TREBALL FINAL DE MÀSTER
EN SOCIETAT DE LA INFORMACIÓ I EL CONEIXEMENT

Knowledge production in e-Research

Autora: Pilar Callau Usón
Director: Agustí Canals Parera
Juny 2011
Universitat Oberta de Catalunya

Abstract

In recent years, there has been a fast increase in is the use of shared networked computational resources for research, generally referred to as e-Research. E-Research claims to have important implications for the nature of scientific practice and, therefore, for scientific knowledge creation.

This work considers some cross-cutting parts of e-Research in order to analyse the ways in which the production and management of knowledge is being transformed by actual e-Research practices and how e-Research projects create cores of durable networks in knowledge production. A theoretical framework for the analysis is proposed, based on three dimensions: major drivers and limitations of e-Research, impact of e-Research key characteristics in the knowledge production process and analysis of the networked nature of e-Research practices. An assessment of these different dimensions of the changes in knowledge production processes is done through a depth analysis of a set of interviews with participants in several European initiatives of e-Research, allowing us to approach actual e-Research practices within different research fields.

-

Resum

En els últims anys hi ha hagut un ràpid augment de l'ús de recursos computacionals compartits en xarxa per a la recerca, al que sovint es refereix com a e-Recerca. S'afirma que l'e-Recerca que té profundes implicacions per a la naturalesa de la pràctica científica i, per tant, en la creació de coneixement científic.

En aquest treball es consideren alguns temes transversals de l'e-Recerca per a analitzar les formes en què la creació i gestió de coneixement es veu transformat per les pràctiques de l'e-Recerca actuals i com els projectes d'e-Recerca formen nuclis de xarxes duradores en la producció de coneixement. Per a tal anàlisi es proposa un marc teòric basat en tres dimensions: principals factors conductors i limitadors de l'e-Recerca, l'impacte de les característiques principals de l'e-Recerca en el procés de producció de coneixement, i l'anàlisi de la naturalesa enxarxa de les pràctiques d'e-Recerca. Una avaluació d'aquestes dimensions dels canvis en els processos de producció de coneixement es realitza a través de l'anàlisi en profunditat d'una sèrie d'entrevistes amb els participants de diverses iniciatives europees en e-Recerca, permetent així una aproximació a les practiques d'e-Recerca actuals en diferents disciplines.

Table of Contents

1 INTRODUCTION

In recent years, there has been a fast increase in the use of the Internet and other computational resources in research. One area that has become particularly prominent is the use of shared networked computational resources for research, or e-Science, what claims to have important implications for the nature of scientific practice and, therefore, for scientific knowledge creation.

E-Science is generally defined as the combination of three different developments in research:

- the sharing of computational resources,
- distributed access to massive datasets,
- the use of digital platforms for collaboration and communication (Wouters, 2006).

More recently, the term 'e-Research' has become more widespread to avoid the restrictive character of the label e-Science and to include other non-experimental fields as e-Social Science or e-Humanities as well. In this work, the term e-Research will be used as a generic label to encompass research under all labels mentioned.

The core idea of the e-Research vision is that knowledge production will be enhanced by combining networks of human expertise, networks of data and other resources, and shared computational and visualization tools. One of the main characteristics of e-Research is that it fosters research collaboration at a global level, involving interdisciplinary and geographically distributed research groups. This means that e-Research tends to be more complex than other research forms, creating more and more complex links between disciplines, institutions and technology than traditional research. Both the interdisciplinary collaboration and the large scale resources sharing provoke a transformation of research modes, transforming the way knowledge is created, distributed and accessed.

E-Research is mainly based in large e-Infrastructures, which can be understood from a two points of view: as the components of the technological systems that form the electronic networks used, and as the organizational networks that arise and collaborate on this technological infrastructures. The analysis of the implications of e-Research in the transformation of knowledge production, access and diffusion should be, therefore, based in the organizational study of these socio-technical networks.

Although research in this subject is still limited and the world of e-Research is continuously moving, there have been recently several studies examining different aspects of how e-Research is modifying research and knowledge production.

In his works from 2007 and 2008, Schroeder address the potential of e-Research regarding its capacity to create global knowledge, analysing several factors as the geopolitical challenge that the globalizing ambition of e-Research implies. According to him, it is not that e-Research revolutionizes the way in which research is conducted or knowledge production, but it amplifies the way in which research is done, if certain limitations are overcome.

The main limitation detected in e-Research is thus, the limited attention space. E-Research infrastructures must dominate the attention space in order to obtain the competitive advantage required in the market of symbolic goods, which is in this case scientific knowledge (Schroeder, 2007). Another limitation, related to this one, is the competition for obtaining resources to support and fund e-Research. By being an emergent infrastructure, e-Research, as any other scientific area, has its own momentum, and the impact that it has in a future depend on actors placed at different levels: research policies, technical capacities and organisational capacity (Schroeder, 2008).

From a very different point of view, other studies address the phenomenon of e-Research focusing on the organisational changes that it carries, both for the increased inter-disciplinariety and for the need of inter-institutional cooperation or the management of the resources used in concrete initiatives. In e-Research projects organizational and managerial issues are often more important than technical issues (Lawrence, 2006; Meyer & Schroeder, 2009).

In the past, generally, research was usually localized within departments and individual research institutes, and within particular scientific fields. e-Research facilitates, by definition, new forms of collaboration linking organisations across institutions and research groups. It does imply, therefore, distributed collaboration and involves a number of scientific disciplines. The barriers of different disciplinary domains are moving, though, in different forms with e-Research (Hine, 2006). Meyer & Schroeder (2009) measure e-Research from a quantitative perspective, making special emphasis in scientific publications and the interdisciplinariety level found in them depending on the research field. Their work shows that disciplines with an high degree of mutual dependence and a low level of uncertainty as high energy physics are more likely to produce and use e-Research resources than the field with low degree of mutual dependence and high uncertainty as may be cultural or social geography. Disciplinary differences are therefore evident when it comes to the velocity of the processes of e-Research adoption. The characteristics of the diverse scientific fields do have an influence in the production and use of digital resources, and as a consequence, in the shift towards e-Research these fields suffer.

It is however important to be careful with this statements, inasmuch as other works argue that the 'structure of the work' as size, distance, interdependencies or scientific competition, are more problematic for the collaboration of scientific teams than the mixture of backgrounds (Walsh & Maloney, 2007). So, a common point of much of the literature about e-Research is about the complexity of its social structure. The case studies that focused on the coordination of large-scale e-Research projects conducted until now demonstrate the technical and organizational obstacles of the process of adoption of e-Research that these projects face. Often, the literature about this question is presented in the form of case studies of some multi-disciplinar and multi-institutional e-Research project, and analyse the organisation of the project, its logistics, interactions and communication methods. Generally, as it is the case of Lawrence (2006), these studies simply reflect the basic challenges of collaboration and management: means of communication, decision-making mechanisms, trust management, etc. Results show that distant collaboration require a considerable overload of time dedicated to coordinate the logistics and other daily issues that collaboration entails (Cummings & Kiesler, 2005).

Still, so far there is no overview covering the extent to which e-Research is transforming scientific research. An exception to this if the work of Wouters (2006), in which the characteristics of e-Research are analysed according to three models of knowledge creation, as well as the implications of large-scale collaboration in actual research practices. Using the notion of epistemic cultures introduced by Knorr-Cetina (Knorr-Cetina, 1999), Wouters underlines the existing tension between research practices actually constructed over the new technological infrastructures and the ambition of e-Research of being the new paradigm for knowledge creation (Wouters & Beaulieu, 2006).

Wouters' work is interesting in the measure that it introduces, in a purely theoretical manner, the challenges posed by e-Research to actual scientific practices (local), and it distinguishes between generic characteristics and the specific characteristics. E-Research generic characteristics emphasize the potential of e-Research as a technological infrastructure that eases collaborative research at a global level, while field-specific characteristics underline the local character of the scientific and technological puzzle (Wouters & Beaulieu, 2006).

When studying the challenges posed by e-Research to actual research practices – i.e. local practices -, the focus can be put in the generic characteristics of e-Research or on the specific characteristics that can be found in concrete e-Research fields or even projects. However, addressing the analysis of e-Research from the perspective of one or the other set of characteristics influences the vision that can be extracted of it (Wouters & Beaulieu, 2006). The intention of this work, is to analyse e-Research by establishing a framework encompassing both perspectives: considering e-Research socio-technical networks in a global manner, and studying its specificities along the three kinds of knowledge production that Wouters identifies in e-Research projects – knowledge production based on computational modelling, knowledge production through data-comparative research, and knowledge production through resource sharing networks.

This work will consider some cross-cutting parts of e-Research in order to analyse the ways in which knowledge, its production and management, is being transformed by actual e-Research practices. The visibility of research results and, more generally, the production of knowledge can be seen, beyond its quantitative measures, from the perspective of the factors that boost, obstruct or structure the knowledge creation processes. We will focus on the analysis of how and in what ways the socio-technical networks in which e-Research is based provide a conduit for knowledge (and what kind of knowledge).

At a first stage, the theoretical basis for a cross-cutting analysis of e-Research are set. First, the major drivers and limitations of e-Research are analyzed from a global perspective, paying special attention to the role of research policy and to the general discourse accompanying e-Research.

The impact that the key characteristics of e-Research have in the knowledge production process and how research practices are affected by them are then analyzed in the light of the general scientific research process.

A theoretical framework is proposed next to address the networked nature of e-Research practices, considering the notion of *Ba* applied to research contexts (Hautala, 2011). This framework focuses in the ways in which e-Research projects create cores of durable networks in knowledge production, and are able to maintain them, connecting research groups, projects, institutions, technological infrastructure, funding mechanisms and discourse.

At a second stage, a group of participants in several European initiatives of e-Research are approached by the means of in-depths interviews, in order to assess the different dimensions of the changes in knowledge production processes analysed previously. The expertise of researchers interviewed based on the different initiatives in which they are involved allows to approach the networked nature of research practices in e-Research for different modes of knowledge production.

2 DEFINING e-RESEARCH AND ITS CHARACTERISTICS

The term e-Research refers to research activities, not uniquely scientific research but also other disciplines, including humanities and social sciences, which use advanced information and communications technologies (ICT) capabilities to support existing and new forms of research.

E-Research is generally defined as the combination of three different developments in research (Wouters, 2006): the sharing of large computational resources and high performance visualization capabilities, distributed access to massive datasets and the use of digital platforms for collaboration and communication. Thus, by its very nature, it facilitates and fosters research collaboration of multi-disciplinary and inter-disciplinary research both nationally and internationally.

E-Research is for a big part about creating new communication, computing and information gathering tools, however e-Research is not only about computing and networked technological infrastructures, but also about distant collaboration and an increased social scale of research. Computational power and capabilities can be considered the key intellectual challenge driving e-Research, aiming to solve two detected needs in research: the arise of more complex problems and the need of handling more data. These two problems are closely linked to the need to coordinate resources for large research projects at a global scale in order to meet higher research goals (Wouters & Beaulieu, 2006).

Therefore, the infrastructures sustaining e-Research, or e-Infrastructure, can be seen as a large technological system consisting of a number of interdependent social and technical systemic parts (Schroeder, 2007). It can be considered, in the first instance, as the physical components of the technological system: the advanced electronic networks and computing tools that make use of the Internet and the Web; and as the organizational networks that support, implement, sustain and are supported by this system. E-Research can be thus considered as a socio-technical system. Technologies and social institutions (organizational capabilities) are intertwined in e-Research, which depend as much on organization capabilities as on technical ones (Meyer & Schroeder, 2009; Lawrence, 2006).

The combination of the use of advanced computing power and tools and the increased creation of research collaborations internationally that characterize e-Research, transforms the way research is conducted, expanding the scale of research. It is usually argued that the 'e' of e-Research stands for 'enhanced' more than 'electronic', as it enhances research practices within different disciplines rather than replacing traditional practices.

The range of e-Research activities is diverse and multi-disciplinary. E-Research methodologies and capabilities can be applied in all research disciplines, not just those with a traditional affiliation with high performance computing as physics or biotechnology, and enables the emergence of entirely new fields of research that are based in the use of new techniques for data mining and analysis, advanced computational algorithms, and resource sharing networks.

What characterizes e-Research practices across all research fields involved can be reduced to three major attributes: interdisciplinary, inter-institutional and international collaboration, data sharing and access to shared computational resources.

