

eResearch 2020

The role of e-Infrastructures in the creation of
global virtual research communities



Impact self-assessments from the eResearch2020 study by Franz Barjak, FHNW

BELIEF Brainstorming Workshop, 24 February 2010

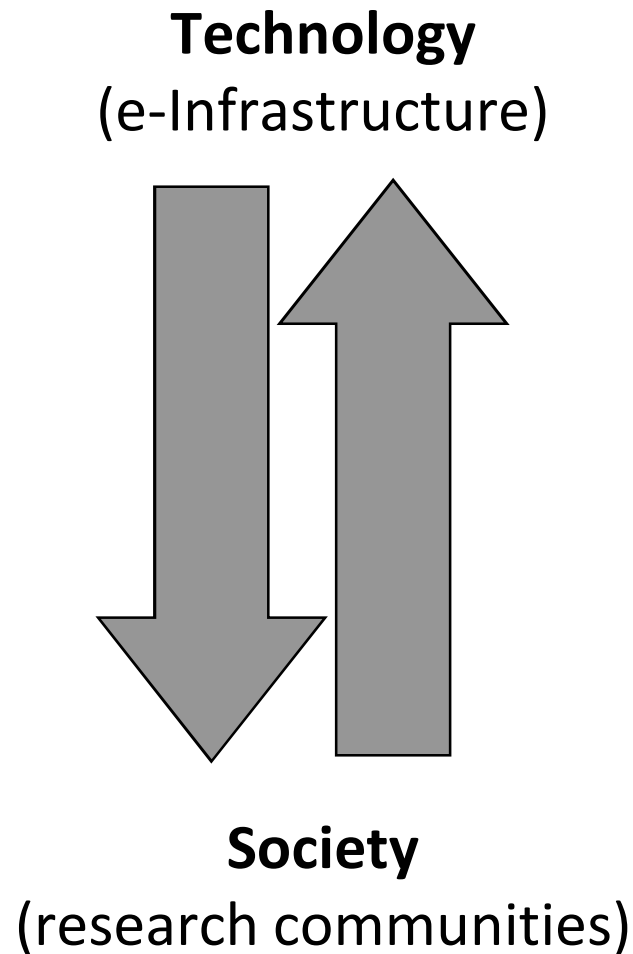


Fachhochschule Nordwestschweiz
Hochschule für Wirtschaft

Content

1. Conceptualizing the relationship between e-Infrastructures and research communities
2. Impact assessments by people running e-Infrastructures
3. Impact assessments by people working with/in e-Infrastructures
4. Conclusions

Conceptualizing the relationship between e-Infrastructures and research communities



Technological determinism:
external, largely independent technologies determine or force change in the social system that can be measured, modelled and predicted

Social shaping of technology:
non-generic, configurable technologies are understood, adapted and used in different ways and according to the (dynamic) needs of specific communities

Data basis

- e-Infrastructure case studies
 - 18 cases
 - (Telephone) interviews
 - Document analysis
 - Extended case descriptions
 - Multi-case comparison
- Research Communities Survey
 - 407 respondents
 - Exploratory online survey among users and developers of e-Infrastructures
 - Descriptive statistics on responses

E-Infrastructure case studies

Impact as stated in case studies

C3-Grid	<ol style="list-style-type: none"> 1. New insights into the interaction of earth subsystems 2. Advanced new paradigm of distributed data 3. Model for a successful Grid project in Europe 	<p>Technical contributions Community building</p>
CineGrid	<ol style="list-style-type: none"> 1. Formation, growth and maturation of the CineGrid community 2. Raised trust and mutual understanding for sharing and joint use of resources 3. Raising awareness about new technologies among culture professionals 4. Demonstrated the feasibility of transmitting high-resolution video data and high-quality sound over very fast optical networks 	
CLARIN	– (infrastructure still in early phase)	
D4science	<ol style="list-style-type: none"> 1. Combined resources and field-specific tools to Virtual Research Environments 2. Initiated a dialogue between different communities 	
DARIAH	– (infrastructure still in early phase)	
DEISA	<ol style="list-style-type: none"> 1. Established collaboration between national supercomputing organizations 2. Facilitated set up of international projects which need supercomputer resources 	
DRIVER	<ol style="list-style-type: none"> 1. Software, portal and support network for digital repositories 2. Contributed to building a European community of digital repositories 	
EELA-2	<ol style="list-style-type: none"> 1. Spread the practice of Grid computing in Latin America 2. Infrastructure facilitates the production, mining, processing and analysis of data 3. Boost research collaboration in the Latin American Grid community 	

