



Umbrella for Photon / Neutron Community

PaNdata Partners



- Alba, Spanish National Sychrotron Facility
- Diamond UK Synchrotron facility
- European Synchrotron Radiation Facility (ESRF)
- Deutsches Elektronen Synchrotron (DESY)
- Institut Laue–Langevin (ILL)
- Max IV Laboratory Lund
- ISIS STFC Neutron source
- HZB, Helmholtz Zentrum Berlin
- Paul Scherrer Institut (PSI), hosting SINQ and SLS
- Soleil, French National Synchrotron Facility

CRISP IT Partners



- European Synchrotron Radiation Facility (ESRF)
- Deutsches Elektronen Synchrotron (DESY)
- European Organisation for Nuclear Research (CERN)
- European Spallation Source (ESS)
- GSI Helmholtz Centre for Heavy Ion Research (GSI)
- Institut Laue-Langevin (ILL)
- European X-ray Free Electron Laser (XFEL)
- Paul Scherrer Institut (PSI)

The user community I

☐ *Photon facilities*

- ✓ Synchrotrons and Free Electron Lasers (FELs)
- ✓ Light of highest brightness
- ✓ About 15 synchrotrons in EU (ESRF + national)

☐ *FELs, even 10^3 to 10^6 times brighter*

- ✓ SLAC/Stanford, DESY/Hamburg, FEL/Spring-8/Japan, PSI/Villigen
- ✓ Membrane proteins; microscopic movies of chemical reactions

☐ *Neutron facilities*

- ✓ Complementary
- ✓ Similar user community

☐ *Small teams, visit for*

- ✓ Few hours (structural biology) to
- ✓ Few weeks (superconductivity, nano investigations)

The user community II

- ❑ ***In EU >> 30'000 visiting users /y***
 - ✓ Organised by local user offices
 - ✓ Large overbooking ($\geq 3:1$), low chance to be accepted
 - ✓ Important to minimize administrative load
- ❑ ***On-site visits***
 - ✓ Short duration
 - ✓ In part spontaneous (keep that attraction)
 - ✓ Part-time users
- ❑ ***Decentralized structure (compare e.g. to CERN)***
 - ✓ Manifold research fields
 - ✓ Many data sources facilities
 - ✓ National character of facilities, report to own governments
- ❑ ***Zoo of research areas***
 - ✓ Archaeology, chemistry, materials + analytical sciences, life sciences
 - ✓ Physics is minority
 - ✓ **Linking element is common use of large facilities (not science field) !**

What are the IT requests?

- ❑ ***Huge datasets***
 - ✓ Novel 2D detectors, quantum leap in data quality, but also data volumes
 - ✓ Multi-image techniques (tomography, lens-less imaging)
 - ✓ Molecular movies at FELs
 - ✓ 'Petabyte' 'normal' unity; time over for 'hard-disk in the trouser pocket'
- ❑ ***Trans-facility experiments***
 - ✓ Single Sign On (SSO)
 - ✓ Standardize proposal procedures on EU scale
- ❑ ***Remote data access***
 - ✓ Analyze data remotely at facility
 - ✓ Combine datasets taken at different facilities
 - ✓ Clouds (commercial, community-based)
 - ✓ Respect confidentiality restrictions
- ❑ ***Remote experiment access***
 - ✓ Basic: passive online access to measured data
 - ✓ Advanced: active control
- ❑ ***PR Issues***
 - ✓ Improve corporate identity
 - ✓ Improve public lobbying

Umbrella as Prototype

- ❑ ***Incorporate confidentiality aspects***
 - ✓ High competition, especially structural biology
 - ✓ Time-window-structured access to experiments and data
- ❑ ***Rely on existing local user office structure***
 - ✓ Great experience
 - ✓ DIY (Do It Yourself) operation
 - Users: manage their personal entries
 - User offices: supervising; manage authorizations
- ❑ ***Base system on professional authentication standard***
 - ✓ Shibboleth, federated Single-Sign-On System (SAML), widely used
 - ✓ Special photon / neutron user federation
 - ✓ Only one identity provider
 - ✓ Supervising by local User Offices
- ❑ ***Concept***
 - ✓ Unique user identification on EU scale
 - ✓ Hybrid information storage
 - ✓ No possibility for cross-facility information pull
 - ✓ Multi-level identification (maximum autonomy to facilities)
 - ✓ Waterproof but slim data protection system

Operation concept

Bottom-up: Delegation and direct feedback

❑ *Facilities*

- ✓ Keep existing administration structures as much as possible
 - o Proposal workflow
 - o Guest house / restaurant, access badges, stock room, ...
- ✓ During implementation parallel operation
 - o smooth transition
 - o No time-zero

❑ *Users*

- ✓ DIY (Do It Yourself) operation
 - o Users: manage their personal entries
 - o User offices: supervising; manage authorizations

❑ *Collaborations*

- ✓ Self organization of data access via collaborations
- ✓ Principal investigator / main proposer controls who is allowed to access data

❑ *Applications*

- ✓ Multi-level trust
- ✓ applications define level
- ✓ Lowest level: Google-type handshake
- ✓ Higher level: authentication at facility user offices, no external ??

