E-Science Models and the Next Generation Grid Infrastructure in the Philippines
Collaborative Efforts in Educational and Research Institutions

Emilie Sales Capellan, Roy R. Consulta
ITPP – TEMEP, Seoul National University
capellan@temep.snu.ac.kr; roy.consulta@gmail.com

Abstract

In a digital information processes, the replica of life cycle are shaping the method and the way the learners' study. In a bigger system, part of the things that will be represented in the life course is the Model, like the research process, through a chain of sequentially interconnected stages or phases in which information is manipulated or produced. This paper presents and discusses methods and ways in which the life cycle approach offers insight into the relationships among the stages and activities in research, especially in the field of technology evolution, like e-Science. And this paper will also present an idea and concept how this research life cycle in e-Science will affect the Philippine community. An understanding of this viewpoint may contribute further insight into the function of e-science in the larger picture of methodical and scientific research.

1. Introduction

Utilization of computers generates many challenges as it expands and develops the field of the possible in methodical and scientific research and many of these challenges are usual to researchers in diverse areas. The insights achieved in one area may catalyze change and accelerate discovery in many others. It is absolutely true in the statement that it is no longer possible to do science without doing computing [1]. Computing in the sciences and humanities has developed a great deal over the past decades. The life cycle method turns us to a more sensitive possible information loss in the gaps between stages. The transition and the evolution points in the life cycle are essential junctions for further important activities in the research field, such as e-Science. Many issues and streams of activity flow throughout the life cycle of research, including project administration, grant procurement, data management, knowledge creation, ethical judgments, intellectual property supervision and technology management as a way e-Science is being implemented. Linking activities across stages requires harmonization and coordination and a sense of continuity in the overall process [2]. In the Philippines, the research undertaken in the Sustainable Technologies Group of the De La Salle University makes use of a highly interdisciplinary approach to providing effective solutions to environmental problems [3]. These problems require an intelligent, integrated approach to yield solutions that are beneficial on a life cycle basis. Also in the Philippines, it makes use of the life cycle framework in most of the projects. Therefore, it makes use of advanced computing techniques such as:

- Knowledge-based and rule-based decision support systems
- Monte Carlo and fuzzy sets
- Pinch analysis
- Artificial neural networks
- Swarm intelligence

To be open and responsive to e-science, researchers must evaluate and assess the services it provided for both research outcomes and data. Given the stages of the life cycle associated with e-science, it needs to determine the services to be provided by research libraries and the partnerships required to implement and sustain these services. There is barely a scientist or scholar remaining who does not use a computer for research purposes. There are distinctive terms in use to point out the fields that are particularly oriented to computing in specific disciplines. In the instinctive and technical sciences, the term "e-Science" has recently become popular, where the "e" of course stands for "electronic" [4]. Science ever more done through distributed and dispersed worldwide collaborations enabled by the Internet, using very large data collections, tera-scale computing resources and high performance visualization.

With the technology today, a very powerful infrastructure is required to support and sustain e-Science. The Grid is an architecture projected to produce all the issues together and make a reality of such a vision for e-Science. In the field of technology, such as Grid computing, architecture examines Grid technology as a standard and generic integration mechanism assembled from Grid Services (GS), which are an extension of Web Services (WS) to comply with additional Grid requirements. The principal extensions from WS to GS are the management of state, identification, sessions and life cycles and the introduction of a notification mechanism in conjunction with Grid service data elements [5]. The e-Science term is intended to confine an idea of the future for scientific research-based on distributed resources especially data-gathering instruments and group researches.

E-Science is scientific investigation performed through distributed global collaborations between scientists and their resources, and the computing infrastructure that enables this. Scientific progress increasingly depends on pooling know-how and results; making connections between ideas, people, and data; and finding and interpreting knowledge generated by strangers in new ways other than that intended at its time of collection. E-Science offers a promising vision of how computer and communication technology can support and enhance the scientific process. It does this by enabling scientists to generate, analyze, share and discuss their insights, experiments and results in a more effective manner.

In the Philippines, as the technology evolved, the agency called ASTI[6] being mandated to conduct scientific research and development in the advanced fields of Information and Communications Technology and Microelectronics, undertake projects committed to the development of its people and country as a whole. ASTI has its project called PSIGrid program, which will initiate the establishment of the necessary infrastructure and community linkages to operate throughout the country its grid facility [7]. ASTI will deploy a reliable and secure grid management system for managing users, nodes and software to ensure the reliability and security of the entire grid.