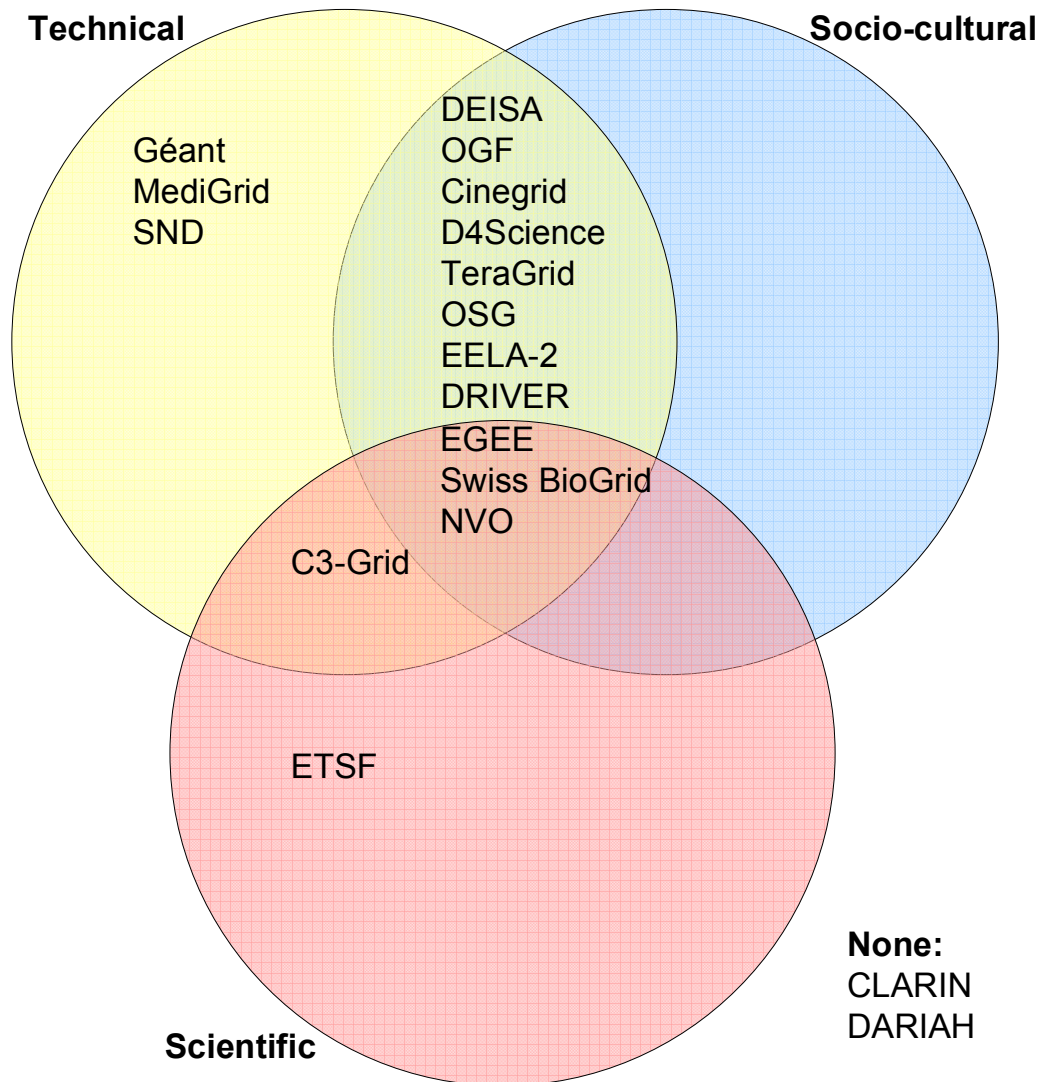
Impact as stated in case studies II

EGEE	<ol style="list-style-type: none"> 1. Establishment of European, globally well-connected, Grid communities 2. Advancement of Grid use in science 3. Contribution to new insights in many scientific fields 4. Development of gLite, a middleware package for grid computing 5. Wide collaboration network in Grid development 	Scientific insights
ETSF	<ol style="list-style-type: none"> 1. Proven a new way of collaboration between theorists and experimentalists 2. Sizable scientific output of publications and conference talks 	
GÉANT	<ol style="list-style-type: none"> 1. Facilitated research through providing data transmission services 2. Increased its spatial reach gradually beyond Western Europe 	
MediGrid	<ol style="list-style-type: none"> 1. Developed software and standards to be used in future projects 2. Raised participants' experience with Grid technology 	
NVO	<ol style="list-style-type: none"> 1. Advanced the integration of data in astronomy 2. Confirmed the existence of rare celestial objects and detected new ones 3. Enabled new web-based access to astronomical data 4. Collaboration support 5. Developed common data standards 6. Involved commercial partners 	
OGF	<ol style="list-style-type: none"> 1. Demonstrated the effectiveness of the open model of standardization 2. Fostered an open Grid community with commercial and academic participants 3. Developed necessary standards for e-Infrastructure providers 	7