The Umbrella Concept

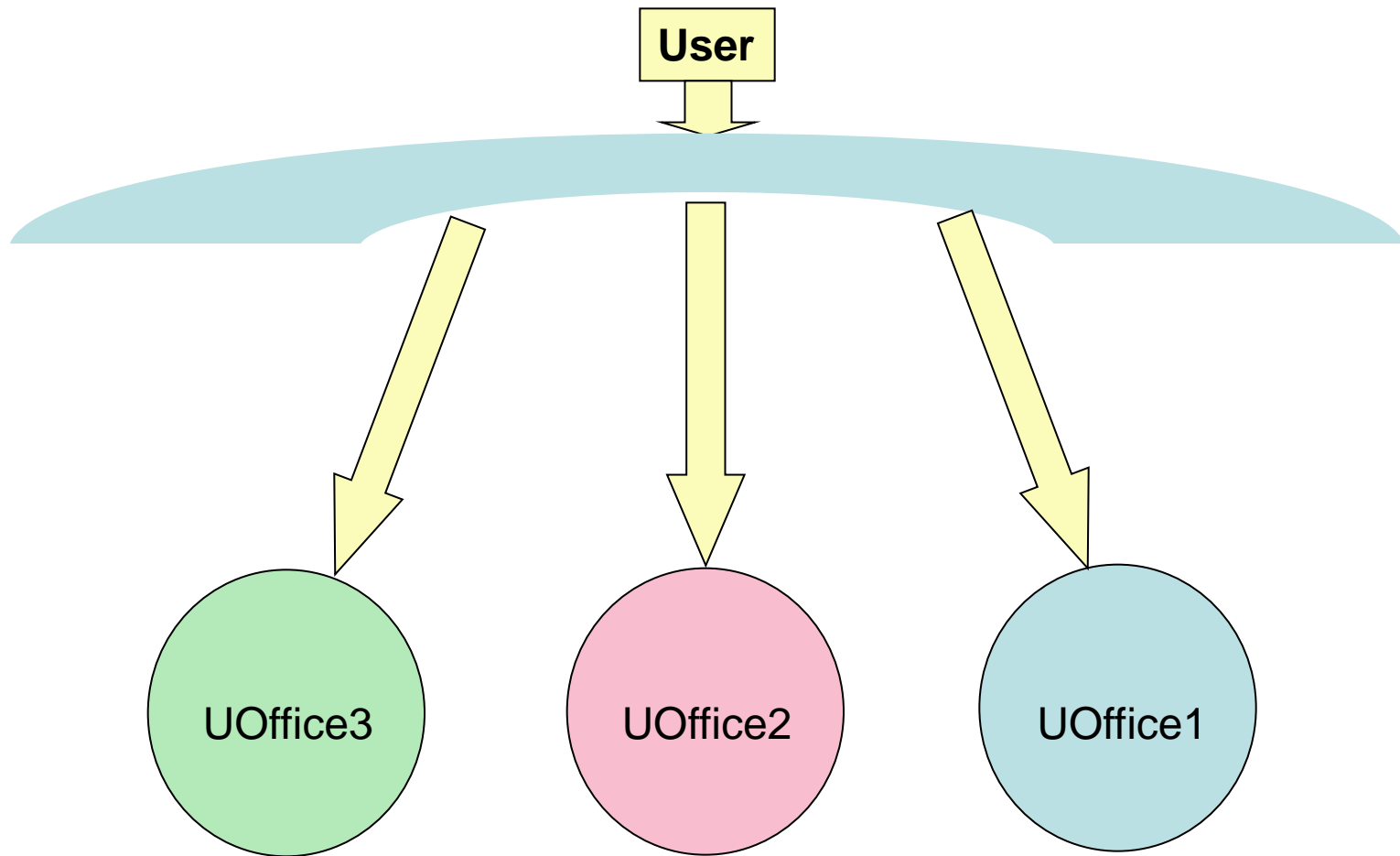
