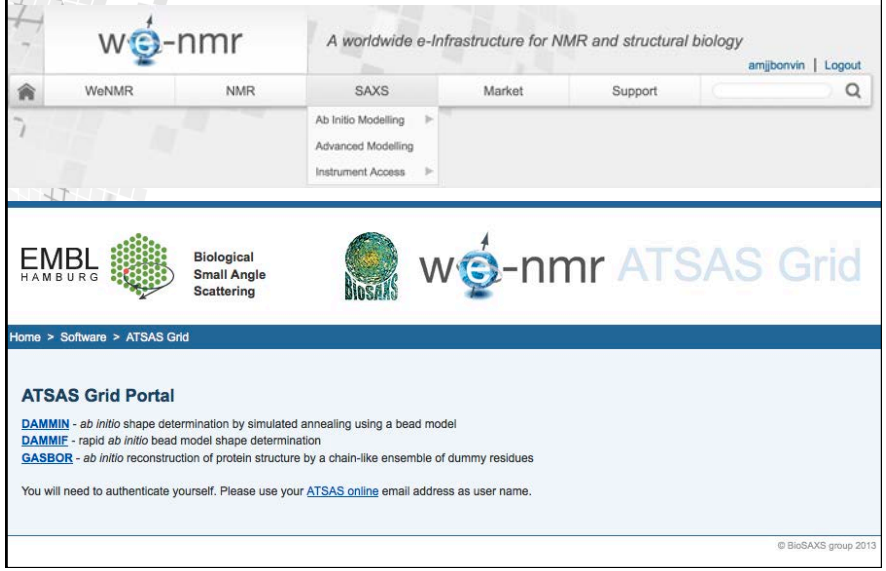

WeNMR services for amjjbonvin
HADDOCK
UNIO
AMBER
SUBSCRIBE
AMBER Molecular Dynamics WEB Portal
All steps complete, submit the request
GROMACS


Your subscription to service AMBER is active
WeNMR services for amjjbonvin
HADDOCK
UNIO
AMBER
GROMACS
ATSAS
ATSAS Grid Portal
XPLOR-NIH
Xplor-NIH is a generalized software for biomolecular structure determination from experimental NMR data combined with geometric data.