Impact as stated in case studies III

OSG	<ol style="list-style-type: none"> 1. Technological provision of robust e-Infrastructure to main stakeholders 2. Sharing and dissemination of the distributed models undergirding the practice of distributed research
SND	Providing access to the uniquely good Swedish datasets
Swiss BioGrid	<ol style="list-style-type: none"> 1. Proof of concept of the feasibility of building bottom-up a Grid system 2. Precursor of the national Grid initiative 3. Technical achievement of bridging computers with different operating systems 4. Scientific successes in virtual screening and proteomics
TeraGrid	<ol style="list-style-type: none"> 1. Advanced the set of technologies required to integrate distributed heterogeneous supercomputers and other high end performing computers into a cohesive and persistent fabric 2. Built the social and organizational fabric for a community that has produced important technology advancements 3. Significant improvement in computational resources for scientists relying upon advanced computational infrastructure

Primary impact and contributions



- Technical contributions are standard
- Socio-cultural (including political) contributions are frequent
- Scientific contributions are rare (but informants were probably not fully aware of them)
- **Absent: economic or commercial effects**

Predominant e-Infrastructure goals

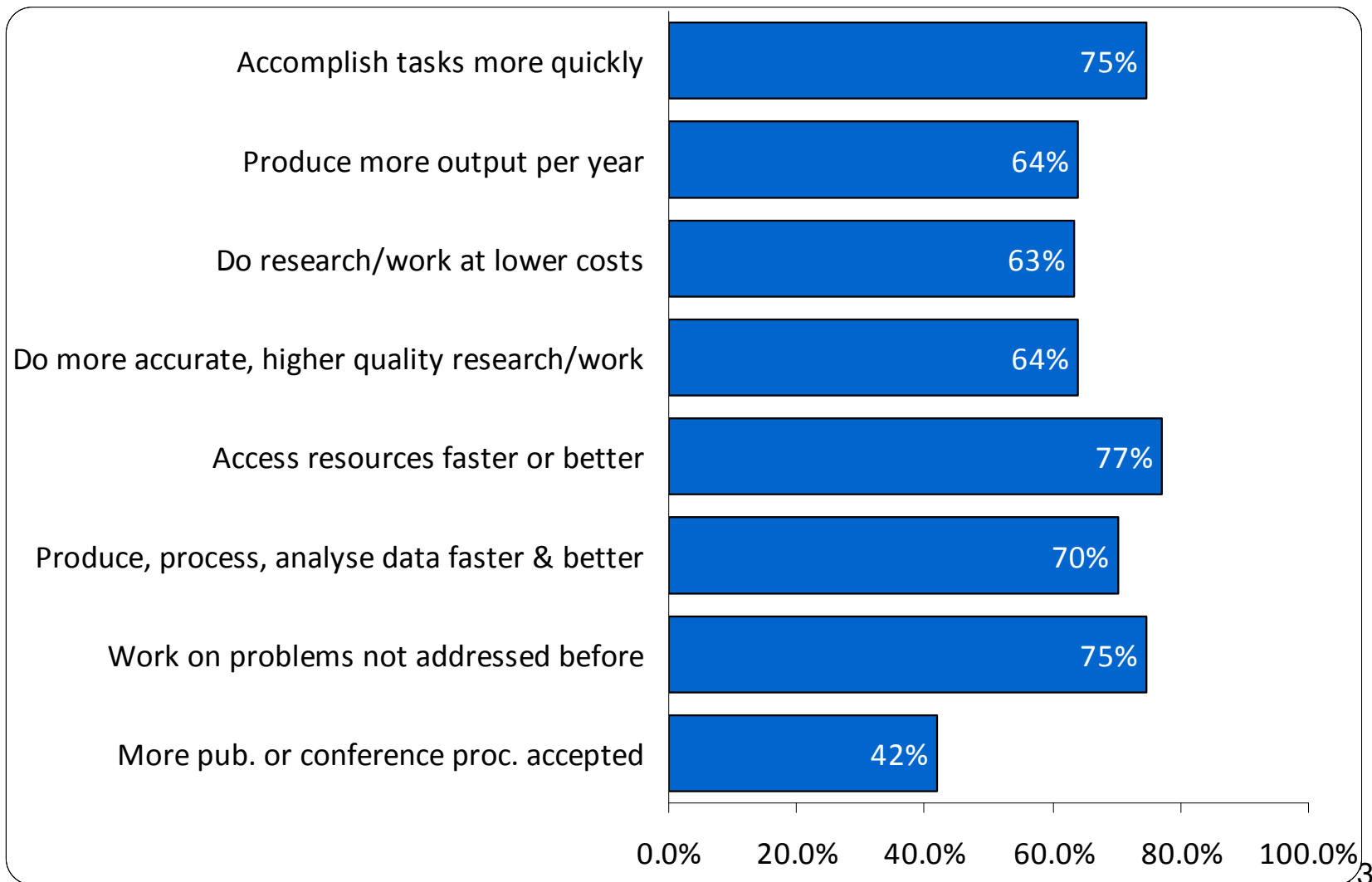
Technical (17)		Socio-cultural (12)		Scientific (9)
C3GRID	CineGrid	CineGrid	CLARIN	C3GRID
CLARIN	D4Science	D4Science	DARIAH	EGEE
DEISA	DRIVER	DEISA	DRIVER	ETSF
EELA-2	EGEE	EELA-2	GEANT	MediGrid
GEANT	MediGrid	OSG	SND	NVO
NVO	OGF	SWISS BIOGRID		OSG
OSG	SND	TeraGrid		SND
SWISS BIOGRID				SWISS BIOGRID
TeraGrid	DARIAH			TeraGrid