This collaboration and sharing can take place between researchers, between projects, between institutions, between disciplines and between different locations. Thus, it reveals new possibilities for collaboration by not only putting research online, but rather using leading edge computing tools—including shared databases and instruments, tools for distributed work, and shared computing resources—to foster it (Schroeder, 2007). Managing the large scale networked research resources for this kind of projects or initiatives needs for an increased cooperation between all the institutions involved, provoking the organizational and management structure of e-Research projects to often be even more challenging than technological questions (Meyer & Schroeder, 2009).

In traditional research, generally, research is often localized in individual departments and research institutes and within concrete fields. E-Research, by its networked nature, implies geographically distributed collaboration, as well as it involves several scientific disciplines (i.e. at least of two research fields are involved in every e-Research initiative: computer science and one other field). Research practices, therefore, are modified by this collaboration in three different aspects:

- Research data and results are shared in an open way between groups and researchers involved in a project by means of the e-infrastructures, instead of being guarded carefully for individual research purposes.
- Adaptation is needed towards more standardized ways of working instead of the idiosyncratic behavior usually present in research practices.

  However, the barriers separating different disciplinary domains are being modified in different forms (Hine, 2006). The epistemic characteristics of different scientific fields influence in the production and forms of use of networked resources and thus, e-Research methodologies are adopted at different velocities and adapted to particular specificities (Meyer & Shroeder, 2009). Not only disciplinary differences influence e-Research practices, but also do the forms in which work is structured. Several factors of the work structure as the project size, participants' distance, interdependencies or scientific competition are often more challenging for collaborative research practices than the degree of inter-disciplinariety (Walsh & Maloney, 2007).

- Supporting distant and inter-institutional collaboration requires an overload of communication.

  One layer of the e-infrastructures is the communication layer. As the efforts dedicated to e-Research become more structured from an organizational point of view, the volume of e-mail communication and meetings necessary to coordinate the logistics, decision-making mechanisms and other organizational activities increase (Cummings & Kiesler, 2005). Scientifics involved in those distributed and collaborative efforts dedicate a major part of their time to communication and organizational activities uniquely to create and sustain this collaboration (Meyer & Schroeder, 2009; Lawrence, 2006).

These transformations across disciplines in research practices entail a change is the way knowledge is created, distributed and accessed. Nevertheless, e-Research ambition of transforming knowledge is not uniquely driven by research communities but also by an increased number of national and international funding programs and initiatives dedicated to the creation and expansion of such e-infrastructures, the ethos of science and the own dynamics of these large technological systems, as will be discussed in the following.

2.1 E-Research and Research Policy

In recent years, there have been a number of initiatives to support collaborative research via electronic networks and to support the development of the large technological infrastructures in support of research fields that require large-scale data processing. E-Research, and more broadly e-Science, has become a buzzword for these initiatives, but these have also gone under other labels as cyberinfrastructure in the US, e-infrastructure in the EU, or the Grid, as we will see.

The dominant discourse about these initiatives, and e-Research in general, is about the way the forms of doing research will be revolutionized by it. These revolutionary transformations are expected to take place as a consequence of the adoption of powerful large technological infrastructures in research practices, creating new forms of knowledge, fostering the creation of globally distributed research communities, and boosting the share of research resources between large interdisciplinary and distributed research groups. Thus, we can say that e-Research is, in a first stage, a discursive construction mixing technical and scientific practices, computer technology design and science policy (Wouters, 2006).

Nevertheless, e-Research has not been defined in a strict way. The proponents of e-Research describe what e-Research is, or will be, by pointing to individual examples and case studies (Lawrence, 2006; Schroeder, 2008). That is the reason for several authors (Wouters & Beaulieu, 2006) to point out that e-Research is a particularly interesting case for the sociology of expectations. The more sceptic school stands that the main characteristic of e-Research actual practices is little more than the creation of a promise (Wouters, 2006). What it is true is that the shaping of e-Research is, at least partially, based on the expectations about the capabilities of the technologies that enable it.

The Grid technology developed by computer scientists working on particle physics settings played an important role for e-infrastructures. The central question in this field was moving from sharing information via the Internet to sharing computer power and data storage capacity via the Grid, turning "*the global network of computers into one vast computational resource*" (CERN, 2005). Indeed, the dream about the Grid is not hidden by its proponents:

> *"... That is the dream. But the reality is that today, the Grid is a 'work in progress', with the underlying technology still in a prototype phase (...). It's future, even if still uncertain, is potentially revolutionary."*
>
> *(CERN, 2005)*

The collaborative nature of e-Research is also reflected in the rhetoric used: "... *They dreamt of being able to collaborate with distant colleagues easily and efficiently, safely sharing with them resources, data, procedures and results*" (CERN, 2005).

The promise of technology capabilities influencing how research will be conducted is important in shaping e-Research development and adoption. The common discourse about the capabilities of e-Research and e-infrastructures can be found along the different national and supranational initiatives and so, the process of defining future e-Research practices includes from the drafting of funding proposals and e-Research programmes, to the creation of demonstrators and pilot projects (Wouters & Beaulieu, 2006).

The early developments of e-Science, can be located in the US at the end of the 1990s in the framework of supercomputer centers with pilot projects in computer science that lead to the emergence of the first modern Grid structure designed to provide access and coordinate resources and other activities, in the framework of the I-WAY project (Wouters & Beaulieu, 2006). This developments where framed in the first US cyberinfrastructure programs created by the US National Science Foundation (NSF). However, US cyberinfrastructure programs have been primarily focused on the development of the technological infrastructure components: computing and communication hardware systems, networking middleware and scientific databases, and therefore mainly driven by computer science and engineering.
Nevertheless, e-Research programmes are actually based on the claim that a number of fields have an increased need of computational resources to be able to deal with complex computational problems, as it is the case of particle physics, astronomy, medical and cognitive science or seismology. The domains that are supposed to be affected by this revolution vary according to the national variety of e-Research. But, so the claim goes, almost all disciplines will be transformed in some way or other.

Albeit computer science research is a core part of every e-Research, e-infrastructure and cyberinfrastructure initiative, other national and international programmes on e-Research as it is the case of European Union initiatives, the UK e-Science programme or, in a more reduced scale, the several initiatives in the Netherlands, contemplate a more broaden perspective of e-Research applications in different fields (Wouters, 2006). The ambitious UK e-Science programme, started in 2001 with a funding of 250 million pounds for a 5-years programme (Hey & Trefethen, 2002), funding over 100 pilot projects along each research council in their own application areas ranging from basic science to engineering[1]. Even if its first project goal was to identify the generic middleware[2] requirements arising from those e-Science pilot projects, from 2005, the UK e-Science program has taken the initiative to spread the reach of e-Research to content-rich domains as the social sciences and humanities[1].

A similar evolution can be found in European initiatives from 2004 until now. The Research Infrastructures programme[3] of the European Community consists actually of several parts, ranging from facing the technical challenges and developing more powerful technological components to e-infrastructures, to providing support to existing research infrastructures or supporting the creation of new research infrastructures. Added to these, the EU programme is also funding initiatives for supporting policy development and the programme implementation itself.

Regardless of the different labels and forms that different initiatives take, research policy documents and agendas across countries are in agreement about the general e-Research ambition and e-Research occupies an important place in them[4], with remarkable similarities in its structure and form.

However, the different actors and components that are determining e-Research infrastructures are located on different levels. There are a number of factors apart from research policy that shape e-Research, as they are the technical possibilities or the organisational capacity (Shroeder, 2008). The future of e-Research will only cohere in the overlapping of research policy agendas, the agenda of the organizations involved, the ethos of science and epistemic cultures involved, and the dynamics of the

1 UK e-Science program: http://www.epsrc.ac.uk/about/progs/rii/escience/Pages/default.aspx
2 Generally referred to as the Grid
3 European Commission Research Infrastructures program: http://ec.europa.eu/research/infrastructures/index_en.cfm
4 In the US, for example, there is now a separate *Office of the Cyberinfrastructure* within the NSF. In the UK a National e-Science Centre has been established together with eight other e-Science Centers. At European level, several e-Infrastructure fora were established, the European Strategy Forum on Research Infrastructures (ESFRI) and the e-Infrastructure Reflection Group (e-IRG) being the most prominent.

large technological systems being developed. And for doing so, greater organizational coherence will be needed among the various actors and groups, as well as having a deep understanding of the main limitations that e-Research's different components have in constructing the e-Research dream.

2.2 E-Research limitations

The scope of the e-Research, the claimed revolutionary changes in the ways research is conducted and knowledge is created depend heavily in whether it is able to overcome its own limitations.

The global scope of e-Research

There are several on-going debates within science policy about the best way of organizing e-infrastructures support in the context of increasingly complex scientific goals, entailing interdisciplinary collaboration while also being distributed among different institutions and geographical locations.

Most research in the e-infrastructures initiatives aims to construct an infrastructure that is supposed to support all forms of e-Research in the future, meaning not only quantitative research or computational oriented research. The ambition of universality, however, cannot be reached unless this(ese) infrastructure(s) are not shared globally between researchers, institutions and research projects around the world. Indeed, the contributions towards extending the global scope of research, reshaping disciplines and developing research technologies are mutually reinforcing (Schroeder, 2008). Yet at the same time, in each case there are also limits.

The globalizing ambition of e-Research entails, to truly create global knowledge in a cooperative manner, an important geopolitical challenge (Schroeder, 2008). An international network of tools, databases or computer applications require, inevitably, the establishment of access coordination policies and the standardization of processes at a global level.

In its analysis of the interrelation of different disciplines and the way knowledge is globalized of three e-Science projects, Schroeder (2008) underline the importance of driving effort towards a common infrastructure between the principal national and international e-Research programmes. The core problem in which most e-Research projects have been focusing is the establishment of common standards in which to share the data in order for researchers to access them uniformly. However, this kind of efforts will be repeated again and again in future projects until global standardisation of e-Research processes is set up. Important efforts should be dedicated to achieve compatibility between the European infrastructure and other large infrastructure projects as TeraGrid[5] or OGF[6] in the US, or the SIENA initiative[7] in Europe. By being compatible we mean setting standards, data sharing policies and not less importantly, establishing strong organizational links. Openness of e-infrastructures would be a requirement for this compatibility (Schroeder, 2007). This means that, at this level, geopolitics play an essential role, as it involves both (scientific) competition and cooperation, providing e-Research projects a global scope (Schroeder, 2008).

Nevertheless, an important limitation to the universality of e-Research is that the main forms that e-

[5] NSF funded Project TeraGrid to build and deploy a large distributed infrastructure for open scientific research: https://www.teragrid.org/

[6] Open Grid Forum: http://www.ogf.org/

[7] SIENA - Standards and Interoperability for eInfrastructure Implementation Initiative: http://www.sienainitiative.eu/

Research takes in practice are not global, but are not totally local in the sense that are not based in individual projects either. The main forms of e-Research are rather national or supranational (in the case of the EU), focused on the resources funded by large-scale research initiatives (Schroeder, 2008).

The competition for resources

In the view of that, a second limitation that has been broadly signalled by e-Research literature is the competition for resources for supporting and funding e-Research projects. This limitation is closely related to the limited attention space. One of the consequences of the universality of scientific knowledge is that, for a particular phenomenon or domain, there is only one vanguard in research. The main constraint in science is, therefore, the existence of a limited attention space (Meyer & Schroeder, 2009). This means that, apart from the intra-scientific priorities (competition to publish research results in the first place), there is also a non-scientific factor (public attention and funding for research) that constitutes the broader context acting as a selective mechanism on how research is prioritized. E-Research has to compete for both kinds of attention for its further development.

E-Research infrastructures must dominate the attention space to obtain competitive advantage in the market for symbolic goods or tools; in this case, scientific knowledge or instruments in their various forms. "*The major gain in academia or innovative research (including tools) in these cases may come from attention rather than economic gain; and second, that building the capacity or the expertise to develop these findings or data or tools within a group is itself part of the gain, including (again) the attention received by this work which secures further talent and/or resources*" (Schroeder, 2007).

This competition for the attention space has two sides in the case of e-Research. Focusing on the development of the large infrastructures and databases in long term, the competition is related to the globalizing ambition of e-Research perspective, in the sense that for the achievement of this ambition, there will be in the future only one or a few dominant infrastructures and databases, and similarly will happen with tools and types of software.