2. Life Cycle Model of Research

The information that will be presented in this section consists of the variety of methods used in communicating and coordinating research outcomes. The research outcomes and data upon which these outcomes are based collectively document the knowledge for an area of study. The life cycle model helps monitor both the digital objects bound within a stage and those objects that flow across stages. This is represented above in the lightly shaded box around Data and Research Outcomes [8]. Figure 1 show how the life cycle model of research knowledge is being created.

Every chevron in the above model symbolizes a stage in the life cycle of research knowledge creation. The spaces between chevrons indicate the transitions between stages. These transitions tend to be vulnerable points in the documentation of a project's life cycle. When a stage is completed, its information may not get systematically preserved and instead end up dead-ended (most often on someone's hard drive.)Shifts in the responsibility for the objects of research also tend to occur at these points of transition. For example, the data collection stage passes along completed interviews or questionnaires to the data processing stage; the data processing stage passes one or more clean data files to the data access and dissemination stage. In each transition, someone else usually becomes responsible for the outcomes of the previous stage. These transition points become important areas in negotiating the digital duration plan for a project as partners in the life cycle of research identify who is responsible for the digital objects created at each stage.

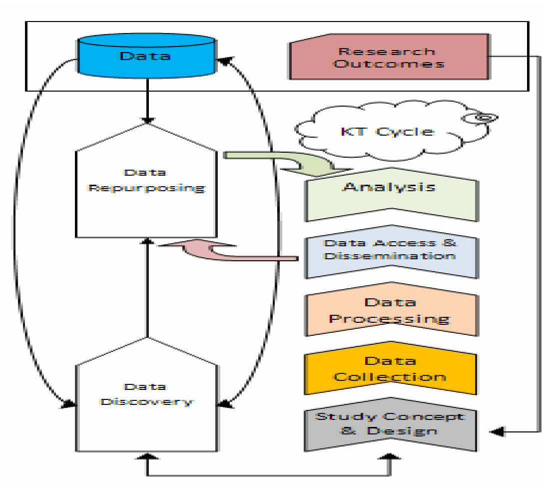

Figure. 1. Life cycle model of research knowledge

Now in e-Science, the knowledge life cycle can be observed as a set of challenges as well as a sequence of stages. Each stage has variously been seen as a blockage. The attempt of acquiring knowledge was one bottleneck recognized early [9]. But so too are modeling, retrieval, reuse publication and maintenance. In this section, we examine the nature of the challenges at each stage in the knowledge life cycle and review the various methods and techniques at our disposal. Although it is often suffer from a deluge of data and too much information, all too often what we have is still insufficient or too poorly specified to address our problems, goals, and objectives. In short, we have insufficient knowledge. Knowledge acquisition sets the challenge of getting hold of the information that is around, and turning it into knowledge by making it functional. This might involve, for instance, making implied knowledge explicit, identifying gaps in the knowledge already held, acquiring and integrating knowledge from multiple sources (e.g. different experts, or distributed sources on the Web), or acquiring knowledge from unstructured media (e.g. natural language or diagrams).

A variety of techniques and methods has been developed ever since to facilitate knowledge acquisition. Much of this work has been carried out in the context of attempts to build knowledge-based or expert systems. Techniques include varieties of interview, different forms of observation of expert problem-solving, methods of building conceptual maps with experts, various forms of document and text analysis, and a range of machine learning methods [10]. Each of these techniques has been found to be suited to the elicitation of different forms of knowledge and to have different consequences in terms of the effort required to capture and model the knowledge [11, 12]. Specific software tools have also been developed to support these various techniques [13] and increasingly these are now Web enabled [14]. However, the process of explicit knowledge acquisition from human experts remains a costly and resource intensive exercise. Hence, the increasing interests in methods that can (semi-) automatically elicit and acquire knowledge that is often implicit or else distributed on the Web [15].

A variety of information extraction tools and methods are being applied to the huge body of textual documents that are now available [16]. Another style of automated acquisition consists of systems that observe user behavior and assumed knowledge from that behavior. Examples include recommender systems that might look at the papers downloaded by a researcher and then detect themes by analyzing the papers using methods such as term frequency analysis [17]. The recommender system then searches other literature sources and suggests papers that might be relevant or else of interest to the user. Methods that can engage in the sort of background knowledge acquisition described above are still in their infancy but with the proven success of pattern directed methods in areas such as data mining, they are likely to assume a greater prominence in our attempts to overcome the knowledge acquisition blockage.