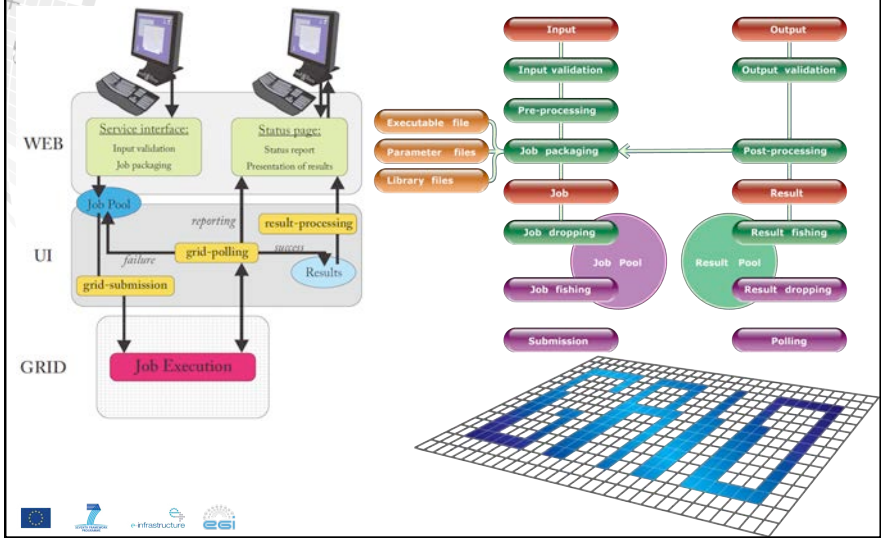

AMPS-NMR
WeNMR GRID-enabled web portal
WeNMR home
NMR services
SAXS services
WeNMR Support Center
WELCOME ALEXANDRE BONVIN
My Account
Amber
Jobs
Projects
Logout
How to start? (Quick run)
To quickly create a new calculation, use the Amber drop down menu from the blue bar above and select New calculation. Then follow the four-steps procedure to input all of your data. Check also the AMPS-NMR paper in Bioinformatics. Pubmed link
Tutorial
We would like to show the very basic of Amps-nmr in a short, simple and interactive tutorial. This tutorial is designed to give you a good basic understanding of the use of mainly functionality of this web portal without loading you down with too much technical details.
Click here to read this tutorial.

The WeNMR services portfolio

And SAXS services as well!



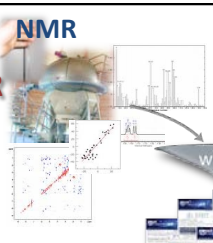
Behind the scene of most portals



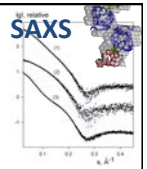
Services and Usage

A worldwide e-Infrastructure for NMR and structural biology

NMR



SAXS



WeNMR VRC (Nov. 2013)

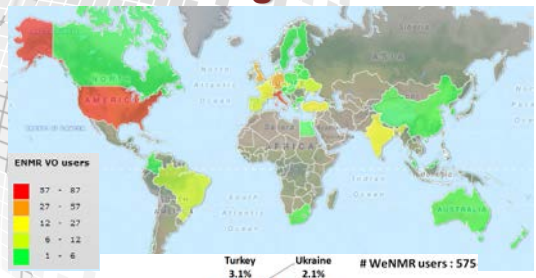
- One of the largest (#users) VO in life sciences
- > 575 VO registered users (36% outside EU)
- > 900 VRC members
- ~ 82 000 CPU cores
- > 2M SI2K CPU years
- > 5M jobs User-friendly access to Grid via web portals

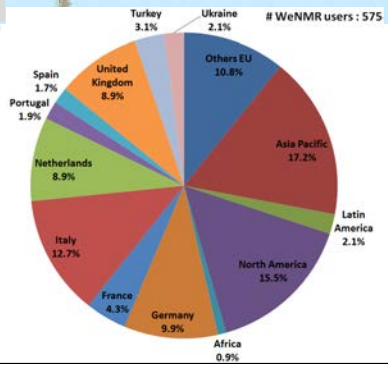
www.wenmr.eu

w-nmr

VO registered user distribution

w-nmr

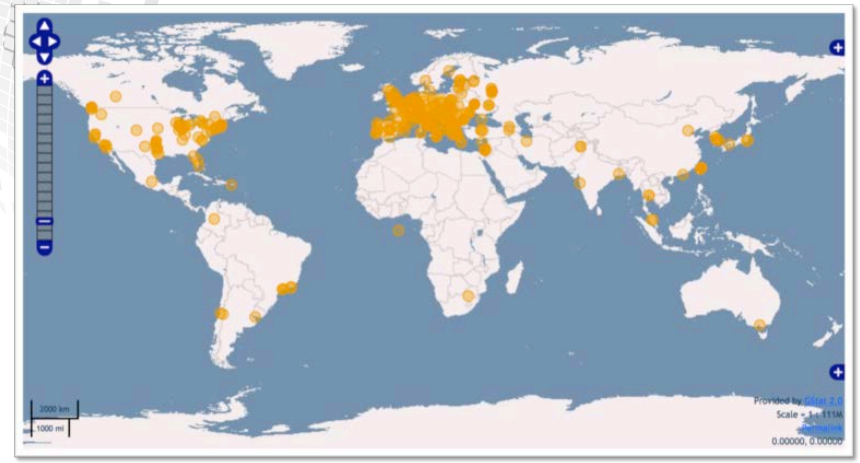




- 575 users from 42 countries
- 36% outside Europe
- 9.7 users/month growth
- enmr.eu largest VO in life sciences