On the other hand, in academic research, it is possible to have increasing rates or volumes of knowledge production and increasing specialization of research outputs coming from e-Research projects. The limitation is nevertheless the focus of attention on a few research results, and this effect persists in e-Research (Schroeder, 2007). It is important to underline, however, that the competition of the attention space not only takes place within the academic community for the visibility of research results, but also at it has been mentioned previously, for competing for research funding and support. This later competition has an important effect in the direction that many research projects take. Competing for research funding implies some kind of adaptation to research policy agendas. This is specially the case of e-Research projects, which by its inter-institutional and interdisciplinary collaborative nature need for largely funded projects (Beaulieu & Wouters, 2009). One of the so-called advantages of agenda-driven research activities is that they serve the wider community, supporting societal challenges. However, there is a real risk that by focusing more on a winning grant-style programme research could be dramatically reduced or fragmented in largely isolated islands of knowledge. It is important to underline that while certain types of research can be strategically driven, others will require a researcher-oriented approach.

The tension between the undefined generic ambition of the e-research dream and the particular pilot projects that embody the dream by means of national and supranational e-Research initiatives and programmes is not an accidental phenomenon but is the very consequence of the way e-Science is constructed as a future mode of scientific research (Wouters, 2006). The interest in the development of e-infrastructures that exist either at national as at international levels is translated in a broad interest in expanding the realm of e-Research, and as so, it is reflected in actual research policy agendas. As a consequence, there seems to be no field of research to which e-Research might not be relevant (Wouters, 2006). In this context, a correct balance between agenda-driven activities and curiosity-driven research seems difficult to reach.

The adoption of technological systems in specific scientific fields

The future of technological research infrastructures is not determined by the extrapolation of their technical possibilities (Wouters, 2006). These technologies cannot be assumed to be extensible without taking into account their limits.

E-Research exemplifies this in various forms. First, the ability of creating new tools does not necessarily mean that a closer relationship between knowledge and technologies will arise. And second, new technologies, in research as well as in any other field, tend to add and complement, rather than to substitute previous technologies or practices related to them (Schroeder, 2008). This leads us to an important point when dealing with e-Research: despite being at the centre of attention of research policy agendas, e-Research is nothing more than an area of knowledge production among other.

What is clear is that e-infrastructures development is concentrated at the leading edge of scientific advance, and this, apart from being an important aspect of the competition for the attention space, also constitutes a limit by itself. Whether scientific disciplines are closely coupled and mutually dependent, and to what extent they are driven by research instruments or machines play one of the most important roles in the processes of adoption of e-Research. Disciplines with an high degree of mutual dependence and a low level of uncertainty as high energy physics are more likely to produce and use e-Research resources than the field with low degree of mutual dependence and high uncertainty as may be cultural or social geography (Meyer & Schroeder, 2009).

However, this should not be misunderstood, there will also be impacts on scientific areas of advance which have low degree of mutual dependence and high uncertainty, but they will face with particular constraints in their intellectual and social organization of these sciences. As we will see later on, the characteristics of the diverse scientific fields have a strong influence in the production and use of digital resources and, as a consequence, in the shift towards e-Research that this fields suffer.

E-Research combines changes at local and global scales, but it is an on-going effort. By now, e-Research is just at the beginning of providing generic devices and tools, and the same happens with providing a common language to both work with and to talk about it.

Thus, being part of an emergent infrastructure, e-Research, as any other major trend in Science, has an own momentum. And, as we have seen, its main limitations are the standardisation of the technological systems and of work, the competition for the attention space and for obtaining resources to fund and support its future development, and the meeting of technical possibilities with scientific epistemic cultures.

The question that arises is whether the frontier between e-infrastructures and other, more traditional, research resources will disappear or not (Schroeder, 2008). In other words, if once e-Research becomes of common use, a differentiated e-Research will disappear as a concept, integrating in dominant research, as it has happened with other changes that research has experienced historically.

3 KNOWLEDGE PRODUCTION IN E-RESEARCH

E-Research and the large technological system that supports it are deemed to be essential for creating what is potentially a new system of knowledge production (Schroeder, 2007). From a less ambitious perspective, e-Research is claimed to be changing the ways in which knowledge is created and managed. While it cannot be denied that the communication layer of e-Research changes communication practices in science, the ways in which e-Research benefits research and/or transforms the general structures of scientific research processes are not deeply studied.

In order to approach how new research practices arising with e-Research are influencing the process of knowledge production, we must understand how the scientific community 'traditionally' produces and manages knowledge.

3.1 Research and knowledge production

"Research can be defined as the search for knowledge, or as any systematic investigation (...) to establish novel facts, solve new or existing problems, prove new ideas, or develop new theories, usually using a scientific method."
"Research", *Wikipedia*[8]

"Science is a systematically organized body of knowledge on a particular subject."
"Science", *Oxford Dictionaries Online*[9]

The goal of the research process is to produce new knowledge. The research process is generally structured in several steps, normally being: 1) field study and hypothesis setting, 2) conceptual and operational definitions, 3) collecting of data, 4) analysis of data, 5) testing and revising hypothesis, and 6) drawing conclusions.

We can identify in this process, three main forms of research, corresponding to three different phases of the research process, but that can also be considered individually. First, we find the structuring and identification of new research problems – exploratory research -; constructive research is built over the results of exploratory research, i.e. developing solutions to the identified problems; and last, empirical research tests the feasibility of a solution using empirical evidence[7].

Moreover, research can also be classified by two more broad types depending in what kind of data it is based. Primary research, thus, is the research dedicated to the collection of non-existent data or data not yet collected or used previously, while it is called secondary research to the research that summarizes, compares and/or synthetizes existing research[7].

From the point of view of epistemology, Science - understanding it as knowledge of any kind, from latin: *scientia* meaning "knowledge" - has a high degree of systematicity. This systematicity is manifested in six aspects of science:

- how it describes,
- how it explains,
- how it establishes knowledge claims,

[8] "Research", *Wikipedia, The Free Encyclopedia.* Retrieved 17 May 2011. [http://en.wikipedia.org/wiki/Research]
[9] "Science", *Oxford Dictionaries Online*. Retrieved 17 May 2011. [http://oxforddictionaries.com/definition/science]

- how it aims for completeness,
- how it expands knowledge and
- how it represents knowledge.

On the other hand, social order in Science is seen, in the first place, as emerging from the scientific ethos, i.e. from a set of norms which govern all scientists' actions and interactions. Merton (Merton, 1942) proposed that the 'ethos of modern science' consists of four sets of institutional imperatives: universalism, communalism, disinterestedness, organized scepticism. Scientific communities are marked by these norms and therefore so it is supposed to do the production of knowledge.

All four norms are interrelated and mark the research process. However, as we will see further on, 'universalism' and 'communism' are especially relevant in the context of e-Research discourse.

The production of scientific knowledge has been traditionally regarded as a collective enterprise, where researchers contribute to a common product: scientific knowledge. Although often the production of new knowledge is located on the individual (Hautala, 2011), scientific knowledge production is not an individual activity, even in the case of non-collaborative research (Gläser, 2003).

Individuals produce new knowledge and circulate it within the scientific community, who decides about validity and recognition for new knowledge claims (Merton's 'organised scepticism'). Through this perspective, the incentives for contributing new knowledge are either found in an intrinsic motivation to advance knowledge (Merton's 'disinterestedness'), or are defined via an exchange theory which assumes that scientists exchange new knowledge for some kind of recognition by references to their work. However, research cannot be done without drawing upon the work of other researchers, what in most cases requires collaborating with others (Merton's 'communalism'). The publication of new knowledge claims, for instance, involves not only the author but also trusted assessors, editors, and reviewers, who all influence the content of the publication.

Scientific communities thus represent a type of collective production which is characterized by decentralized coordination of contributions via local interpretations of the shared subject matter of work (Gläser, 2003). It can be said that scientific communities are knowledge production collectivities where:

- Coordination of actions is achieved by scientific knowledge rather than norms, namely by the community's shared body of knowledge.
- Scientists derive problems and methods of problem solving from their interpretation of the common body of knowledge. It is also important to notice that a part from publications, other important part of the research process is sustained on research materials, protocols, methods, etc., shared by researchers, as well as continuous informal communication between them.
- The integration of new knowledge claims proposed by individual researchers is accomplished by other scientists using the offered new knowledge, i.e. treating it as belonging to the common body of knowledge (Merton's 'universalism').

An important conclusion can be drawn from this, that is that the transmission of material and components (results, hypothesis, data, etc..) between researchers is part of the knowledge production process. Sharing of research material, by means of publication or by any other means, is never separated of previously produced knowledge.

3.2 How e-Research changes the knowledge production process

There are some new elements in how networks of geographically dispersed researchers communicate, collaborate and share data and computational resources via e-infrastructures that imply some changes in the process of producing knowledge. In particular, e-Research sharing of computational and data resources raises new communalization processes that will affect scientific communities depending on the specific mode of knowledge production they use. In the following, we will analyse how the three key distinctive characteristics of e-Research as per Wouters' definition (Wouters, 2006) impact the knowledge production process and affect research practices.

Use of digital platforms for collaboration and communication

One of the most studied impacts of e-Research and ICT-mediated research are related to the communication layer. Generally, these are study cases of e-Research projects that reflect the basic challenges of ICT-mediated collaboration such as the forms of communication, decision-making or trust management (Lawrence, 2006).

As it is the case with personal communication, the internet makes things easier, faster and more global also in e-Research. As discussed when analysing the collaborative nature of e-Research (see Section 4), e-Research requires an overload of communication. However it cannot be claimed that e-Research communication practices had changed the quality of the social order of scientific communities or the way scientific knowledge is produced.

For instance, the informal communication that accompanies the application of experimental methods in research has always been an important part of researchers' everyday work. It represents, indeed, a specific type of collaboration: the exchange of know-how. E-Research practices, although performed collaboratively by several distributed groups and institutions, do not change that. Informal communication is done both by e-mail communication and in face-to-face meetings held to coordinate the work (Lawrence, 2006).

Another aspect of the use of digital platforms for collaboration is that powerful networks of information and resource storage are being created. The idea of resources sharing is not always focused on scientific data in the traditional sense but may also include other type of resources such as publications. Wouters defines the knowledge creation that is based in access to distributed and heterogeneous digital resources as '*knowledge creation as reading in a digital library'* (Wouters, 2006). The way researchers process those resources might be affected fundamentally. However, this modality of e-Research is the one that is affecting less the knowledge production process, as these changes are the most external to the process of knowledge creation.

Distributed access to massive datasets

Research resources, data and results are shared in an open way between groups and researchers involved in e-Research projects, and often with researchers external to the project.

Databases are different to research reports and publications in their main content. Publications' main contents are knowledge claims, although often some data is included to support them. Databases, on the contrary, contain raw data which is not connected to any knowledge claim.

Sharing access to massive datasets enhances research by providing researchers with new independent basis for the construction of new knowledge (Gläser, 2003) without the need of collecting their own data - that would be, with high probability, a smaller sample -. Shared datasets provide a new basis on which knowledge can be established, a part from own data and publications.

On the other hand, since raw data is created in an earlier stage of the knowledge production process – primary research -, their public accessibility means that communalization is occurring in an earlier stage of knowledge production. This has two consequences:

- Quality control mechanisms of published knowledge may be enhanced in several ways, as the accessibility to raw research data will ease detecting selective use of data to support knowledge claims, or will eliminate the individual competitive advantage resulting from exclusive use of unpublished data. And whether quality control mechanisms are or are not using these new capabilities, researchers perception is about a more controlled research process.
- The emergence of new or separate research activities in many scientific fields where it does not yet exist, generally related to the discovery mode.

  We can find, for instance, within this emerging research modes, knowledge creation through comparison of data or comparative research[10]. Thus, the emphasis is not on complicated analyses of data but on relatively simple comparisons across huge data sets of possibly very heterogeneous types (Wouters, 2006).

Hence, sharing of large datasets rise a more rigid divide between primary research (data production) and secondary research practices (data analysis) (Gläser, 2003). Pattern recognition, visualisation, search tools, simulation or modelling are other fields that take advantage of the large-scale (raw) data sharing enabled by e-Research. Some of these practices are closely related with computational tools and applications as we will see next.

Sharing of computational resources

Apart from developing the technological infrastructure itself, an important aspect of e-Research grand vision of enhancing knowledge production is based in the ability of e-Infrastructures to facilitate remote access to research equipment and distributed computational power in the form of operational time.