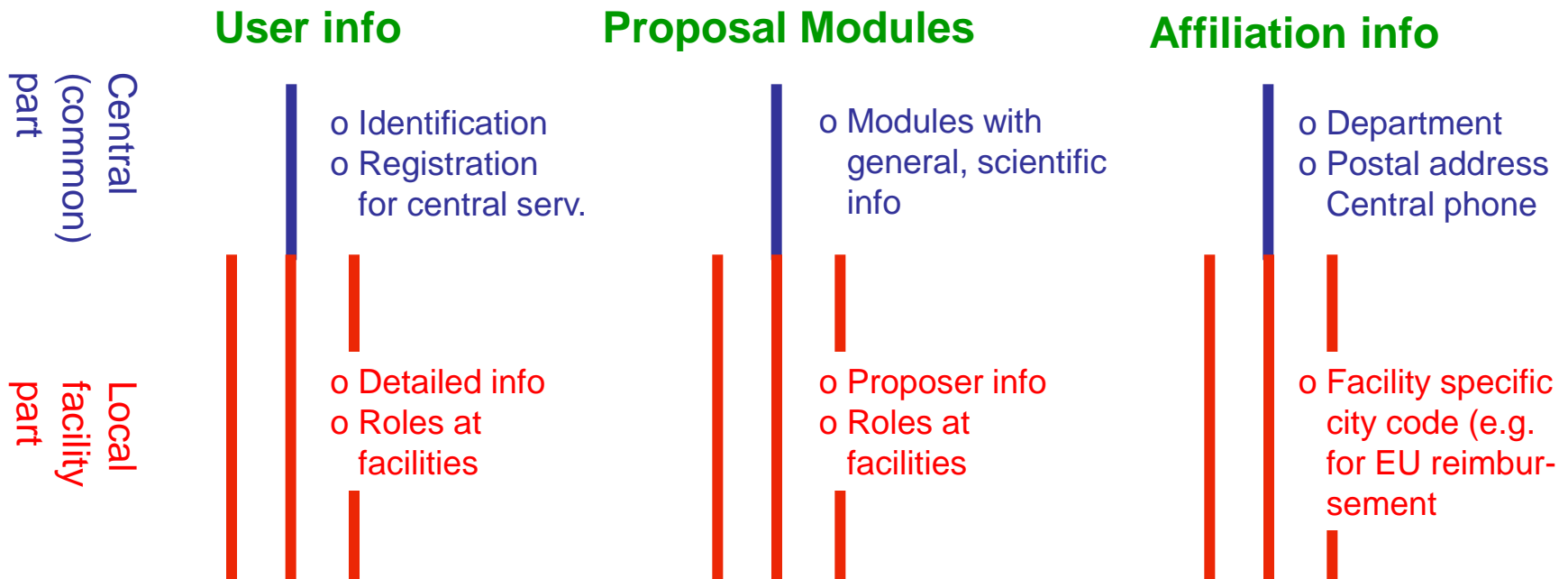