GRID resource centers worldwide

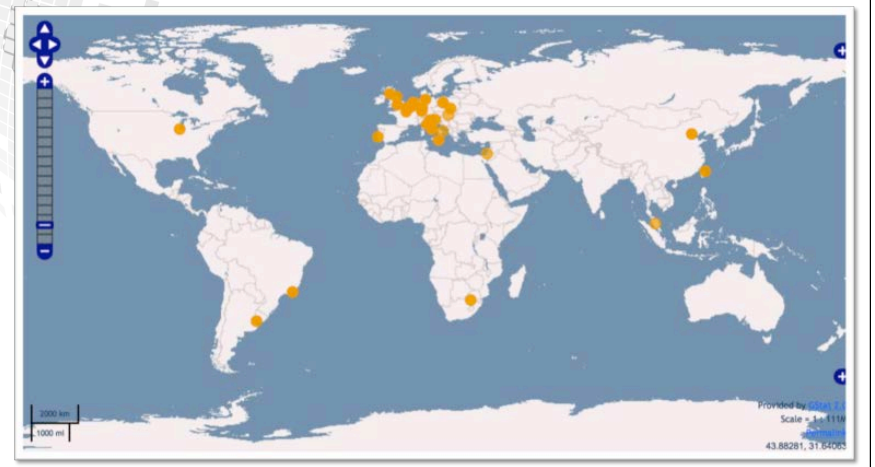
w-nmr



http://gstat2.grid.sinica.edu.tw

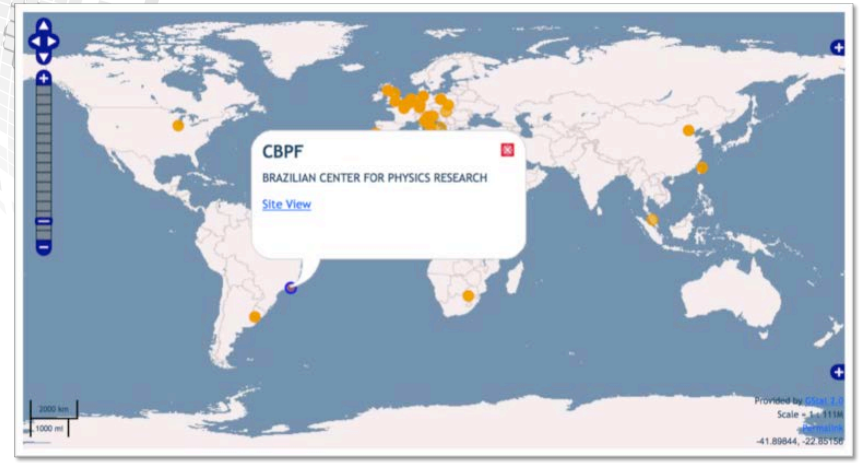
GRID resource centers supporting the enmr.eu VO

w-nmr



http://gstat2.grid.sinica.edu.tw

GRID resource centers supporting the enmr.eu VO



http://gstat2.grid.sinica.edu.tw

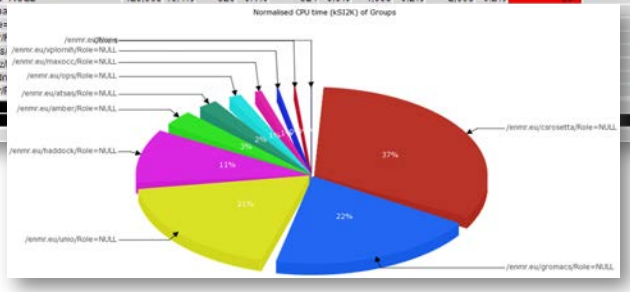
Application accounting

May 2013 - October 2013

VO Groups ordered by Normalised CPU time (kSI2K)