Nevertheless, it cannot be ignored that the research that profits from computational resources is, therefore, computationally oriented. E-Research is thus changing two types of research. First, some fields as astronomy, medical and cognitive science or biotechnology, that were already confronted to computational problems, can tackle with e-Research more complex computational problems that were not possible to address previously (Wouters & Beaulieu, 2006). The development of fields as biotechnology, for instance, would be unthinkable without the enhanced data processing capabilities of distributed computing (Schroeder, 2007). But the major novelty in e-Research is that access to shared computational resources is being extended also to areas that were not computationally oriented in the past as social sciences or humanities, so to mention the most extreme cases.

Access to computational time and equipment represents a communalization in an even earlier stage of knowledge production. It would then not be a communalization of the creation of contributions to the shared knowledge (as in the classical case of publication), nor a communalization of raw data analysis (as in the case of datasets sharing), but a communalization of data production and analysis (Gläser, 2003), thus covering both primary and secondary research.

[10] For example, comparison plays a large role in astronomy to discover new phenomena in the universe.

The availability of more powerful computational tools and applications has consequences at two stages of the research process, depending on the objectives for which computational resources are used:

- The use of computational resources for manipulating and analysing data in new, more powerful forms increases the possibilities of identifying new research problems -exploratory research –.
- The use of computational resources for experimentation and testing of hypothesis - simulation, modelling, etc. -, enhances empirical research.

This share of computational resources has heavy implications in the organisational level of e-Research. Researchers who want to use such resources need to be in close contact with the institutions controlling them. It is important to notice that one of the aspects of socio-technical networks in e-Research is the government of the assignment of computing time and the support needed to operate it[11]. The establishment of procedures for granting access to research equipment is, thus, intrinsic of e-Research. The question that arises is then whether new social structures emerge because of remote access to research equipment. We will address this question in the following sections when approaching the analysis of socio-technical networks in e-Research.

It is important to notice at this point that Merton's norms of 'universalism' and 'communalism' are at the core of e-Research process and a globally shared e-Research infrastructure. Schroeder (2007) goes one step beyond, claiming the openness of both the technological infrastructure and its organization, arguing that open e-infrastructures and openness as a research ideal are mutually constitutive, as openness is a necessary condition for the advance and refinement of scientific knowledge (Schroeder, 2007).

[11] Indeed, there is an Euroean Commission FP7 project, RAMIRI, which aim is to deliver a training and networking program for people involved in planning and managing international research infrastructures within the European Union [http://www.ramiri.eu/]

4 E-RESEARCH SOCIO-TECHNICAL NETWORKS

E-Research is mainly based in large e-Infrastructures, which can be understood from a two points of view:

- as the components of the technological systems that form the electronic networks used, and
- as the organizational networks that arise and collaborate on this technological infrastructures.

This twofold networked structure can be considered as a socio-technical network. Considering e-infrastructures as socio-technical networks will allow us to approach the complex organisational structure of e-Research practices by analysing the interrelations between the social and technical aspects of its organisation.

We have analysed in the previous section how the three defining elements of e-Research - the inclusion of large platforms for communication, massive data sharing and access to large computational resources - impact the knowledge production process and affect research practices by basing the analysis in the scientific method. However, the scientific method is not the only feature facilitating knowledge generation (van der Walt, 2006). Indeed, one of the key factors of any epistemic culture, as it may be e-Research, refers to the social relations between units (Knorr-Cetina, 1999), standing for the social organization of the culture and addressing both organisational and institutional issues in the production of knowledge within that culture (Wouters, 2006).

The organisational study of e-Research socio-technical networks is a key element in the study of e-Research phenomena, to understand how e-Research provides a conduit for knowledge beyond the general changes in the scientific knowledge production process.

In this section, we will address the ways in which e-Research projects create cores of durable networks in knowledge production, and are able to maintain them, connecting research groups, projects, institutions, technological infrastructure, funding mechanisms and discourse, basing analysis in the notion of *Ba* applied to research (Hautala, 2011).

4.1 Nodes in e-Research networks

Some authors argue that complexity is indeed, a property of the actual nature of knowledge generation (Van der Walt, 2006; Van der Walt & de Wet, 2008). Knowledge requires an (or various) agent(s) as knowledge source. Hence, knowledge production can be considered a social system and, as a consequence, can be seen as un-deterministic complex system (Van der Walt & de Wet, 2008).

Considering e-infrastructures as socio-technical networks will allow us to approach the complex organisational structure of e-Research practices through the analysis of the interrelations between the social and technical aspects of its organisation. The nodes, or units, of e-Research networks can be an individual researcher, a research group, an institution or even, in the case of e-Research the technological infrastructure for research - middleware, grid services and resource scheduling - and the funding initiatives.

Socio-technical systems are built then by the functions of its social and/or technical elements working for a common goal. Still, organizations have boundaries and interactions occur within the system and its sub-systems (local practices), and between the wider context and the dynamics of the environment[12].

[12] Socio-technical Systems Roundtable (STS/RT) outputs. Principles of STS Analysis and Design:

Hence, we can identify three groups of nodes depending on the level of context they belong to in e-Research socio-technical networks:

1. Nodes involved in the local context of a project
   - Individual researchers and research groups within research institutes, universities and companies
2. Nodes of the wider context of the project
   - Technological infrastructure
   - Other related e-Research projects
3. Nodes external to the research itself but also driving e-Research dynamics
   - Funding agencies and governments
   - e-Research discourse

All the nodes are interrelated in very diverse ways and the linkages between them are essential for the articulation of the knowledge generation process in e-Research. Indeed, as discussed previously, researchers may be working in collaboration with other specialists and with new technological tools and funding schemes, so there exist both material and symbolic relations between nodes in order to perform research and to communicate with others. Literature offers a number of possibilities for the study of socio-technical networks, but the concept of *Ba* (Nonaka & Takeuchi, 1995, and applied to international research groups in Hautala, 2011) is used here as basis, since it approaches specifically the problem.

4.2 Interrelations between nodes in e-Research networks

> "Ba *is a physical, mental and/or virtual context of interactive knowledge creation that exists situationally in relation to the knowledge type created, individual characteristics of the participants and structural factors of an organisation"*
>
> *(Hautala, 2011)*

Ba is considered the physical (location), mental (the share of experience or ideas) and/or virtual (e-mail communication, distributed collaborative platforms) arena of knowledge creation. According to its proponents, knowledge is created through interaction in *Ba*, and this interaction is bound by time, space, participants, their different contexts and their trust for each other (Hautala, 2011).

Ba exists in individual, group and organisational levels. Hence, we can differentiate between generic and specific *ba*. Generic *Ba* would describe the general arenas of interaction in an organisation, while specific *Ba* would exist uniquely in contexts with a common goal, such as a project (Hautala, 2011).

A *Ba*'s more structured definition is provided by Nakamori (Nakamori, 2006, cited in Hautala, 2011), for whom *ba* is a system of elements (e.g. infrastructure, actors), characteristics (e.g. hierarchy, communication) and relationships. In the following, a framework for analysing e-Research socio-technical networks is provided where the nodes in the three contexts identified above are considered the elements of the generic *Ba* of e-Research networks. Each of the contexts can be seen as specific *Ba* on their own, with its own characteristics and relationships; and are at the same time interconnected, creating the generic arena for knowledge creation in e-Research (Figure 1).

[http://stsroundtable.com/wiki/Reference:Principles_of_STS_Analysis_and_Design]

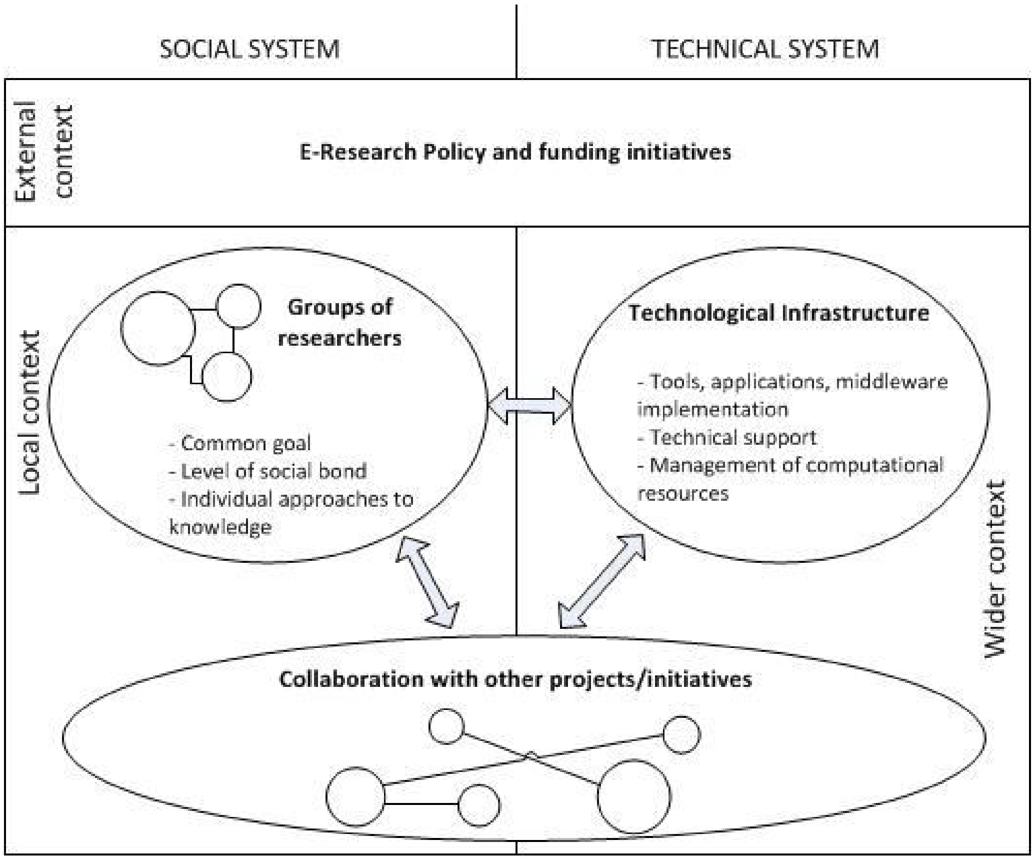

Figure 1. e-Research socio-technical network

Knowledge creation in local context nodes

Researchers, or groups of researchers, involved in any particular e-Research project develop practical modes of organising themselves that support their knowledge creation. According to Hautala (Hautala, 2011) some of the factors framing the creation of a local context include:

- the level of social bond: trust is one of the factors increasing the social bond between researchers, being related mostly to the know-how of the researchers, especially in multi-disciplinary environments. Another factor pointed by some authors studying local practices is the previous knowledge between researchers (Lawrence, 2006). The amount of viewpoints and knowledge exchanges between researchers is thus important to avoid the general tendency of researchers of viewing the research problem individually or having very individualistic modes of operation (Wouters & Beaulieu, 2006).
  Hautala argues that heterogeneous groups - referred to both inter-disciplinary and international research groups - become more innovative than internally homogeneous groups if they are diverse and exist long enough to create common understanding (i.e. they share common interest and commitment) (Hautala, 2011).
- the relationship of researchers to knowledge and research subjects: the role of the researchers in relation to knowledge and the object of research vary for different researchers/groups. Participants' context is an 'individual interpretation of a situation' (Hautala, 2011) and it is based on each individual's values, culture and experiences. Knowledge created in e-Research projects combines several knowledge types and subjective and objective approaches to

knowledge depending on the discipline the different groups come from. Combining individual contexts, each focused in a concrete discipline and/or part of research, and with its own experience and vision, creates the local context of an e-Research project.

Wider context nodes

1) Technological infrastructure relation with local context

The local context of an e-Research project is linked with the technological infrastructure in two different aspects. Often it has to understand the functioning of the infrastructure itself in order to cooperate with computer science researchers in order to adapt some of its technological components to their particular needs. This is the case for e-Research pilot projects. As some studies have shown, technical research work is distributed differently than other types of research, often by clearly dividing each of the tasks. This is one of the challenges that face e-Research pilot projects actually (Hautala, 2011).

In addition, e-Research local nodes must be in close contact with the technological infrastructure for using it and getting support for this use. That means the establishment of access timeframes and power management systems, inter-institutional agreements, etc.

E-Research local contexts depend as much in one and the other aspects of its relation with the technological infrastructures for advance knowledge creation.

2) Other related projects relation with local context

Both internal (within the local context) and external (with other projects) networks have a very important function in sustaining e-Research, and the technological infrastructures in which their work is sustained are thoroughly enmeshed with both types of networking.

Having international links is considered a criterion of excellence (Wouters&Beaulieu, 2006). Also, making contacts outside the own local context is essential for the organization of group strategic activities. This wider network enables obtaining funding, or holding events (Wouters&Beaulieu, 2006).