3. Research Trends – e-Science in a Transparency

The fascinating e-Science concept illustrates changes that information technology is bringing to the methodology of scientific research [18]. e-Science is a relatively new expression that has become particularly accepted after the launch of the major United Kingdom initiative [19]. e-Science describes the new approach to science involving distributed global and international collaborations enabled by the Internet and using very large data collections, terascale computing resources and high-performance visualizations. e-Science is about global collaboration in key areas of science, and the next generation of infrastructure, namely the Grid, that will enable it. Figure 2 summarizes the e-Scientific method.

In a simplest manner, it can illustrate and characterize the last decade as directing simulation and its integration with science and engineering – this is computational science. e-Science builds on this adding data from all sources with the needed information technology to analyze and incorporate the data into the simulations.

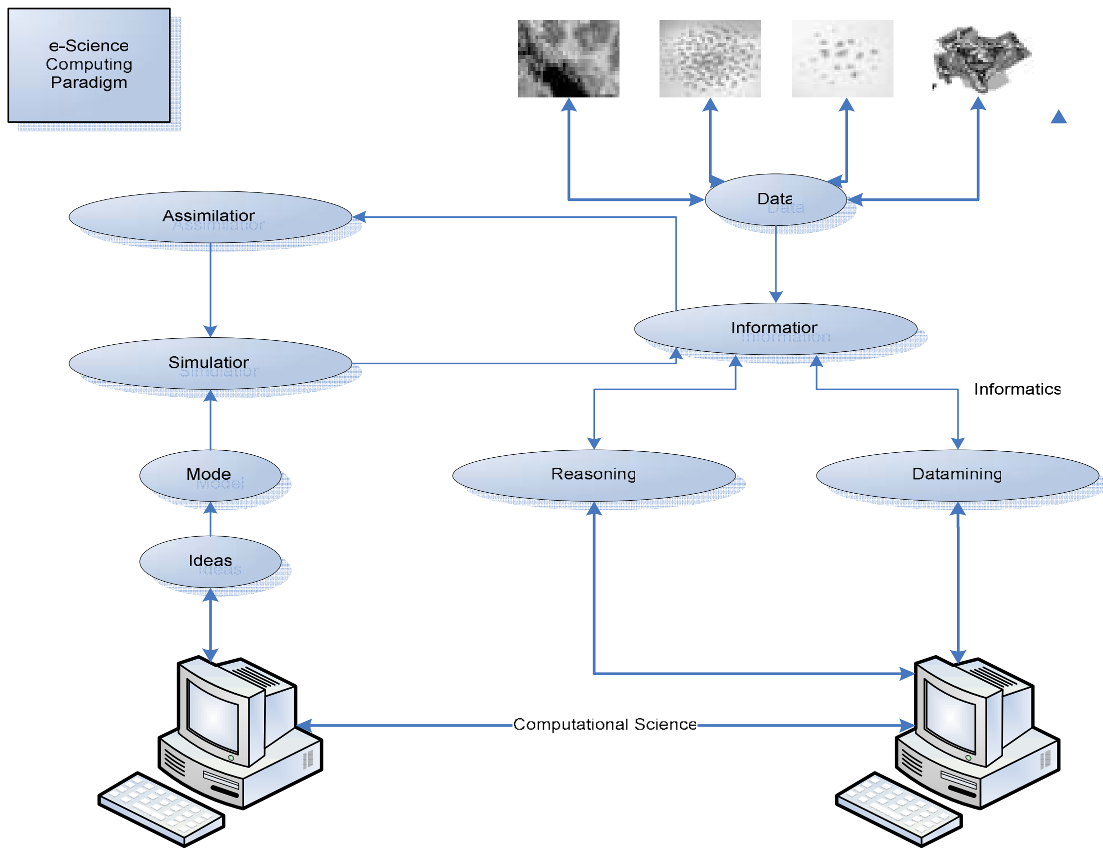

Figure. 2. Computational science and information technology merge in e-Science

Fifty years ago, scientific performance has evolved to reflect the growing power of communication and the importance of collective wisdom in scientific discovery. Originally scientists collaborated by sailing ships and carrier pigeons. At the present aircraft, phone, e-mail and the Web have greatly enhanced communication and therefore the quality and real-time nature of scientific collaboration. The cooperation can be both "real" and enabled electronically [20,21] early influential work on the scientific collaboration. e-Science and hence the Grid is the infrastructure that enables collaborative science. The Grid can provide the basic building blocks to support real-time distance interaction, which has been exploited in distance education. Particularly important is the infrastructure to support shared resources – this includes many key services including security, scheduling and management, registration and search services and the message-based interfaces of Web services to allow powerful sharing (collaboration) mechanisms. All of the basic Grid services and infrastructure provide a critical venue for collaboration and will be highly important to the community.