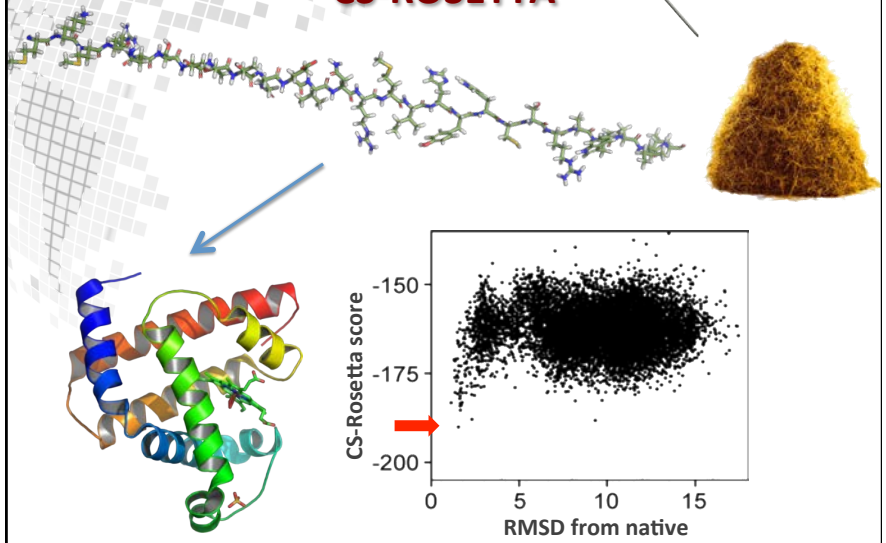
#	Group	Jobs		CPU time		Norm. CPU time		WCT		Norm. WCT		CPU Efficiency	Avg. CPU time	Avg. WCT
		#	%	Hrs	%	Hrs	%	Hrs	%	Hrs	%	%	Hrs	Hrs
1	/enmr.eu/csrosetta/Role=NULL	131,052	16.6%	239,691	42.5%	565,256	37.4%	243,823	46.2%	572,776	44.3%	98.3	1.83	1.86
2	/enmr.eu/gromacs/Role=NULL	1,339	0.2%	76,121	13.5%	339,187	22.5%	13,766	2.6%	59,002	4.6%	553.0	56.85	10.28
3	/enmr.eu/unio/Role=NULL	93,024	11.8%	103,178	18.3%	315,093	20.9%	105,778	20.1%	320,924	24.8%	97.6	1.11	1.14
4	/enmr.eu/haddock/Role=NULL	167,703	21.3%	71,224	12.6%	172,611	11.4%	76,346	14.5%	184,484	14.3%	93.3	0.42	0.46
5	/enmr.eu/amber/Role=NULL	3,639	0.5%	18,395	3.3%	48,539	3.2%	18,571	3.5%	48,960	3.8%	99.1	5.05	5.10
6	/enmr.eu/atsas/Role=NULL	51,634	6.5%	12,997	2.3%	31,899	2.1%	14,263	2.7%	34,688	2.7%	91.1	0.25	0.28
7	/enmr.eu/ops/Role=NULL	191,822	24.3%	7,539	1.3%	20,305	1.3%	16,940	3.2%	46,526	3.6%	44.5	0.04	0.09
8	/enmr.eu/maxocc/Role=NULL	5,453	0.7%	32,298	5.7%	10,963	0.7%	34,940	6.6%	15,235	1.2%	92.4	5.92	6.41
9	/enmr.eu/xplornih/Role=NULL	12,498	1.6%	953	0.2%	3,046	0.2%	1,017	0.2%	3,176	0.2%	93.7	0.08	0.08
10	/enmr.eu/None	1,418	0.2%	584	0.1%	2,433	0.2%	602	0.1%	2,500	0.2%	97.0	0.41	0.42
11	/enmr.eu/Role=NULL	129,065	16.4%	320	0.1%	524	0.0%	1,099	0.2%	2,608	0.2%	29	0.00	0.01
12	/enmr.eu/cyana												1.25	1.25
13	/enmr.eu/Role=												0.22	0.30
14	/enmr.eu/bcbr/R												0.46	0.49
15	/enmr.eu/mars/												0.00	0.00
16	/enmr.eu/bmrz/												0.00	0.00
17	/enmr.eu/mddn												0.00	0.00
18	/enmr.eu/bcbr/R												0.00	0.00
													0.71	0.67