External context nodes

It often seems that for funding agencies or governmental bodies, knowledge creation is no longer aimed at generating knowledge as such but, for enhancing countries' economic and social performance by combining universities, businesses and public policies in the internationalising context (Hautala, 2011). In Section 5, an overview of the major axis of the e-Research discourse built by some major e-Research funding programmes was provided. One important point to highlight here is that language, discourse and knowledge are related (Hautala, 2011).

e-Research knowledge creation, therefore, present some issues because its dependency on funding schemes, where attention is demanded to be paid to added value, to social relevance and other forms of intervention (Beaulieu & Wouters, 2009). New technologies, in this case, large technological infrastructures for research, carry a particular set of hopes about increased efficiency, relevance or novelty (Hine, 2006). Funding bodies also foster particular expectations about doing something for the global goal of shaping an e-Research future. These expectations are constitutive of e-Research local contexts (Beaulieu & Wouters, 2009).

In addition to expectations, the nodes in e-Research socio-technical networks also share obligations. The relations between funders, researchers and those who believe they will benefit from the interventions promised can be thought of as part of the whole network (Beaulieu & Wouters, 2009).

5 EUROPEAN E-RESEARCH INITIATIVES STUDY

A qualitative research methodology has been used, through existing and conducted in-depth interviews detailed analysis, to achieve a broader understanding of the changes being produced by e-Research practices from the inside.

5.1 Interviews

In order to obtain information directly from the participants of e-Research on-going projects, a total of 19 European projects' coordinators have been requested for an interview. The nature of the projects approached is heterogeneous, this heterogeneity ranging from the individual experience of the interviewees to the main fields of research of projects or the different use they make from e-infrastructures.

From the 19 projects approached, slightly less than two thirds agreed on collaborating with the research. Other researchers participating on some of the projects previously approached were then also requested for collaboration in order to obtain a bigger sample for the study.

However, due to time schedules coordination issues, only 6 of the requests came into a conducted interview, and only one was an spoken interview (see Annex). Other 3 researchers approached sent previously published interviews, and 5 more published interviews available from the project's website were also used for analysis.

A first set of interviews focus on experimental sciences, in which e-infrastructures capabilities are being applied in different forms. These disciplines are an interesting case of study because their use of computational resources impact different research processes and stages, from primary and derived research through data storage and access, to the use of computation and visualization tools for modeling and prediction. The interviewees of this set are involved in 6 different projects focused on: a) Life Sciences: Protein modeling, Biodiversity research, Molecular biology and Bio-medical science (systemic phenotyping research based on mouse models); and b) Earth Sciences: Earth System Modeling addressing climate change research and Ecosystem research focused in the impact of climate change on natural ecosystems.

A second set of interviews focus on social sciences and humanities e-Research projects in various fields, such as linguistics, social and economic history, and textual studies. These cases are interesting because e-Research within these disciplines is mainly focused in building digital libraries, encompassing both data obtained directly in digital forms and the digitalisation of previously collected material. Hence, as we will see in the following, this kind of e-Research emphasizes more data sharing and data comparison than the availability of large computational power, through for example semantically enriched data and methodologies.

Last, two in-depth interviews to participants of an European project which objective is to build an open access resource sharing network, the OpenAIRE project, are analysed. Resource sharing in this case is not focused on data or computational resources sharing, but in the creation *of* an open access research publications infrastructure, facilitating the peer-review process as well as other important forms of publication exploring also the way in which to manipulate research datasets of various forms in combination with research publications. This last study stands for the more external form of e-Research changes in the knowledge creation process, but still important in the e-Research analysis.

The interviews conducted and the analysis of existing documentation and published interviews is designed to emphasize some of the aspects of the theoretical framework presented in previous sections to address the impacts of e-Research in knowledge production. Two levels of analysis are set:

- Assessment of the main characteristics and limitations detected in e-Research
  - Sharing data and use and dependency on computational resources
  - Higher research goals and e-Research sustainability

- Networked nature of e-Research
  - Relations between participants
  - Relations with the technological infrastructure
  - Relations with other projects and initiatives
  - Relations with funding agencies and governmental bodies

5.2 Assessment of the main characteristics and limitations detected in e-Research

Sharing data

Sharing distributed research data in an open way between groups of researchers is one of the main pillars of e-Research. One common point between all the interviewees is that making available collections of research data "allows avoiding repetition of effort in a variety of experiments" and, as a consequence, "(research) material can be enhanced". In addition, sharing data improves comparative research, enriching knowledge of the area of study:

> *"These databases* (on sequences and structures of biopolymers) *can help move research results from one organism to another and compare them with each other, for example, to draw conclusions about evolution",*
>
> **Prof. Piotr Zielenkiewicz**, ELIXIR coordinator (ELIXIR, 2011)

> *"Our knowledge of slavery, for example, has been transformed, widened and deepened because of the availability of databases on the slave voyages. (...) Bringing these disparate sources of data together allows researchers to gain a more detailed picture of their subject",*
>
> **Peter Doorm**, DARIAH coordinator (DARIAH, 2010)

In some cases, sharing access to raw data fosters a "*more crossdisciplinary approach to research material, which in return encourages more collaboration among researchers",* says Ioli Kalavrezou from Harvard University (GridTalk, 2009):

> *"Those involved in the project consortium range from physiologists to environmental scientists, with a large number of participants involved in agronomy ecology and biogeochemistry. This range of disciplines is essential to the long-term development of ecosystem science (...) in order to develop a true systems approach, resulting in models that transfer information across scales"*,
>
> **Abad Chabbi**, EXPEER coordinator (EXPEER, 2011)

Another consequence of the availability of research data pointed previously is that it affects the quality control mechanisms of research. As Peter Doorn (DARIAH,2010), says: "*Allowing access to data also opens it up to further scrutiny, meaning that it can be checked and verified, and its credibility established. Although the risk of data being misused cannot be removed entirely, it can be minimised with the imposition of certain limitations*". This new panorama requires often the inclusion of legal aspects discussion within e-Research projects in terms of licences to collect data and also to use it, as it is the case in OpenAIRE (see Annex (c)).

Use and dependence of computational resources

The actual use of computational resources in experimental sciences and in social sciences differs substantially and, as a consequence, the dependency on them is also perceived differently between one and the other group of researchers interviewed.

The role of computational resources for manipulating and analysing data has become an important part in the experimental fields approached:

> *"The role is increasing significantly every year. Most of non-computer scientist are adding computer modeling in a direct or indirect manner to their work.(...) Soon, I believe, it will have a main role in describing protein dynamics, leaving experimental data to a second role",*
>
> **senior researcher in protein modeling** (see Annex (a))

Intensive computation tools used in these fields range from system modelling, validation, prediction or visualisation tools, thus integrating both observation and experimentation approaches to their research problems, enhancing empirical research and increasing the possibilities of identifying new research problems:

> *"The EXPEER research facilities are a combination of research platforms (...) all linked: long-term, highly instrumented field scale experiments are connected to networks of instrumented observation sites, controlled environmental ecotrons, and analytical facilities, data management and modelling infrastructure, enabling the experimental study of ecosystem responses to change",*
>
> **Abad Chabbi,** EXPEER coordinator (EXPEER, 2011)

> *"Biodiversity research will increasingly depend on the new infrastructure capabilities. The infrastructure will complement experimental reductionist methods with 'holistic' modelling approaches on very large data sets. Such analysis will help to find patterns of correlating phenomenon and possible underlying processes. This will assist in more targeted experimentation, which in turn provide new data for analysis in the infrastructure",*
>
> **coordinator of a biodiversity e-Research project** (see Annex (d))

Social sciences and Humanities are in a different stage in the process of adopting e-Research computational capabilities, originally mainly interested in building digital libraries:

> *"there is of course a preliminary disctintion to be aware of, namely the different needs of scientists of "hard sciences" (eg physics) and "soft sciences" (eg history)"*
>
> **OpenAIRE project researcher (I)** see (Annex (c))

Although Ioli Kalavrezou, from Harvard University, argues that "*with access to the digital infrastructures the humanities are becoming almost a new discipline*" (GridTalk, 2009), other social scientists point to the difficulties these disciplines face when adapting research infrastructures:

> *"A lot of material has been digitised over the years and made accessible, but this work has faced all kinds of difficulties. The main problem is that this material is widely scattered – all kinds of institutions put material on the web, but this doesn't mean that it is easy to find, easy to reach or easy to use",*
>
> **Peter Doorn**, DARIAH coordinator (DARIAH, 2010)

And the adoption of e-infrastructures in social sciences and humanities is still more focused on the expectations of their technical capabilities rather than on their actual use:

> *"E-Science, and specifically the Data Grid, have the potential to markedly assist this change* (a methodological revolution in humanities)*",*
>
> **Paul Ell**, Queen's University Belfast (GridTalk, 2009)

> *"Once these* (materials) *are digital, we want to be able to find them easily, to process and analyse them with user-friendly yet innovative software tools, and to publish our findings, in order to share and discuss our new insights with our colleagues",*
>
> **Peter Doorn**, DARIAH coordinator (DARIAH, 2010)

Higher research goals and e-Research sustainability

Hence, we can see how the promise of the technology capabilities of e-infrastructures is important in the shaping of e-Research projects and the adoption of e-Research in general. The ambition of e-Research of transforming the way research is conducted by expanding its scale is not only reflected in disciplines in which its adoption has not yet totally arrived as the Humanities, but also in those in which e-Research has already an important role. A common point observed in all the fields in which the interviewees do their research, is a shift towards higher research goals:

> "(e-infrastructures) *are enabling entirely new ways of encountering the world, raising questions about what it is to be human that have simply not occurred to us before",*
>
> **John Byron**, Australian Academy of the Humanities (GridTalk, 2009)

> *"With infinite computational power we would have a detailed and complete view of all protein dynamics. (...) On the other site, the enormous amount of information that we are generating require complex infrastructures to store it and to access it easily",*
>
> **senior researcher in protein modeling** (see Annex (a))

The possibility of having infinite computational power and data storage capacity, as well as of collaborating with other researchers from around the globe and from different specialities, allows researchers to start thinking about higher research goals, facing global challenges:

> *"The development of international coordinated experiments is required to enhance model reliability and estimate uncertainties. The development of comprehensive Earth System Models that combine the physical, biological and chemical dimensions is underway to improve our understanding and prediction of both natural variability and human-induced climate changes",*
>
> **Dr. Sylvie Joussaume**, IS-ENES coordinator (IS-ENES, 2010)

> *"Sequencing the genome is becoming easier, because the amount of relevant data increases exponentially. Human genome, which has more than 3 billion nucleotides, will also soon be possible to sequence for a relatively small price",*
>
> **Prof. Piotr Zielenkiewicz**, ELIXIR coordinator (ELIXIR, 2011)

> *"it is essential that proper characterisation (systemic phenotyping), archiving and distribution of mouse models for the biomedical research community is achieved, but this cannot be fulfilled by individual research facilities or on the national level alon",*
>
> **Prof. Martin Hrabé de Angelis**, INFRAFRONTIER coordinator (INFRAFRONTIER, 2010)

How these common goals of e-Research projects are compatible with individual research goals seems not to be a problem for the scientists interviewed. Moreover, "*usually the common goal* (of an e-Research project) *refers to a strong overlap of interest*" and therefore the individual and the project's goals "*are quite compatible*". In other words:

> *"Each collaborating group of researchers has its own scientific goal. This can be a PhD project, a department, an EU project or global programmes on societal problems. The* X *project is interacting with the user community to identify emerging needs and to tune the infrastructure capabilities towards these needs",*
>
> **coordinator of a biodiversity e-Research project** (see Annex (d))

Having a common goal, however, is now enough for achieving the success, and as a consequence, the sustainability of e-Research projects:

> *"Humanities computing has a long history of creating tools, resources and services which don't work with each other, and which are not sustainable in the long term",*
> **Martin Wynne**, Oxford e-Research Centre (GridTalk, 2009)

A major limitation for the achievement of e-Research projects' objectives is the standardization of data formats and the interoperability between computational, data infrastructures and tools in order to facilitate their interchangeability between different research groups:

> *"The problem is that progress in obtaining the data is not accompanied by appropriate development standards for their processing and storage. There is a lack of infrastructure and bioinformatics tools that would allow for efficient access to desired results and their analysis. Existing tools are not sufficient to deal with such enormous and heterogeneous databases",*
> **Prof. Piotr Zielenkiewicz**, ELIXIR coordinator (ELIXIR, 2011)