European effort strengthens community-building

	Scientific	Technical	Socio-cultural	Other	None
International	EGEE ETSF	CineGrid D4Science DEISA Driver EGEE EELA-2 Géant OGF	CineGrid D4Science DEISA Driver EELA-2 EGEE ETSF OGF	Géant	CLARIN DARIAH
National	C3-Grid NVO Swiss BioGrid	C3-Grid MediGrid NVO OSG SND Swiss BioGrid TeraGrid	NVO OSG Swiss BioGrid TeraGrid	NVO	

Research Communities Survey

Statements on impact of working with one particular e-Infrastructure



Impact by type of e-Infrastructure

	Geographic scope		Type of service		Driver	
	National	International	Computing	Data	Developer	Community
Accomplish tasks more quickly	65.3%	77.0%	76.1%	67.9%	77.9%	67.1%
Produce more output per year	50.0%	65.8%	71.1%	37.0%	68.5%	50.0%
Do research/work at lower costs	50.0%	62.5%	63.8%	47.3%	66.4%	50.7%
Do more accurate, higher quality research/work	39.6%	66.5%	67.3%	42.9%	75.2%	42.3%
Access resources for my research/work faster or better	62.0%	80.2%	73.7%	82.8%	86.8%	63.5%
Produce, process or analyse data faster and better	56.3%	71.2%	71.4%	57.4%	72.2%	58.0%
Work on problems not addressed before	68.0%	74.3%	78.2%	56.9%	77.8%	61.6%
Have more publications or conference proceedings accepted	27.9%	43.7%	49.0%	14.6%	48.5%	26.2%