Normalised CPU time (kSI2K) of Groups



A few examples: CS-Rosetta

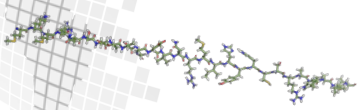
CS-ROSETTA

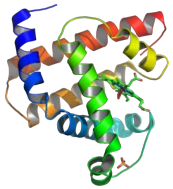


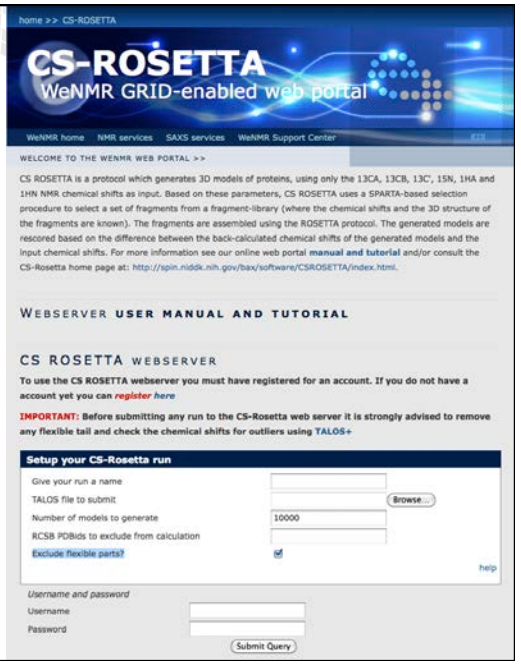
CS-ROSETTA (not) web portal

/Volumes/Home/Users/christophe/minirosetta_svn/bin/minirosetta.macosgccrelease -abinitio::increase_cycles 1 -nstruct 10 -database /Volumes/Home/Users/christophe/minirosetta_database -frag3 /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/DATA/1NKU/FRAG_CS-ROSETTA/frag3.t000.rosetta.tab.gz -frag9 /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/DATA/1NKU/FRAG_CS-ROSETTA/frag9.t000.rosetta.tab.gz -abinitio::stage1_patch /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/SCORE/score0_pcs1.wts_patch -abinitio::stage2_patch /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/SCORE/score1_pcs1.wts_patch -abinitio::stage3a_patch /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/SCORE/score2_pcs1.wts_patch -abinitio::stage3b_patch /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/SCORE/score5_pcs1.wts_patch -abinitio::stage4_patch /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/SCORE/score3_pcs1.wts_patch -native /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/DATA/1NKU/idealized_1NKU.pdb -out:file:silent /Volumes/Home/Users/christophe/PCS_ROSETTA_RESULT/ABINITIO//1NKU_exact_N1_PCS1.silent -out:file:scorefile /Volumes/Home/Users/christophe/PCS_ROSETTA_RESULT/ABINITIO//1NKU_exact_N1_PCS1.sc -abinitio::rg_reweight 0.5 -abinitio::rsd_wt_helix 0.5 -abinitio::rsd_wt_loop 0.5 -abinitio::use_filters false -broker::setup /Volumes/Home/Users/christophe/Dropbox/minirosetta_pcs_file/RUN/./../1NKU_exact/setup_pcs1.txt -run:protocol broker -overwrite -PCS:normalization_id 1 -in::file::native_exclude_res 1 2 3 4 5 6 7 8 9 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 -mute core.scoring.methods.PCS.PcsEnergy -mute core.optimization.LineMinimizer