> *"There is no doubt that achieving interoperability from a multi-layered perspective is essential not only in the shift towards an advanced information space and cross-disciplinary research but also to foster innovation across the board. (...) Standards are also effective in formalising the goals and outputs of any research project underpinned by advanced infrastructures, tools and Systems (...) enabling users to relate to and connect with others",*
> **Tobias Blanke**, Kings College (DL.ORG, 2010)

5.3 Networked nature of e-Research

Relations between participants

E-Research projects integrate diverse researchers, research groups, institutions and disciplines in order to achieve their research goals. The creation of a local context relies heavily on the level of social bond between researchers and their mutual trust and relationship with their own research subjects:

> *"Communication is a vital part of this process, whereby expertise, knowledge and best practices should be interwoven to create a point of gravity (...) while also ensuring that any new knowledge acquired in this area is passed on to others, thereby ensuring collective knowledge and awareness",*
> **Tobias Blanke**, Kings College (DL.ORG, 2010)

> *"For this reason we've brought groups together with different expertise (...) relevant for the health spectrum. (...) Scientists were very keen to get involved in the network because they see the benefit for every single partner, and can bring additional capacities and expertise",*
> **Prof. Martin Hrabé de Angelis**, INFRAFRONTIER coordinator (INFRAFRONTIER, 2010)

The implementation of a project's work plan, however, requires some level of standardisation of the ways in which the different actors work, in order to avoid possible "*bottlenecks that could even make fail the project*", although "there are clear differences in the roles and functions of the individual partners":

> "(the project) *highlighted the need for common methods, tools and approaches across sites to be defined and implemented",*
> **Abad Chabbi**, EXPEER coordinator (EXPEER, 2011)

And, as a consequence, a more complex, and usually distributed, organisation is needed:

> *"These labs work under certain standard operation protocols and agreed quality controls coordinated from a central point to manage the capacity. (...) The problem with not having protocols is that there are different standards between university labs",*
>
> **Prof. Martin Hrabé de Angelis**, INFRAFRONTIER coordinator (INFRAFRONTIER, 2010)

One point highlighted by several interviewees is that these projects "*integrate and enlarge the community* (that is working on some research field) *even more*".

Relations with the technological infrastructure

> *"Within technological infrastructure it is a continuous learning project. (...) Of course it always add a small challenge to the in-house individual research, but a challenge that allows your research to grow.",*
>
> **senior researcher in protein modeling** (see Annex (a))

Understanding the functioning of the infrastructures and acquiring knowledge on how to use and profit computation tools in different research fields become necessary for conducting e-Research. "*Probably we should start educating our students so they can take advantage of them and make good use*", argues one of the scientist interviewed. Indeed, one of the activities that most of the projects approached consider as an inseparable part of them is to develop "*training activities for young researchers*".

Another important aspect highlighted in the interviews is the need of joint collaboration with technical research, both for getting support in the use of technological tools and access to infrastructures:

> *"This is a very new instrument and there are no showcases or time frames to demonstrate how it might work, he said, "so in that sense it is a challenge",*
>
> **Prof. Martin Hrabé de Angelis**, INFRAFRONTIER coordinator (INFRAFRONTIER, 2010)

And for getting the development of new tools or enhancing technological resources in a manner that eases the research:

> *"will be achieved through joint technical developments of software tools, exchange of expertise, access to model information and data (...) This should help enhance the efficiency (...) and reduce the strong technical burden required for these models, while keeping model diversity, which is a key element to address uncertainties in climate projections",*
>
> **Dr. Sylvie Joussaume**, IS-ENES coordinator (IS-ENES, 2010)

A last aspect of the relationship of e-Research internal context with technological infrastructures is the establishment of access and use agreements with the physical sites and bodies managing these infrastructures:

> *"Such services (the ones that the project is setting up) imply that (the project) must have optimized links with other supporting infrastructures, such as GEANT, PRACE and EGI, as well as data facilities such as GBIF, EBI or ESA",*
>
> **coordinator of a biodiversity e-Research project** (see Annex (d))

Relations with other projects and initiatives

Collaborations with other European projects and initiatives are crucial to approach the global goals of e-Research projects. One of the reasons that these projects have in dedicating such important efforts for creating global network is achieving the interoperability we have mentioned before in order to address the fragmentation that exists in some fields:

> *"its aim (referring to the relation of a project with other projects) was primarily to create synergies and interactions between existing experimental platforms and large uncoordinated and fragmented networks of ecosystem observation sites across Europe",*
> **Abad Chabbi**, EXPEER coordinator (EXPEER, 2011)

> *"Community engagement is key to bringing interoperability benefits and best practices into sharp relief",*
> **Tobias Blanke**, Kings College (DL.ORG, 2010)

Thus, establishing an high level governance structure requires not only from organisational efforts at an internal level (within the project), but also in a wider context.

Hence, what motivates having strong synergies with a wider community of researchers, projects and initiatives is both "*benefiting from the expertise of the other communities" and by this, achieving excellence in a concrete field:*

> *"(The project) has established relationships with national infrastructures to ensure a mutually beneficial approach to research infrastructures for the Arts and Humanities. (...) Showcasing user models and perspectives in this context is also of interest as we widen the scope and ensure a bottom-up approach and community-driven input, which will also help to gain a critical mass",*
> **Tobias Blanke**, Kings College (DL.ORG, 2010)

Excellence that in many cases is essential for a strategic point of view, as these wider networks often enable obtaining future funding for the continuation of the research and, therefore, positions the research better in the competition for resources:

> *"The seed of these collaborations has conducted to new EU funded projects (...) or to National e-Science calls",*
> **Participant of GISELA project** (see Annex (e))

> *"Through participation (...) we are well positioned among leaders in the worldwide competition for resources, knowledge, and innovation",*
> **Prof. Martin Hrabé de Angelis**, INFRAFRONTIER coordinator (INFRAFRONTIER, 2010)

Relations with funding agencies and governmental bodies

This need for collaboration has not gone unnoticed. Many funding bodies and governmental agencies dedicate major efforts to foster and enhance it. The role that large funding schemes at national and trans-national levels play in the shift towards more global research goals through large technological infrastructures follows a two-ways causation:

> *"We continue negotiations with colleagues in Europe, but we also hope for funds from the Ministry of Science and Higher Education",*
> **Prof. Piotr Zielenkiewicz**, ELIXIR coordinator (ELIXIR, 2011)

Funding schemes on e-infrastructures, on one side, foster the expectations of doing something for the global goal of shaping e-Research future by addressing global societal problems. On the other direction of causality, e-Research projects' scientific objectives focus often in the creation of one global or European e-infrastructure, which in most cases requires from large funds and supporting infrastructures and data facilities governed by transnational agencies:

> *"we are in very close contact with the Commission and governments and the funding agencies, because they are all part of it, and they see what they can do to help us"*
>
> **Prof. Martin Hrabé de Angelis**, INFRAFRONTIER coordinator (INFRAFRONTIER, 2010)

As Wouter Los, coordinator of LifeWatch project argues, political commitment is one of the three essential conditions – together with technology and scientific collaboration - for the implementation of these kind of projects, in order to fund their relatively high costs (LifeWatch, 2010). One can even extract from some of the answers, than the unique role of funding bodies and governments is about funding: "*we* (the community created from several initiatives around the OpenAIRE project) *know* (now) *how to start this kind of projects, with or without EC assistance"* (see Annex (c)).

However, when interrogated about the major drivers of e-Research future shaping, and whether e-Research is or not focused on funding or their objectives seek for an adaptation to research policy agendas in order to obtain resources, a common response is that the major concerns of the scientific community and technology 'push' are the ones driving the direction that funding bodies take in e-Research:

> *"They enable the interpretation of observed changes and provide the future climate projections upon which EU policies on climate change are formulated",*
>
> **Dr. Sylvie Joussaume**, IS-ENES coordinator (IS-ENES, 2010)

> *"Mostly the driver* (of e-Research) *is based on 'technology push' and the scientific community must learn from initial frontrunners what are the potential opportunities",*
>
> **coordinator of a biodiversity e-Research project** (see Annex (d))

6 CONCLUSIONS

The main objectives of this work were to analyze the ways in which the production and management of knowledge is being transformed by e-Research and how e-Research projects create cores of durable networks for knowledge production connecting research groups, projects, institutions, technological infrastructure and funding mechanisms.

For this, some cross-cutting aspects of e-Research have been analyzed by the means of a theoretical overview of its major drivers and limitations and its impact on the knowledge production process. At a second stage a set of in-depth interviews to senior researchers from diverse fields that are actually involved in e-Research projects has been analysed in order to approach these aspects from a practical point of view.

Some major conclusions can be drawn:

- E-Research fosters the globalization of knowledge production by allowing more ambitious research goals, based both on the possibility of large-scale inter-disciplinary collaborations and the possibilities provided by e-infrastructures technical capabilities for sharing computational resources and massive research data.
- The adoption of e-Research practices reaches different levels in different research fields and disciplines. There are certain areas of scientific advance, as it is the case of experimental disciplines, whose practices have been more impacted by the use of e-Research technical instruments**,** increasing the dependency these fields have with powerful research infrastructures in order to advance knowledge.
- E-Research, thus, impact several levels of the knowledge production process and affect research practices by changing the methodologies used or the access modes to raw material. The transmission of material and components (results, hypothesis, data, etc..) between researchers is part of the knowledge production process, and is never separated of previously produced knowledge. However, primary and derivative research are, for instance, modified by the availability of large shared datasets, and research fields based on experimentation and monitoring approaches to data start sharing research practices. In the words of one of the interviewees: “With access to massive computing, massive storage, network information… we will access a totally different level of research in almost every field (sooner or later)”.
- E-Research involves a number of players situated at different levels and, thus, forming different contexts of practice. This networked nature of e-Research practices add complexity to the organisational structures needed for sustaining those projects and pose several challenges at different levels. An important effort of these networks is dedicated to facilitate the interconnections between its different nodes and provide a sustainable conduit for advancing knowledge.

7 BIBLIOGRAPHY

BEAULIEU, Anne; WOUTERS, Paul (2009). e-Research as Intervention. In 'e-Research: Transformations in Scholarly Practice', edited by Nicholas Jankowski, 54-69. New York: Routledge.

CERN (2005). What is the Grid?,
available online [http://gridcafe.web.cern.ch/gridcafe/whatisgrid/whatis.html][last accessed: 10 May 2011]

CUMMINGS, J.; KIESLER, S. (2005). Collaborative research across disciplinary and institutional boundaries. Social Studies of science 35(5):703-22.

DARIAH (2010). "Virtual infrastructure for humanities research", interview to Peter Doorn; published in Projects Magazine, 2010, available online
[http://www.dariah.eu/index.php?option=com_docman&task=doc_details&gid=372&Itemid=200][last accessed: 9 June 2011]

DL.ORG (2010). Interview with Tobias Blanke, Kings College; in DL.org Newsletter: Interoperability in focus, April/May 2010; available online [http://www.dlorg.eu/index.php/on-line-media-room/enewsetters][last accessed: 15 June 2011]

ELIXIR (2011). ELIXIR will facilitate biologists' access to biological databases; article published in Chem-Bond magazine, available online [http://www.chem-bond.com/portal_news.php?id=211][last accessed: 11 June 2011]

EXPEER (2011). Implementing infrastructure, the EXPEER project; in International Innovation magazine-Environment 2011, Issue 1, pp. 23-25;
available online [http://www.research-europe.com/magazine/ENVIRONMENT/2011-1/][last accessed: 9 June 2011]

GLÄSER, Jochen (2003). What Internet use does and does not change in scientific communities. Science Studies, Vol. 16 (1), pp. 38-51, 2003.

GRIDTALK (2009). GridBriefing: "Digitising culture: Grids and eHumanities", December 2009; available online [http://www.gridtalk.org/Documents/Grids-and-eHumanities.pdf][last accessed: 11 June 2011]

HAUTALA, Johanna (2011). International academic knowledge creation and ba. A case study from Finland. Knowledge Management Research & Practice (2011) 9, 4-16.

HEY, Tony; TREFETHEN, Anne E. (2992). The UK e-Science Core Programme and the Grid. Future Generation Computer Systems (2002) 18, 1017–1031.

HINE, Christine (2006). New infrastructures for knowledge production: understanding e-science. Hershey, PA: Information Science Publishing.