In Philippine perspective, researchers created what they say are the first generic system for Grid Computing that utilizes an industry-standard Web service infrastructure. The system, called Bayanihan Computing .NET [22] is a generic Grid computing framework based on Microsoft .NET that uses Web services to harness computing resources through "volunteer" computing similar to projects such as SETI@Home [23], and to make the computing resources easily accessible through easy-to-use and interoperable computational Web services. As mentioned in the preceding section that ASTI agency from the Philippines is managing a project called Philippine e-Science Grid Program (PSIGrid). This emerging computing model provides the ability to perform higher throughput computing by taking advantage of many networked computers to model a virtual computer architecture that is able to distribute process execution across a parallel infrastructure. The establishment and planning of this PSIGrid is expected to foster collaboration among local research groups as they share computing resources to further their research efforts. This is also expected to enable efficient utilization of local computing resources.

4. e-Science Practical Model Application

With the global advancement of technology, new advances in networking and computing technology have produced an explosive growth in networked applications and information services. Applications are getting more complex, heterogeneous and dynamic. In the recently concluded forum regarding national e-Science development strategy, held on August 24 at the Westin Chosun Seoul under the supervision of KISTI and under the joint auspices of the Ministry of Science and Technology (MOST) and the Korea e-Science Forum, has reported significance of R&D activity changing into e-Science system.

The necessity of national e-Science is becoming more & more important because of its new research method which challenges huge applications and research in limited environments, improvement in research productivity which enables us to utilize research resource at remote places and collaborate between researchers, education learning trait which enables diverse learning equipment's utilization with networked studying environment, and finally economic development's new growth engine with cutting-edge technology innovation. [24]

One major impact that it had made contributed was on the medical field, for instance on reduction on the period for drug development; enabling global research projects in fields of aerospace development, nuclear fusion research, tsunami and SARS prevention; boosting national science technology competitiveness by developing a new methodology model in which IT and science technologies are converged by securing convergence research's core technology and cooperation and collaboration among nations, regions and fields in that the researchers can have access to cutting edge equipment, data and research manpower. By means of cutting-edge technology innovation, this national e-Science can serve as a new growth engine of economic development, provoking astronomical economic ripple effect.

Aside from the R&D applications, e-Science has also proven its importance by its introduction to classroom. In UK, a pilot project has begun to explore the potential benefits of collecting and sharing scientific data within and across schools and closer collaborations between schools and research scientists with a view to running a national project involving multiple schools. [25]. This pilot project has begun to reveal the educational potential through the collaboration of teachers and students, in a way they input, manipulate their collected data and share this Grid-like technologies, such activities can provide and a larger scale project would have the potential to begin to feed schools-sampled local pollution data into a more significant GRID-based data set which scientists could use to build up a picture of pollution levels across the country.

Another major contribution that UK, being the first country to develop a national e-Science Grid, developed was used in the diagnosis and treatment of breast cancer as in one of the pilot projects, they developed a digital mammographic archive together with an intelligent medical decision support system which an individual hospital without supercomputing facilities could access through the use of grid. This project is called e-DiaMoND and Integrative Biology. [26]

In Australia, they have introduced the world's first degrees in e-Science. Two Australian computer science departments namely Australian National University and RMIT have worked together, established a program called "Science Lectureships Initiative" designed to foster linkages between academia and industry with the idea of attracting students into science-related areas which would then benefit emerging industries. [27] At RMIT, the eScience Graduate Diploma started with only 10 students in the first year, but thereafter struggled to gain enough extra students to become self sustaining as a separate program while at ANU there was a large influx of overseas students particularly from Indian subcontinent and from East Asia. With these initiatives, this can provide guidance and attract other universities to set up similar education programs.