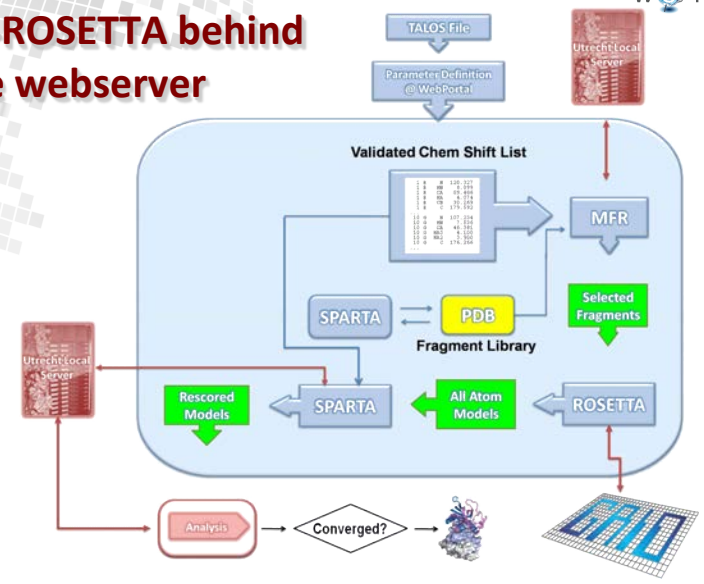
CS-ROSETTA web portal





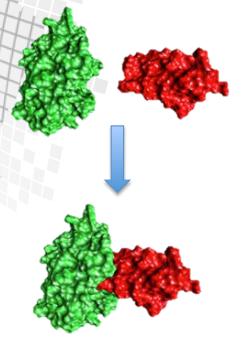


CS-ROSETTA behind the webserver

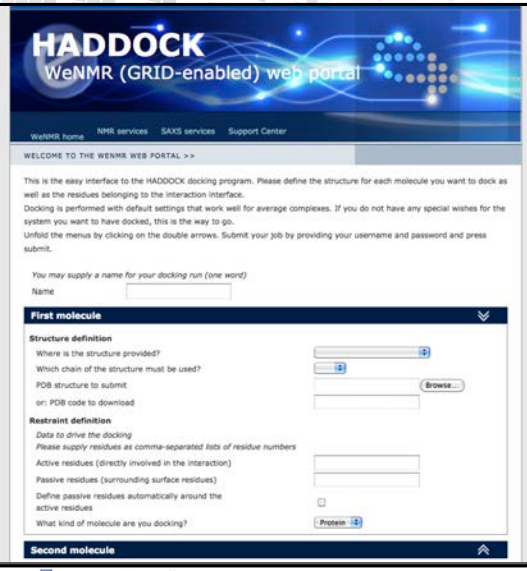


HADDOCK: Shedding light on biomolecular interactions

Haddock web portal

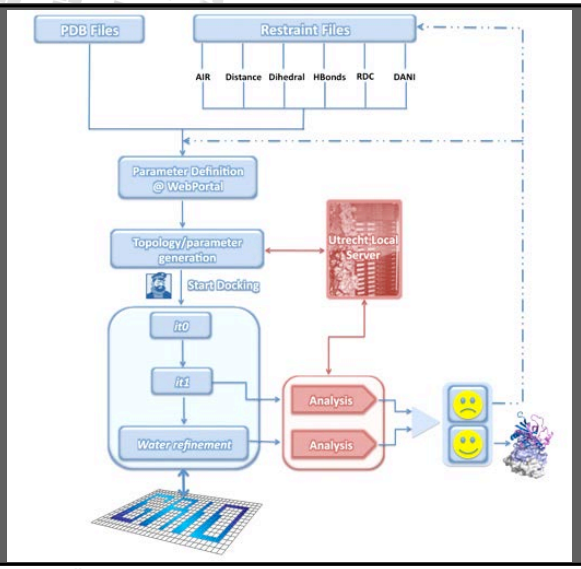


User friendly easy interface



- 3600 registered users
- > 60000 served runs since June 2008
- ~ 12% on the GRID

What is happening behind the scene?