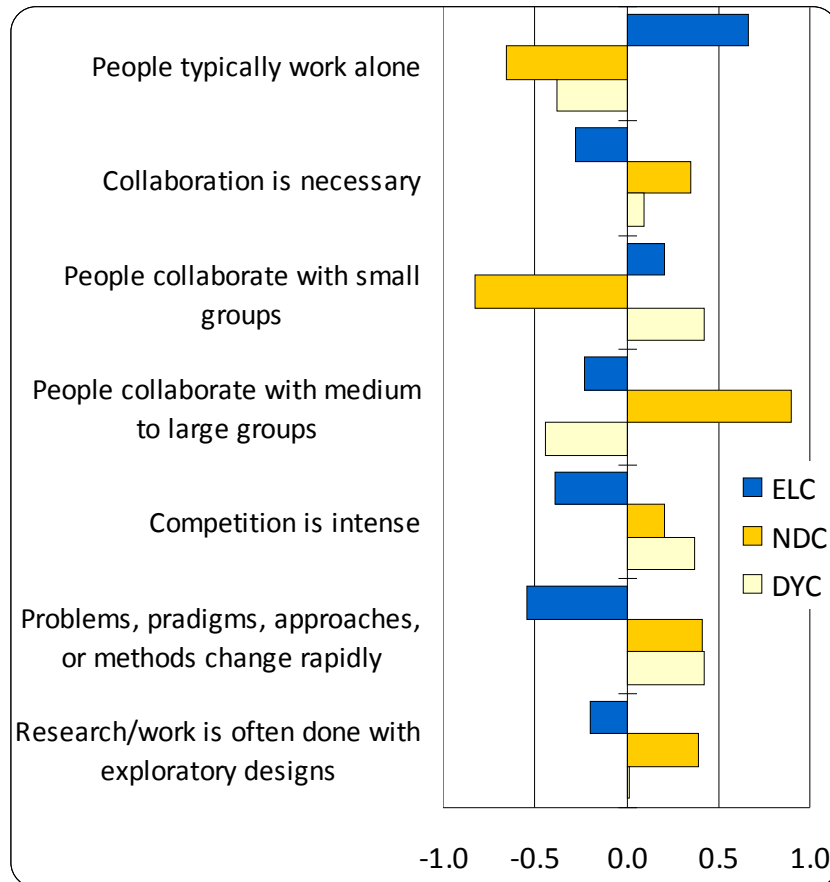
Impact by type of field

Three clusters of fields

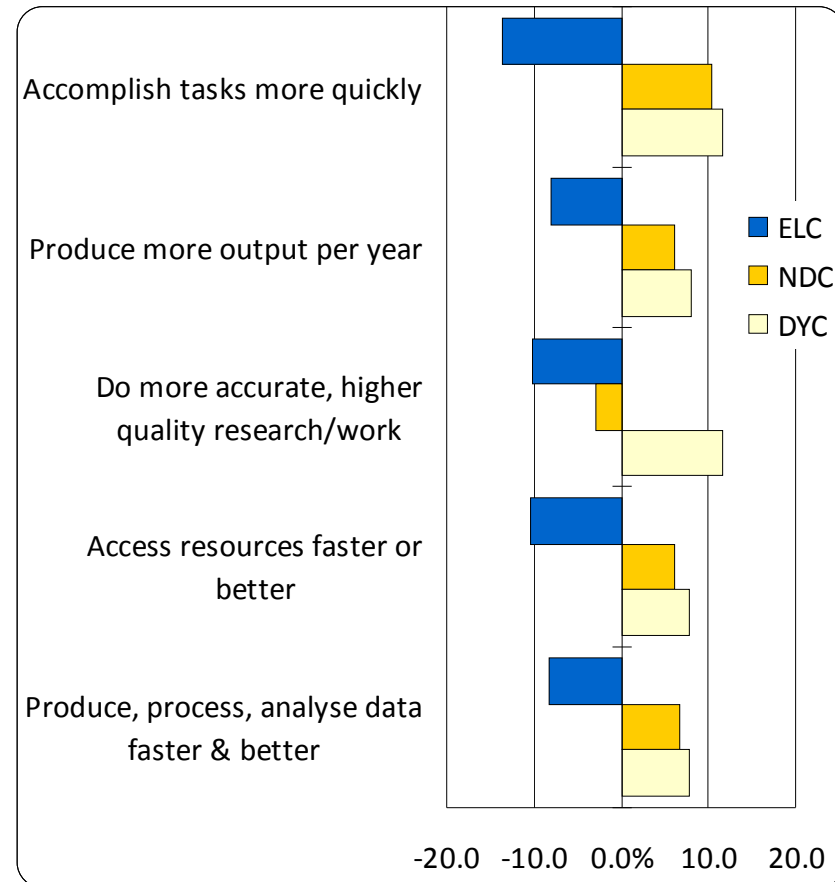
ELC: Established low collaboration fields