5. Definition and Relevance of e-Science in the Philippine Perspective

As noted by many different studies and researches that been done by the different authors in all parts of the world, we can say that e-Science have its own role, function, and relevance in this modern society. Many developed countries have gone far in this field. However in the Philippines, it is just on the introduction phase. Thus, e-Science could be defined as a solution that can guarantee the Philippines, through international collaboration, improve its technological innovation in researches and discoveries within the applied technological approach.

This paper serves as the driving force in addressing the issue on three most important application of ICT which is education, health and governance. With its direct connectivity to a number of international research and education networks such as Asia Pacific Advanced Network (APAN) and the Trans-Eurasia Information Network 2 (TEIN2), this will benefit researchers on the academe sectors to collaborate in the global research community.

6. Research Life Cycle Model in the Philippines

In Fig. 3, this model PREGINET[28] will be the network backbone that will support the key major players in the whole system flow of the scientific research arising from the academe and the government's research and development institutes. Thus in this research life cycle model, this will serve as the highway that will serve the applications on which, in this research, is the e-Science.

As shown in the model, e-Science will take part as the heart of these important areas of researches, as this will become the central application to researchers from the academe and other R&D institutions, e-Library and distance learning. This platform will allow linkages among its partners in the network locally and globally.

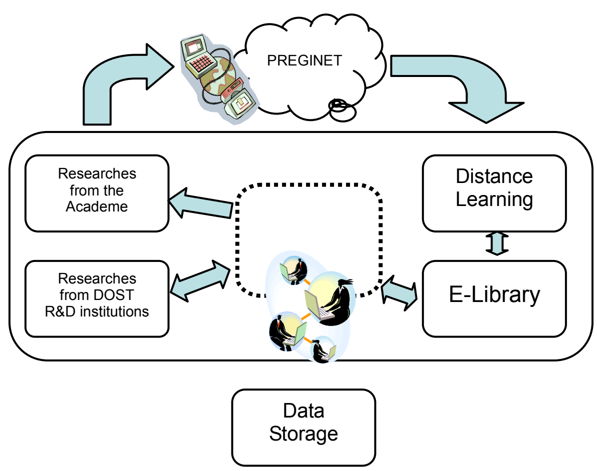

Figure. 3. Development Model and Work Processes

And this will be linked to a central depository which will be managed and controlled by a policy making body or the technical working group.

7. Effect of Research Life Cycle Utilizing Grid Technology – the e-Science in the Philippines

By understanding the full research life cycle allows us to identify gaps in services, technologies and partnerships that could. There is also a need to understand the process of collaboration in e-Science in order to fully and accurately define requirements for next generation Access Grids [29]. The emergence of e-Science systems raises also challenging issues concerning design and usability of representations of information, knowledge or expertise across variety of potential users that could lead to a scientific discovery. [30].

The grid is transforming science, business which in effect, e-Science research, business and commerce will significantly benefit from grid based technologies which will potential increase abilities, efficiency and effectiveness through leading edge technology applications and solving large scientific and business computing problems. Although on the part of socio-economic aspects, this will demand investigation to address issues such as ethics, privacy, liability, risk and responsibility for future public policies. In addition, for the envisaged new forms of business models, economic and legal issues are also at stake which will require interdisciplinary research.

For e-science projects like PSIGrid program in the Philippines, the majority (if not all) of the funding is from government sources of all types. For this cooperation to be sustainable, however, especially in commercial or government settings, participants need to have an economic incentive. With the promising vision that e-Science have, there is a great chance for the PSIGrid program to also participate on the global world and thus come up with technologies that would be beneficial to its citizens.

8. Conclusion

To conclude, given the above viewpoints on lifecycle and e-Science models, there have been important changes how technology especially on scientific researches can be successfully managed.

The trend in technology goes towards increasingly global collaborations for scientific research. In every country that initiated implementing its vision for e-Science, it can be seen that each had its own strategy to face the challenges not only with regards to technical issues such as dependability, interoperability, resource management, etc. but also more on people-centric relating to its collaboration and sharing of its resources and data. For example in the case of United Kingdom (UK), they had established nine e-Science centers and eight other regional centers covering most of UK which primarily aimed to allocate substantial computing and data resources and run standard set of Grid middleware to form the basis for the construction of UK Grid testbed, to generate portfolio industrial Grid middleware and tools and lastly to disseminate information and experience of Grid. [31]

The ideas presented in this paper on both the e-Science models and lifecycle approach will have impact in giving insights, directions and encouragement for policy makers along with valuable contribution to serving the Filipino people especially those scientists and researchers in coming up with technological breakthrough.