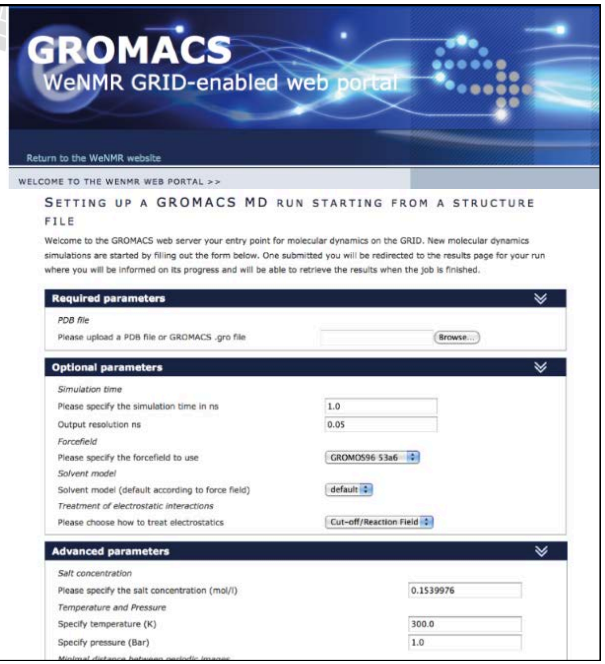
A few examples: GROMACS

GROMACS web portal

Runs in parallel on the Grid using GROMACS multithreading capability

Currently for efficiency runs on 6 processors

CREAM-CE features used for parallel job submission



w-nmr

Involving software developers and pushing new developments: CASD-NMR

CORRESPONDENCE

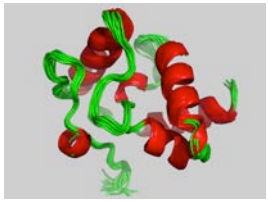
w-nmr

Nature Methods **6**, 625 - 626 (2009)
doi:10.1038/nmeth0909-625

CASD-NMR: critical assessment of automated structure determination by NMR

Antonio Rosato[1,2], Anurag Bagaria[3,4], David Baker[5], Benjamin Bardiaux[6], Andrea Cavalli[2], Jurgen F Doreleijers[6], Andrea Giachetti[1], Paul Guerry[9], Peter Güntert[3,4], Torsten Herrmann[9], Yuanpeng J Huang[10], Hendrik R A Jonker[4,11], Binchen Mao[10], Thérèse E Malliavin[6], Gaetano T Montelione[10], Michael Nilges[6], Srivatsan Raman[5], Gijs van der Schot[12], Wim F Vranken[13], Geerten W Vuister[8] & Alexandre M J J Bonvin[12]

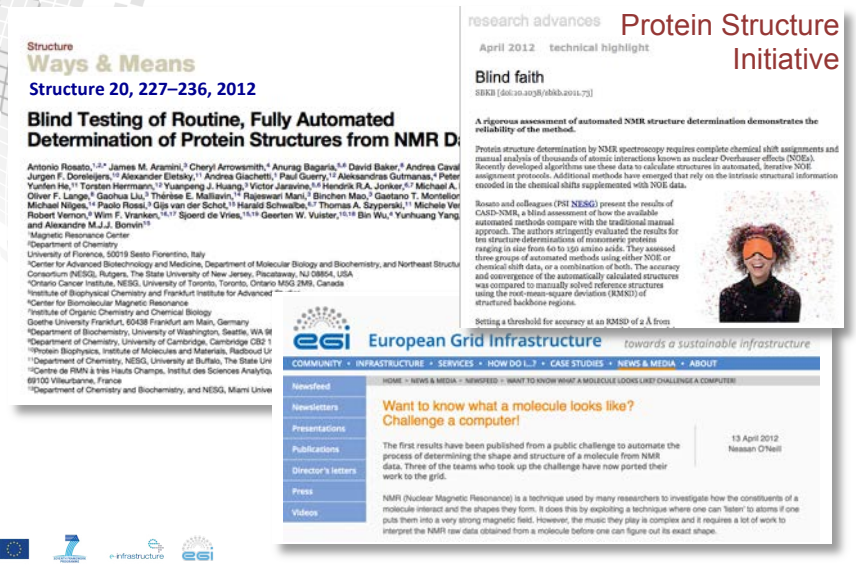
nature methods
Techniques for life scientists and chemists