NDC: Novel dynamic collaborative fields

DYC: Dynamic competitive fields

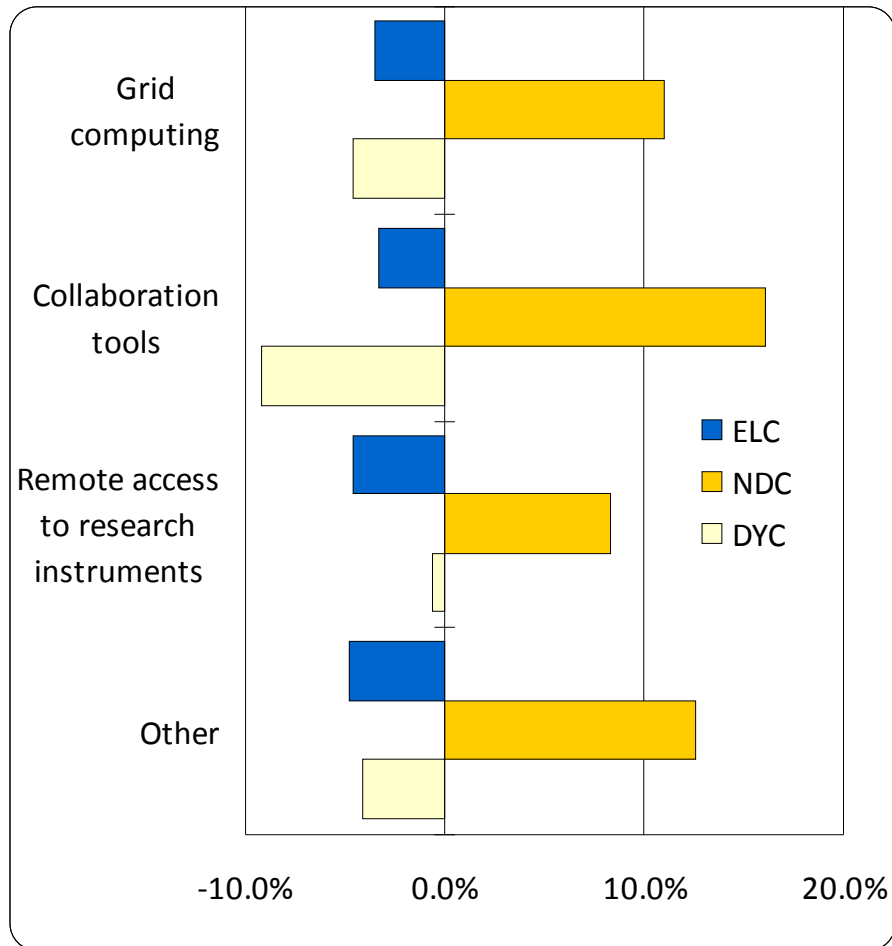


Impact ratings per field cluster (diff. to total in %)