Fig.1

Hybrid concept (central and federated)

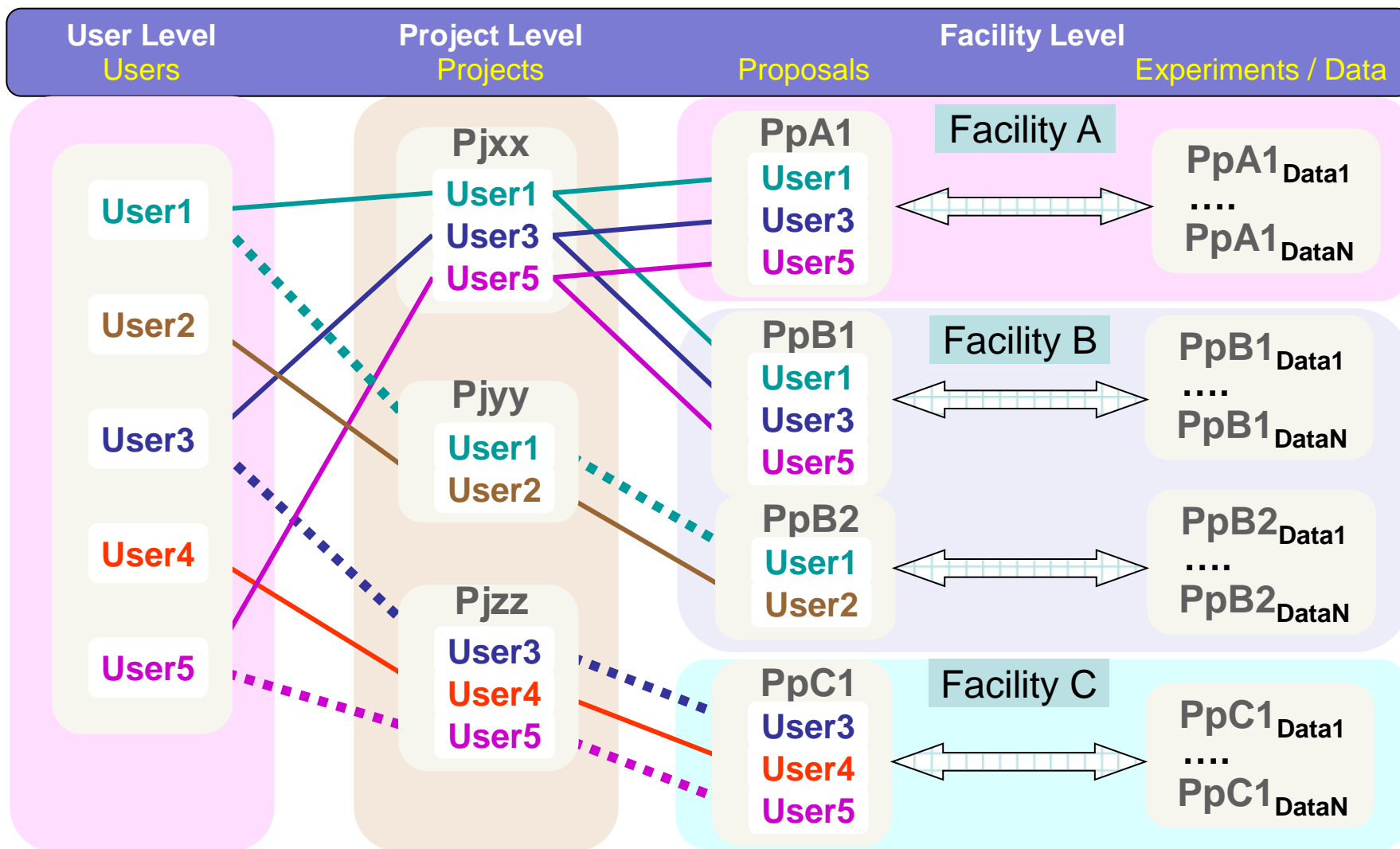
Answer to conflicting requests:

- Efficient technology
- Confidentiality
- Consequent distinction of authentication and authorisation



Remote data access, concept proposed

- ❑ ***Embargo vs. post-embargo period***
 - ✓ Here only embargo (most critical, confidentiality)
- ❑ ***Standard access rights rule***
 - ✓ No chance for manual central authorization
 - ✓ 1'000s of experiments, 10'000s of users
- ❑ ***Identity by Umbrella***
 - ✓ Unique, EU-wide user authentication
 - ✓ Allows trans-facility actions, Single Sign On
- ❑ ***Keep Role of proposal as organizing element***
 - ✓ Who participates in experiment, has access right to data
 - ✓ Principal investigator / main proposer



Bridging

□ *Umbrella Plus*

❖ Proposal-based user administration

- Linking via Umbrella to local WUOs: includes full user services
- Remote file access, remote experiment access + ...

❖ Non-proposal-based user administration

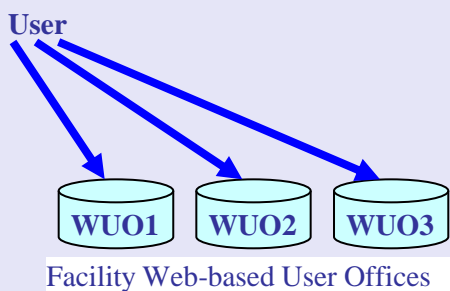
- HEP-type operation (very long-term proposals)
- Small facilities (e.g. university labs, ...)
- May have need for user db, but not for the rest
- Umbrella + stripped-down version of a WUO
 - Core user db
 - Shibboleth communication
 - Green / red lamp at the output