- **10 blind targets** (structures to be determined by fully automated methods
- **15 teams worldwide** involved with their respective developer teams
- Assessment of results by independent team
- **Impact: WeNMR & EU as drivers of innovation, leading to to new software developments, increased trust in automated methods and improved services for end users**

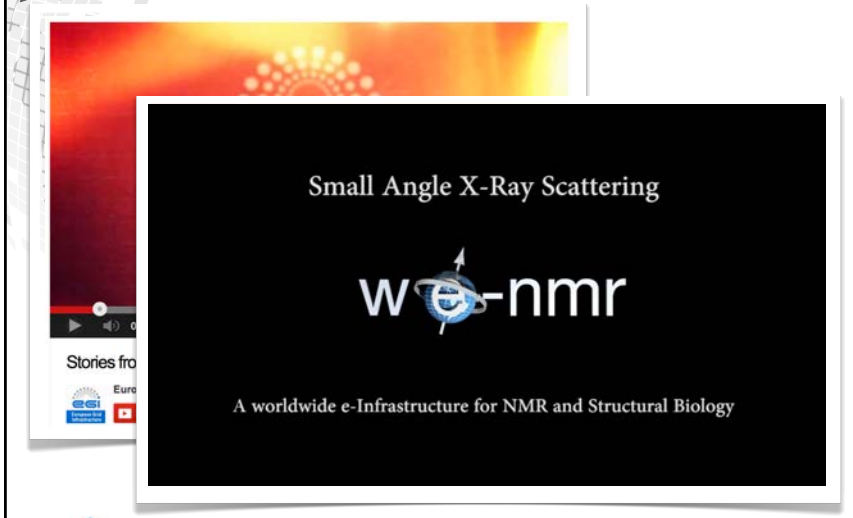
CASD-NMR

w-nmr



Dissemination highlights

Dissemination highlights: general public



Acknowledgments

The team

Alexandre M.J.J. Bonvin
Andrea Giachetti
Antonio Rosato
Anurag Bagaria
Christophe Schmitz
Eric Frizziero
Gijs van der Schot
Harald Schwalbe
Hendrik R. A. Jonker
Ivano Bertini
Johan van der Zwan
Lucio Ferella
Marc van Dijk
Marco Verlato
Mikael Trellet
Mirco Mazzucato
Nuno Loureiro-Ferreira
Peter Güntert
Rolf Boelens
Sjoerd J. de Vries
Stefano Dal Pra
Torsten Herrmann
Tsjerk A. Wassenaar
Victor Jaravine
Wim F. Vranken
Danny Hsu
Simon Lin

The partners

Utrecht University, Bijvoet Center for Biomolecular Research, NL
Johann Wolfgang Goethe Universität Frankfurt a.M., Center for Biomolecular Magnetic Resonance DE
University of Florence, Magnetic Resonance Center, IT
Istituto Nazionale di Fisica Nucleare , Padova, IT
Raboud University, Nijmegen, NL
University of Cambridge UK
European Molecular Biology Laboratory, Hamburg, DE
Spronk NMR Consultancy, LT
Academia Sinica

Contract n°: RI-261572
Project type: CP-CSA