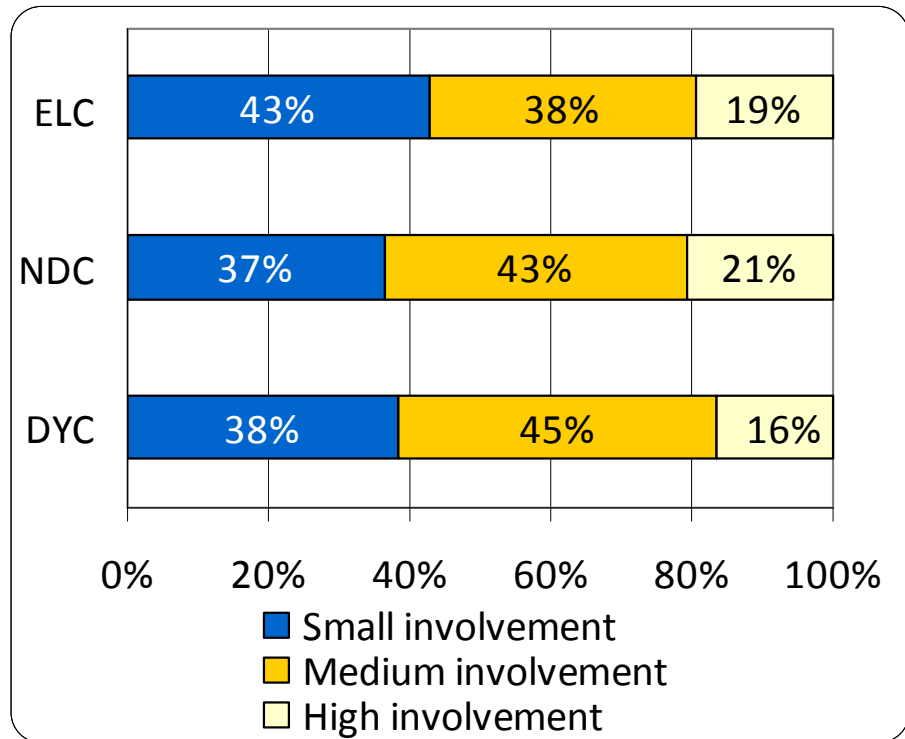


Type of field and involvement

Use of services/resources

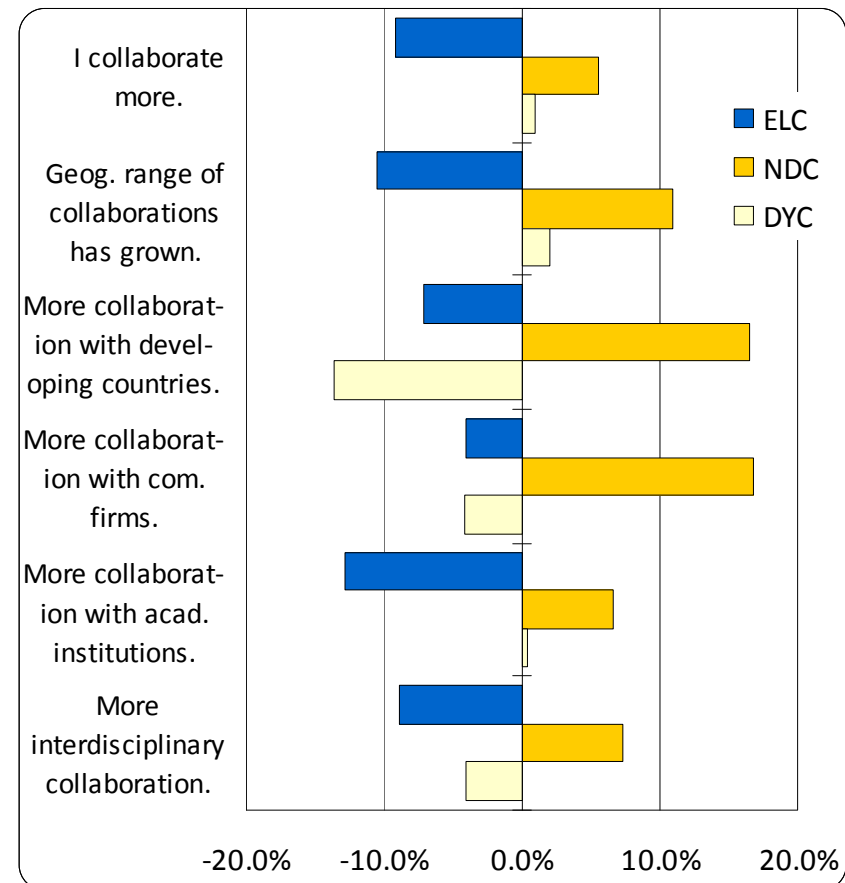
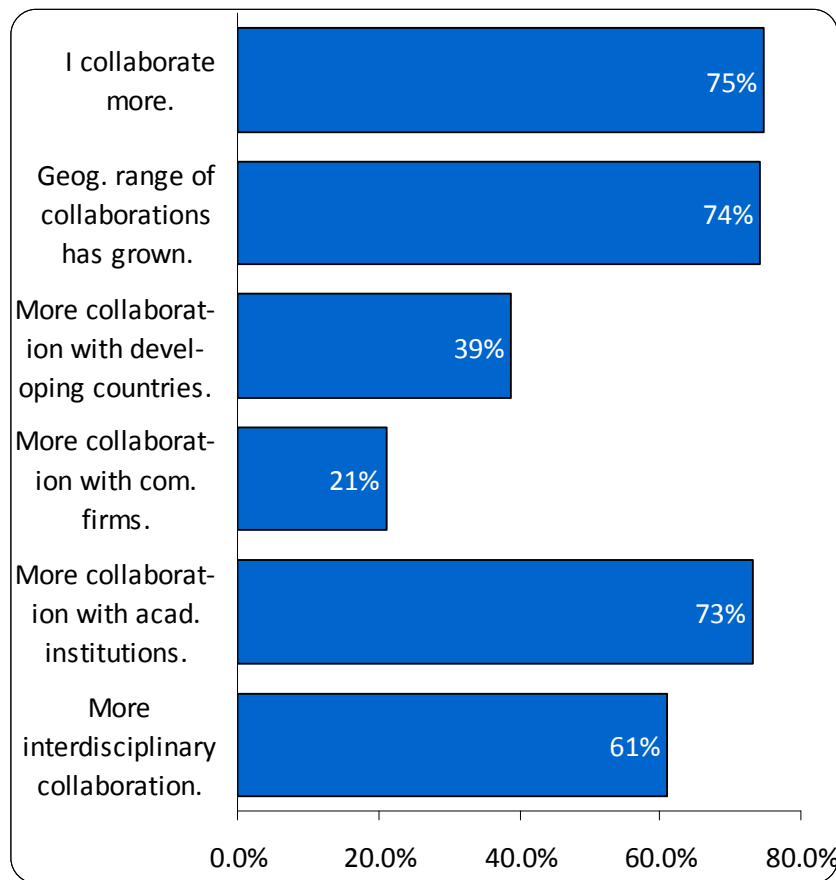