INFRAFRONTIER (2010). "From mice to men", interview TO Prof. Hrabé de Angelis; IN Projects Magazine, issue 16, April 2010;

available online [http://www.infrafrontier.eu/docs/news/48_Projects_April_2010.pdf][last accessed: 11 June 2011]

IS-ENES (2010). Innovation through integration, interview to Sylvie Joussaume; published in researchmedia.eu, 2010 ; available online [https://is.enes.org/the-project/documents/communication/IS_ENES_Sylvie_Joussaume-1.pdf/view][last accessed: 9 June 2011]

KNORR-CETINA, Karin (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.

LAWRENCE, Katherine A. (2006). *Walking the Tightrope: The balancing acts of a large e-research project*. Computer supported cooperative work (2006) 15, 385-411.

LIFEWATCH (2010). Keeping Watch, interview to Wouter Los; published in researchmedia.eu, 2010 ; available online [http://www.research-europe.com/index.php/international-innovation/][last accessed: 15 June 2011]

MERTON, R. (1942). *The normative structure of science*, in 'The Sociology of Science', University of Chicago Press, Chicago, 1973, pp. 267–78.

MEYER, Eric T.; SCHROEDER, Ralph (2009). *Untangling the web of e-Research: Towards a sociology of online knowledge*. Journal of Informetrics 3 (2009) 246-260.

SCHROEDER, Ralph (2007) e-Research Infrastructures and Open Science: Towards a New System of Knowledge Production?. Prometheus, Vol. 25, No. 1, March 2007.

SCHROEDER, Ralph (2008). *e-Sciences as research technologies: reconfiguring disciplines, globalizing knowledge*. Social Science Information (2008) vol. 47(2), pp. 131-157.

VAN DER WALT, Mariana; DE WET, Gideon (2008). *A framework for scientific knowledge generation*. Knowledge Management Research & Practice (2008) 6, 141-154.

WALSH, J. P.; MALONEY, N. G. (2007) *Collaboration Structure, Communication Media, and Problems in Scientific Work Teams*. Journal of Computer-Mediated Communication, 12:712–732.

WOUTERS, Paul (2006). What is the matter with e-Science? - thinking aloud about informatisation in knowledge creation. Speech at (?).

WOUTERS, Paul; BEAULIEU, Anne (2006). *Imaging e-Science beyond Computation*. In 'New Infrastructure for Knowledge Production: Understanding E-Science', edited by Christine Hine, 48-70. London: Information Science Publishing.

Annex – Interviews

<u>a. Interview to a senior researcher in protein modeling</u>

- **Why or how did you start being involved in computationally oriented research in your field?**

I was always quite interested in programming, computers and physics since a kid. Then the last two years of my chemistry degree I saw there was a field joining all them. So I entered...

- **What is the role of computational resources for protein dynamics research? How they are being used actually?**

The role is increasing significantly every year. Most of non-computer scientist are adding computer modeling in a direct or indirect manner to their work. From a validation tool, it is being "upgraded" to an atomic scale prediction tool.
Soon, I believe, it will have a main role in describing protein dynamics, leaving experimental data to a second role.

- **What kind of impact do you think this global research infrastructures have or will have on protein modelling research? In which modes do you think research in this field will change, if it is the case, due to the access to computational resources? And to what extent do you think life sciences will depend, if it would, on research infrastructures in order to advance?**

Our main limitation to obtain an accurate view of protein dynamics is computational resources. With infinite computational power we would have a detailed and complete view of all protein dynamics. Thus, improving computational infrastructures we get closer to this aim. On the other site, the enormous amount of information that we are generating require complex infrastructures to store it and to access it easily

- **And, considering also other disciplines, what kind of impact do you think research infrastructures have or will have on research practices, as a whole?**

Similar to ours. With access to massive computing, massive storage, network information... we will access a totally different level of research in almost every field (sooner or later). Probably we should start educating our students so they can take advantage of them and make good use.

E-Science projects make researchers from different locations, research institutions and backgrounds work together towards a common goal.

- **How do you perceive such 'common goal' within the participants of some of your projects?**

Well, usually the common goal refers to a strong overlap of interest. If there is very mild overlap then the projects tends to fade away.

- **Did you have any contact with the participants of your (inter-institutional) projects before embarking in the project?**

Sometimes I did sometimes not. Most of times these inter-institutionals projects originate from exposeru of your research (or the other group/s) in conferences, articles, etc without any previous contact

- **How this 'common goal' is compatible with individual research?**

Since usually there is a strong overlap of interest they are quite compatible. Furthermore many times the individual research shifts slightly from the inter-institutional collaborations.

However, distant and inter-institutional (& inter-disciplinary) collaboration are not the only challenges of computationally oriented projects. This kind of projects involve links not only within groups of researchers but with the technological infrastructure itself and with funding bodies.

- **How would you describe these relations/ties between these different actors? How the different actors involved in these projects influence each other's research (i.e. between user communities and technological infrastructure, with funding bodies, ...)?**

Within technological infrastructure it is a continuous learning project. Furthermore, compatibilities have improved significantly in the last years. Of course it always add a small challenge to the in-house individual research, but a challenge that allows your research to grow. From the funding site it might add some complexity at the beginning, but usually it is well defined and controlled by the institutions.
At the same time it is true that research no involves many more players than before, technicians, managers, etc. which in some particular cases might be a bottleneck and even make fail the project.

- **Do this relations provoke a sense of some kind of e-research 'community' / 'network'?**

Well, I guess we do feel closer to each other than in the past. Communications are instantaneous and quite frequent. Information is constantly there, in a device that we use in your daily personal life, at the reach of our hand. Now there is no problem in asking a person at the other site of the world about this or that aspect of your research while buying tickets for the movies in the internet. Many companies are specialized in sending your information related to your research, and performing a follow up of novel publications related to you. It does build a network feeling.

In the same line but moving further, collaborative projects require from an important dedication for creating networks.

- **In the light of this and from your experience, how would you say these projects create cores of durable networks for research advance?**

Some do some do not. I read recently that 65% of the project in the world fail. From many different reasons. Some, of course will create stable cores, but I would say on the 20% or less.

The promise of technology capabilities influencing how research will be conducted are important in shaping research infrastructures development and adoption. We have previously mentioned funding bodies as national governemnts or the EC as being actors of research infrastructures newtorks. Research infrastructure projects are usually large projects that depend heavily on funding programmes and, therefore, on research policy.

- **Do you think the actual direction that take e-Science efforts reflect the concerns of the scientific community? What is, in your opinion, the major driver of e-research?**

The major drivers are, in my experience, better infrastructures and specialized knowledge

And to finish,

- **How would you define e-Research?**

Without tacking in consideration any previous question in this interview I would say it is research that is done in the web.

b. Interview to a participant of the OpenAIRE project (I)

According to OpenAIRE website, the goal of the project is to establish *"an infrastructure for researchers to support them in complying with the European Guidelines on Open Access"*, and to build a *"portal and e-Infrastructure for the repository networks"*.

- **What is the role of open repositories for research practice? How they are being used actually?**

La generación actual de repositorios busca aumentar el número de publicaciones científicas disponibles en abierto y con ello aumentar la difusión del conocimiento generado en universidades y centros de investigación. En este último caso, no se trta solo de almacenes de artículos sino que se procura un valor añadido mediante el uso de metadatos y la aplicación de estándares que permiten el intercambio y recolección de contenidos.

- **What kind of impact do you think an European repository facility have or will have on research? In which modes do you think research will change, if it is the case, due to the access to this infrastructure? And to what extent do you think research practices will depend in the future on these infrastructures in order to advance?**

Era imprescindible que la Comisión Europea asumiera un papel más activo en las iniciativas de acceso abierto. Esta infraestructura ha generado no solo estándares y herramientas sino que sirve como modelo y ejemplo y puede ser muy relevante facilitando la integración de muchos más repositorios incluso fuera de la región. Para los científicos el depósito en abierto se debe convertir en una práctica habitual.

E-Science projects such as OpenAIRE make researchers from different locations, research institutions and backgrounds work together towards a common goal.

- **How do you perceive such 'common goal' within the participants of the project?**

Hay una estrategia común para compartir recursos, establecer normas y desarrollar herramientas. Se está formando una red de cooperación y generando sinergias.

Distant and inter-institutional (& inter-disciplinary) collaboration are not the only challenges of e-infrastructure projects. This kind of project involves links not only within groups of researchers but with the technological infrastructure itself and with funding bodies.

- **How would you describe these relations/ties between these different actors? How the different actors involved in these projects influence each other's research (i.e. between user communities and technological infrastructure, with funding bodies, ...)?**

El mayor logro del proyecto es que ahora la red OpenAire se ha convertido en uno de los principles interlocutores de la Comisión Europea sobre los temas de acceso abierto y ello tendrá consecuencias positivas a largo plazo.

- **Do this relations provoke a sense of some kind of e-research 'community'/'network'?**

Sin duda. La preparación de un proyecto de continuidad, OpenAire+ asi lo demuestra

In the same line but going further, OpenAIRE collaborations with other European and national initiatives are crucial to approach the goal of the project,

- **In the light of this and from your experience, how would you say these projects create cores of durable networks for research advance?**

No se puede generalizar, pero en este caso ya existía una base sólida. La iniciativa de la CE ha permtido consolidarla.

We have previously mentioned funding bodies as national governments or the EC as being actors of e-infrastructures networks. E-infrastructure projects are usually large projects that depend heavily on funding programs and, therefore, on research policy.

- **Do you think the actual direction that take e-Science efforts reflect the concerns of the scientific community? What is, in your opinion, the major driver of e-research?**

Las políticas europeas se han decantado ultimamente por la innovacin, en detrimento quizá de la ciencia básica. La e-investigación es una opción pero también una necesidad para rentabilizar los recursos y potenciar los resultados de la ciencia básica.

c. Interview to a participant of the OpenAIRE project (II)

According to OpenAIRE website, the goal of the project is to establish *"an infrastructure for researchers to support them in complying with the European Guidelines on Open Access"*, and to build a *"portal and e-Infrastructure for the repository networks".*

– **What is the role of open repositories for research practice? How they are being used actually?**

*a) Till present Austrias research community do not have the possibility to use Austrian repositories. Therefore the community is relying on repositories located in other countries. B) there is of course a preliminary distinction to be aware of, namely the different needs of scientists of "hard sciences" (eg physics) and "soft sciences" (eg history) c) from our point of view (university of vienna) we think that open repositories can play a major role in near future, for many reasons. Therefore we created an task force (open access group) at our university, reuniting in this group all competences we think are needed, *) to enable our scientific community to comprehend the role of OA-repositories,*) to give them an advice in publication matters (eg on how to deal with publishers), *) we are launching our first universitarian institutional repository (October 2011) *) and we are organizing periodically helded workshops in legal issues. Now the specific answer to your question: a) we think that open repositories are very important, that is the reason why we are running one, and b) we are encouraging our scientist and researchers (as well as our administrative staff) to use our own repository and to use other repositories*

- **What kind of impact do you think an European repository facility have or will have on research? In which modes do you think research will change, if it is the case, due to the access to this infrastructure? And to what extent do you think research practices will depend in the future on these infrastructures in order to advance?**

The impact will be very strong, and not only in research. It will also change the ways on qualifying for a project, will change the attitude of students (because they will need to learn on how to use these systems), will change the way on how to apply for proejcts (you can define a restricted area inside a repository, in where to "publish" the proposals), it will also change the way we are now using local persistent identifiers (and my be lead to a creation of European permanent identifiers). And it will maybe lead to and adoption of European wide DAI (digital author identifications)

E-Science projects such as OpenAIRE make researchers from different locations, research institutions and backgrounds work together towards a common goal.

- **How do you perceive such 'common goal' within the participants of the project?**

In the same way it is defined by the project. It is the same way it leaded to the definition of the following project openaireplus

Distant and inter-institutional (& inter-disciplinary) collaboration are not the only challenges of e-infrastructure projects. This kind of project involves links not only within groups of researchers but with the technological infrastructure itself and with funding bodies.

- **How would you describe these relations/ties between these different actors? How the different actors involved in these projects influence each other's research (i.e. between user communities and technological infrastructure, with funding bodies, ...)?**

We think this kind of impulses are very positive (eg: as we were informed that Portugal has abouto 20 repositories compliant to openaire, we asked ourselves why Austria will have just one). Another impulse is given by the fact, that the acting bodies are not currently involved in research, they are s e r v I c I n g research, and it is very fruitful to compare our local situation with the situation of similar entities in other countries (eg Austrian situation vs Czech Republic ituation in matters like Open Access Policy or use of URN

- **Do this relations provoke a sense of some kind of e-research 'community'/'network'?**

Yes, definitely

In the same line but going further, OpenAIRE collaborations with other European and national initiatives are crucial to approach the goal of the project,

- **In the light of this and from your experience, how would you say these projects create cores of durable**

networks for research advance?