□ *Umbrella Bio*

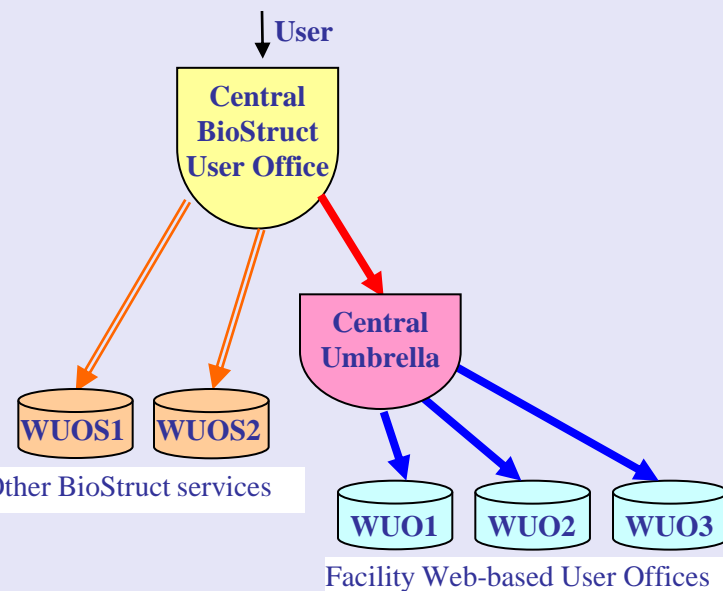
- ❖ Currently 2 decoupled user review/access schemes
- ❖ Combine Umbrella + BioStruct

Umbrella and BioStruct

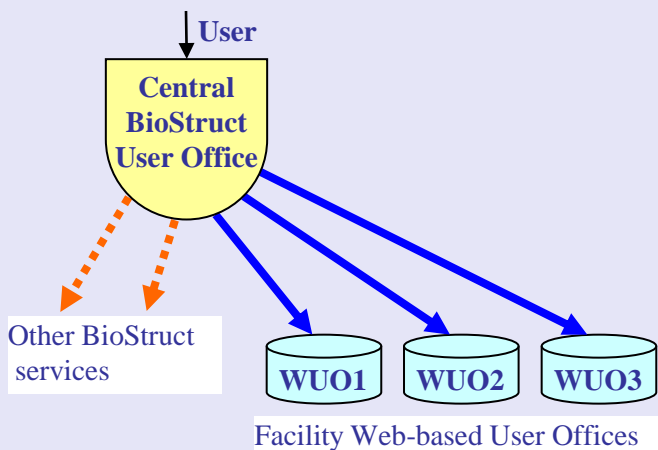
a) Standard



c) BioStruct with Umbrella



b) BioStruct as present present



Friendly user phase

- ❑ **Goal and duration**
 - ❖ Test of the system by future users
 - ❖ February 1 – March 31
- ❑ **Central Applications**
 - ❖ Prototype of central web site
 - ❖ EAA: registration, mutation
 - ❖ Alfresco, Indico, Issue tracker, Wiki
- ❑ **Federated applications**
 - ❖ Umbrella + WUO clone versions
- ❑ **Participants**
 - ❖ Facilities
 - DESY
 - Diamond (iCAT service, Moonshot?)
 - ESRF
 - PSI
 - ❖ 'Friendly' users
 - ~30, all over EU
 - External expert users (ESUO, ETH, BioStruct, ??)
 - Local facility experts (DESY)

Umbrella road map

- ❑ ***till January, 31: Umbrella preparation***
 - ❖ Definition of active participants
 - ❖ Definition of elements to offer to users
 - ❖ Definition of web portal
 - ❖ Documentation
 - ❖ Final developments

- ❑ ***from February 1, Friendly user phase***
 - ❖ Contact of users
 - ❖ Umbrella + WUO test versions (DESY, PSI, ESRF, Diamond)

- ❑ ***from May 31***
 - ❖ Workshop with all participants
 - ❖ Concluding feedback document
 - ❖ Implementation of feedback
 - ❖ **Legal work (trust issues, MoUs, ...)**

- ❑ ***from September 1, Ready for implementation***

Conclusion

- ❑ ***Clear demands at large photon / neutron facilities***
 - ✓ Unique user ID
 - ✓ Remote data and experiment access
 - ✓ Need for user and facility friendliness
 - ✓ Very large number of visiting scientists: Need slim and efficient system

- ❑ ***Limited excitement on management (and user?) side***
 - ✓ Resources
 - ✓ Confidentiality
 - ✓ Scientific competition

- ❑ ***Overlapping IT communities, bridging***
 - ✓ Large facilities and universities (educational sector)
 - ✓ Large facilities and university labs
 - ✓ Different communities

- ❑ ***Umbrella as prototype***
 - ✓ Common web portal
 - ✓ Slim solution, no top down organization, self service elements
 - ✓ Build on existing infrastructure, clear topology, avoid parallel worlds