Intensity of involvement



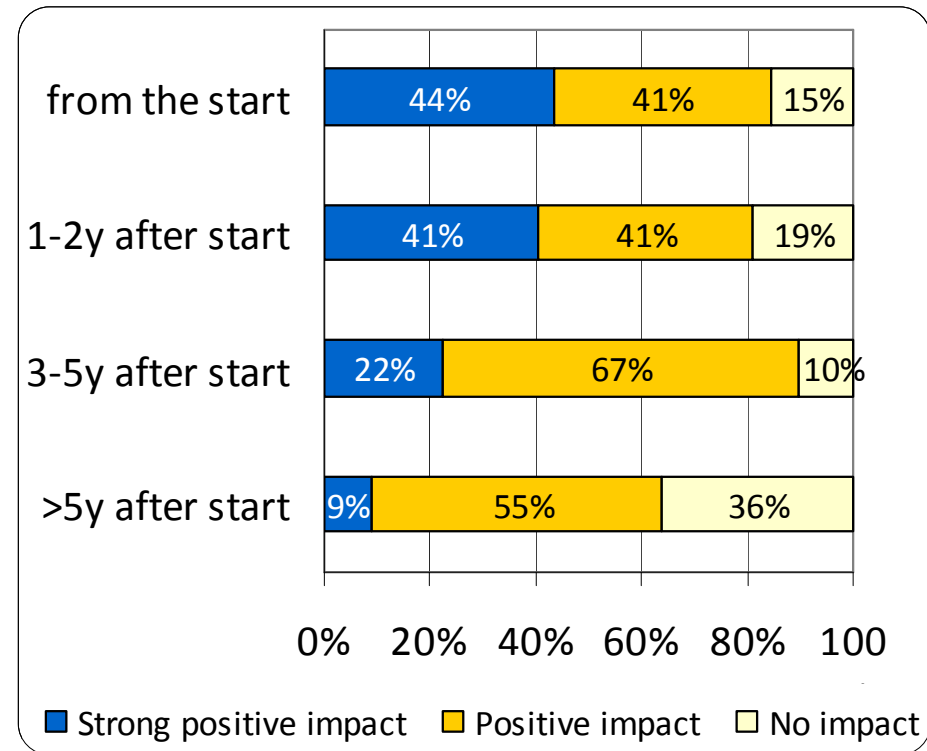
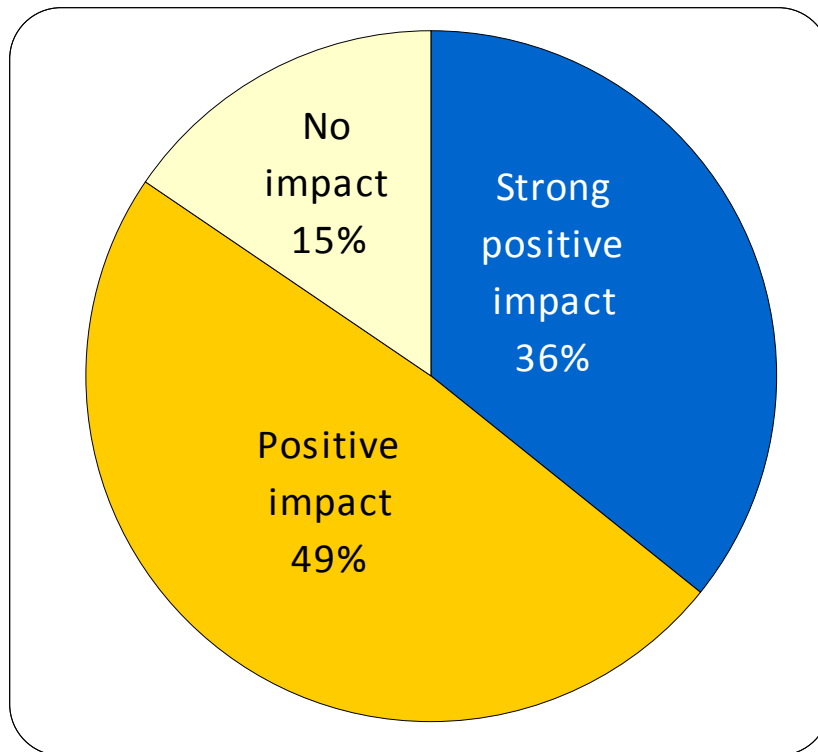
Statements on impact of working with one particular e-Infrastructure on collaboration

Ratings per field cluster (diff. to total in %)



Impact clusters

by involvement
at/after project start



Conclusions

1. The included e-Infrastructure projects contribute to technology development, community-building and scientific progress, but not (yet) commercial success.
2. e-Infrastructures are more often than not considered as a valuable enhancement to scientific research and collaboration.
3. We do not yet have more than anecdotal and non-representative evidence on how e-Infrastructures are defined, adopted and adapted in different fields.
4. Future studies should...
 - ...find out what ICT/e-Infrastructures researchers in different communities currently use and would need for their research,
 - ...explore the link between social and technological change in research further,
 - ...assemble a representative data basis on technology-enhanced research for several fields across different domains.

Categorization of e-Infrastructures

	Geographic scope	Disciplinary scope	Type of service	Driver
C3-Grid	National	Multidisciplinary	Data	Community
CineGrid	International	Disciplinary	Data	Community
CLARIN	International	Disciplinary	Data	Community
D4Science	International	Disciplinary	Data	Developers
DARIAH	International	Disciplinary	Data	Developers
DEISA	International	Multidisciplinary	Computing	Developers
DRIVER	International	Multidisciplinary	Data	Developers
EELA-2	International	Multidisciplinary	Computing	Developers
EGEE	International	Multidisciplinary	Computing	Community
ETSF	International	Disciplinary	Data	Developers
GEANT	International	Multidisciplinary	Computing	Developers
MediGrid	National	Disciplinary	Computing	Developers
NVO	National	Disciplinary	Data	Community
OGF	International	Disciplinary	Data	Community
OSG	National	Multidisciplinary	Computing	Community
SND	National	Multidisciplinary	Data	Developers
Swiss Biogrid	National	Disciplinary	Computing	Community
TeraGrid	National	Multidisciplinary	Computing	Developers

eResearch 2020

**The role of e-Infrastructures in the creation of
global virtual research communities**



Thank you!