We know better each other, and we know how to start this kind of projects, with or without EC assistance

We have previously mentioned funding bodies as national governments or the EC as being actors of e-infrastructures networks. E-infrastructure projects are usually large projects that depend heavily on funding programs and, therefore, on research policy.

- **Do you think the actual direction that take e-Science efforts reflect the concerns of the scientific community? What is, in your opinion, the major driver of e-research?**

The major driver of e-research is the competition with similar US initiatives (in non profit and in for profit sector)

d. Interview to the coordinator of a biodiversity e-Research project

According to X website, the goal of the project is to build "a world leading e-infrastructure" to support all aspects of biodiversity research. The goal of the X infrastructure is to provide services for scientists and policy users of biodiversity data, integrating large-scale data resources, advanced algorithms and computational capability to support the study of patterns and mechanisms across different levels of biodiversity.

- **What is the role of computational resources for biodiversity research? How they are being used actually?**

a. for (primary and derived) data storage
b. for intensive computation and visualization (i.e. modeling)
c. for workflow processing
d. for Grid management

- **What kind of impact do you think this global e-infrastructure have or will have on biodiversity research? In which modes do you think research in biodiversity will change, if it is the case, due to the access to this infrastructure? And to what extent do you think biodiversity research will depend, if it would, on these infrastructures in order to advance?**

It will change biodiversity research since complex analysis and modeling can be done at much larger scale and speeds with large data sets. Instead of for example modeling a single outcome in a few days, it becomes within our reach to generate thousands of more outcomes within minutes. This will allow studying the effect of for example changing parameters, or the effect of varying data. This is important since biodiversity data and their interrelations are often fuzzy and incomplete, and reliability van be tested. In applications this is relevant for decision support when scenario analyses provides alternative probable results.
Biodiversity research will increasingly depend on the new infrastructure capabilities. The infrastructure will complement experimental reductionist methods with 'holistic' modeling approaches on very large data sets. Such analysis will help to find patterns of correlating phenomenons and possible underlying processes. This will assist in more targeted experimentation, which in turn provide new data for analysis in the infrastructure.

E-Science projects such as X makes researchers from different locations, research institutions and backgrounds work together towards a common goal.

- **How do you perceive such 'common goal' within the participants of the project?**

Each collaborating group of researchers has its own scientific goal. This can be a PhD project, a department, an EU project or global programmes on societal problems. X is interacting with the user community to identify emerging needs and to tune the infrastructure capabilities towards these needs.

Distant and inter-institutional (& inter-disciplinary) collaboration are not the only challenges of e-infrastructure projects. This kind of project involve links not only within groups of researchers but with the technological infrastructure itself and with funding bodies.

- **How would you describe these relations/ties between these different actors? How the different actors involved in these projects influence each other's research (i.e. between user communities and technological infrastructure, with funding bodies, ...)?**

e-Infrastructures allow researchers to cooperate from different locations. X will do this with shared data, workflows and computational capacity, and with 'virtual labs' allowing researchers to discuss running workflows and the results. Such services imply that X must have optimized links with other supporting infrastructures, such as GEANT, PRACE and EGI, as well as data facilities such as GBIF, EBI or ESA. Funding bodies are supported with emerging scientific trends as visible in the usage of X.

- **Do this relations provoke a sense of some kind of e-research 'community'/'network'?**

This possible since X is planning to provide technical support for establishing and operating e-research communities.

X dedicates an important effort for creating a global network: (from the website) "*... its distributed nature results from data generation in networks of research sites and the implementation of regional Service Centres. X builds* *upon the facilities of other infrastructures (such as GBIF or EBI), and a variety of organisations with scientific collections, monitoring sites, or observatories.*"

- **In the light of this and from your experience, how would you say these projects create cores of durable networks for research advance?**

These projects have already created cores of durable networks. X is promoting their effective functioning, as this is crucial to support X' capabilities.

We have previously mentioned funding bodies as national governments or the EC as being actors of e-infrastructures networks. E-infrastructure projects are usually large projects that depend heavily on funding programs and, therefore, on research policy.

- **Do you think the actual direction that take e-Science efforts reflect the concerns of the scientific community? What is, in your opinion, the major driver of e-research?**

This is an interesting question. Mostly the driver is based on 'technology push' and the scientific community must learn from initial frontrunners what are the potential opportunities. Not unless a new group of scientists is exploring new applications within the X architecture, it can be expected that a 'market pull' mechanism takes over as driver.

e. Interview to a participant of the GISELA project

According to GISELA website, the goal of the project is to "*implement the Latin American Grid Initiative sustainability model*" and to "*provide Virtual Research Communities with the e-Infrastructure and Application-related Services required to improve the effectiveness of their research*"

- **What kind of impact do you think a Latin American research infrastructure have or will have on research? In which modes do you think research will change, if it is the case, due to the access to this infrastructure? And to what extent do you think research practices will depend in the future on these infrastructures in order to advance?**

In my opinion, such an e-Infrastructure is of outmost importance since it will enable to work Latin American researchers in a collaborative way with other colleagues settled worldwide. At the same time, it will permit them to access larger resources. The extent of research practices is even a necessity since many lines of work can nowadays be only performed by means of simulation techniques, which could be heavy computational consuming.

E-infrastructure projects such as GISELA make researchers from different locations, research institutions and backgrounds work together towards a common goal.

- **How do you perceive such 'common goal' within the participants of the project?**

In GISELA, collaborative practices has been fostered in several ways: dissemination of the technology and the developments; research conferences for presenting results; specific activities and Work Packages devoted to such a task... But I consider as a main step forward the organization of gridification schools where researchers and ICT Grid experts sat together for two weeks in order to port to the Grid scientific codes. Of course, the created synergies went beyond.

Distant and inter-institutional (& inter-disciplinary) collaboration are not the only challenges of e-infrastructure projects. This kind of project involves links not only within groups of researchers but with the technological infrastructure itself and with funding bodies.

- **How would you describe these relations/ties between these different actors? How the different actors involved in these projects influence each other's research (i.e. between user communities and technological infrastructure, with funding bodies, ...)?**

The emerged relationships are still on the way and producing successful results even when some improvements should be done (mainly, the huge use of the infrastructure). The seed of these collaborations has conducted to new EU funded projects such as CHAIN or to National e-Science calls (CEDIA one, for example).

- **Do this relations provoke a sense of some kind of e-research 'community'/'network'?**

Yes, GISELA is involving RedCLARA since the very beginning and the later is promoting the creation of a network by funding scientific/social e-communities (ComCLARA calls in 2010 and 2011).

In the same line but going further, GISELA collaborations with other European, latin american and national initiatives are crucial to approach the goal of the project,

- **In the light of this and from your experience, how would you say these projects create cores of durable networks for research advance?**

GISELA is collaborating with EGI, CHAIN, EMI and other regional projects. In all of them, sustainability is a must, and the involvement of NREN is crucial. In this sense, a durable network would be consolidated.

We have previously mentioned funding bodies as national governments or the EC as being actors of e-infrastructures networks. E-infrastructure projects are usually large projects that depend heavily on funding programs and, therefore, on research policy.

- **Do you think the actual direction that take e-Science efforts reflect the concerns of the scientific community? What is, in your opinion, the major driver of e-research?**

Yes, it is, although specific actions should be performed in order to bear in mind that not every region has the same development as Europe, i.e. AfricaROC cannot deploy by now all the specifications that an European ROC or, in Latin America, for example, Brazilian NREN has a much advanced development that some Central America Regions.

f. Interview to a senior researcher in molecular modelling and bioinformatics

Your research group uses computational and experimental approaches to for molecular modeling.

- **What is the role of computational resources for life sciences research? How they are being used actually?**

Els recursos computacionals ens permeten processar un volum gran de dades (e.g. projecte de seqüenciació del càncer de Leucèmia -> David Torrents) o simular simular experiments que resultarien massa cars o impossibles (e.g. estudi massiu de la flexibilitat de les estructures de les proteïnes -> Projecte MoDEL; Modesto Orozco). En general s'usa la computació en camps de la genòmica, la simulació de biomolècules (fàrmacs, proteïnes, ADN, ...), en la biologia de sistemes i en les simulacions mèdiques (simulació de teixits)

- **What kind of impact do you think research e-infrastructures have or will have on molecular & bioinformatics research? In which modes do you think research in this field will change, if it is the case, due to the access to computational resources? And to what extent do you think molecular research will depend, if it would, on research infrastructures in order to advance?**

Actualment existeixen molts camps en les ciències de la vida que tenen una dependència amb els recursos de computació. Per exemple, el volum de dades d'un experiment de seqüenciació supera els 4TeraBytes, i es un centre de seqüenciació mitjà pot arribar a produir 2 PetaBytes/any. Gracies a aquest recursos la genòmica ha donat importants fruits en la identificació de noves dianes terapèutiques. En el cas de la simulació de biomolècules, aquestes tècniques han arrelat fortament de el disseny racional de fàrmacs que va canviar al indústria en els anys 80. La biologia de sistemes es sens dubte l'àrea que experimentarà un increment mes important de recursos de supercomputació ja que es treballen en complexos models puguin simular els sistemes de les cél·lules. També caldria esmentar l'ambiciós projecte que es proposa Europa per simular un cervell humà i que requerirà potències de càlcul molt superiors a l'actual.

E-Science projects make researchers from different locations, research institutions and backgrounds work together towards a common goal.

- **How do you perceive such 'common goal' within the participants of some of your projects you have been involved?**

Europa està lluny de competir amb el poder de supercomputació en Estast Units o Àsia. el punt fort d'Europa son les "aplicacions" en ciències de la vida. entre es socis de projectes (europeus) hi ha sempre la sensació de pressió de ser una punta de fletxa del nostre continent (que competeix molts cops amb iniciatives paral·les d'altres regions). Es treballa sempre amb l'objectiu de construir plataformes que incentivin la recerca europea de qualitat i que donin visibilitat internacional.

However, distant and inter-institutional (& inter-disciplinary) collaboration are not the only challenges of computationally oriented projects. This kind of projects involve links not only within groups of researchers but with the technological infrastructure itself and with funding bodies.

- **How would you describe these relations/ties between these different actors? How the different actors involved in these projects influence each other's research (i.e. between user communities and technological infrastructure, with funding bodies, ...)?**

Quan parlem de projectes que involucren diferents actors parlem principalment de projectes europeus. En la meva experiència (projectes orientats aplicacions) es distingeixen 3~4 tipus de socis: Proveïdors de tecnologia (computació / hardware i software), experts en aplicacions i usuaris finals. Es busca que entre els usuaris finals hi hagi empreses rellevants del sector que tinguin efecte tractor sobre la comunitat d'usuaris. Sol haver una relació molt forta entre proveïdors de tecnologia i experts en aplicacions. En alguns casos hi han membres que pertanyen als dos sectors.

- **Do this relations provoke a sense of some kind of e-research 'community'/'network'?**

Jo diria que molts cops les relacions son producte de la e-recerca.

In the same line but moving further, collaborative projects require from an important dedication for creating networks.

- **In the light of this and from your experience, how would you say these projects create cores of durable networks for research advance?**

Com he comentat abans en els projectes es poden classificar els partners entre diferentes clases. Les relacions entre partners de la mateixa classe es perdurable. Això no es tant freqüent en relacions entre proveïdors de tecnologia i usuaris finals (estem parlant sempre en el marc de relació de projectes de recerca)

We have previously mentioned funding bodies as national governemnts or the EC as being actors of research infrastructures newtorks. E-infrastructure projects are usually large projects that depend heavily on funding programmes and, therefore, on research policy.

- **Do you think the actual direction that take e-Science efforts reflect the concerns of the scientific community? What is, in your opinion, the major driver of e-research?**

Jo penso que la força que dirigeix la e-recerca es principalment la notorietat en aspectes tècnics (aconseguir un càlcul superior en complexitat, temps o en mida, etc.). Aquest es un aspecte que no sempre està relacionat amb la qualitat científica. A més aquest es un aspecte que representa una petita part de l'activitat científica. En la meva opinió la comunitat científica busca en la eCiència recursos que resolguin de manera eficient problemes fora de l'àmbit computacional (simular un sistema cel·lular amb alta reproductibilitat, millorar la capacitat predictiva, etc.)