Acknowledgements

This research was supported in accordance with requirements of International IT Policy Program (ITPP) and Technology Management, Economics and Policy Program (TEMEP) of Seoul National University in Seoul, South Korea. ITPP and TEMEP a scholarship program made in close cooperation with Ministry of Information and Communication (MIC) of the government of South Korea.

References

[1] The 2007 Microsoft e-Science Workshop at RENCI, https://www.mses07.net/main.aspx

[2] Conceptualizing the Digital Life Cycle, http://iassistblog.org/?p=26

[3] Sustainable Technologies Research Group, http://www.dlsu.edu.ph/research/centers/cesdr/strg.asp

[4] Boonstra, O; Breure, L; Doorn, P; Past, Present and Future of Historical Information Science, Netherland Institute for Scientific Information

Royal Netherlands Academy of Arts and Sciences, 2004.
[5] Berman, F; Hey, A. J. G; Fox, G. C; Grid Computing – Making the Global Infrastructure a Reality, Wiley Series of Communications Networking & Distributed Systems, 2003.
[6] Advanced Science and Technology Institute, R & D agency forearm of Department of Science and Technology (DOST) in the Philippines
[7] http://www.psigrid.gov.ph/index.php
[8] Humphrey, Charles; e-Science and the Life Cycle of Research, IASSIST Communiqué, 2006.
[9] Hayes-Roth, F.; Waterman, D. A; Lenat, D. B; Building Expert Systems, Reading, Addison-Wesley, 1983.
[10] Shadbolt, N. R; Burton, M, Knowledge elicitation: A systematic approach, in evaluation of human work, and Wilson, J. R; Corlett, E. N. (eds) A Practical Ergonomics Methodology, London, UK: Taylor & Francis, 1995.
[11] Hoffman, R.; Shadbolt, N. R.; Burton, A. M; Klein, G., Eliciting knowledge from experts: A methodological analysis. Organizational Behavior and Decision Processes, 1995.
[12] Shadbolt, N. R.; O'Hara, K.; Crow, L., The experimental evaluation of knowledge acquisition techniques and methods: history, problems and new directions, International Journal of Human Computer Studies, 1999.
[13] Milton, N.; Shadbolt, N.; Cottam, H.; Hammersley, M.; Towards a knowledge technology for knowledge management. International Journal of Human Computer Studies, 1999.
[14] Shaw, M. L. G.; Gaines, B. R., WebGrid-II: Developing hierarchical knowledge structures from flat grids. Proceedings of the 11th Knowledge Acquisition Workshop (KAW '98), 1998 Banff, Canada, April, 1998, http://repgrid.com/reports/KBS/WG/.
[15] Crow, L.; Shadbolt, N. R., Extracting focused knowledge from the semantic web. International Journal of Human Computer Studies, 2001.
[16] Ciravegna, F.; Adaptive information extraction from text by rule induction and generalization, Proceedings of 17th International Joint Conference on Artificial Intelligence (IJCAI2001), Seattle, 2001.
[17] Middleton, S. E.; De Roure, D.; Shadbolt, N. R., Capturing knowledge of user preferences: Ontologies in recommender systems, Proceedings of the First International Conference on Knowledge Capture, K-CAP2001. New York: ACM Press, 2001.
[18] Fox, G; e-Science meets computational science and information technology. Computing in Science and Engineering, 2002, http://www.computer.org/cise/cs2002/c4toc.htm.
[19] Taylor, J. M. and e-Science, http://www.e-science.clrc.ac.uk and http://www.escience-grid.org.uk/
[20] W. Wulf; The National Collaboratory – A White Paper in Towards a National Collaboratory, Unpublished report of a NSF workshop, Rockefeller University, New York, 1989.
[21] Kouzes, R. T.; Myers, J. D.; Wulf, W. A.; Collaboratories: Doing science on the Internet, IEEE Computer August 1996, IEEE Fifth Workshops on Enabling Technology: Infrastructure for Collaborative Enterprises (WET ICE '96), 1996, http://www.emsl.pnl.gov:2080/docs/collab/presentations/papers/IEEECollaboratories.html.
[22] Sarmenta, L. F. G; Bayanihan Computing .NET: Grid Computing with XML Web Services, 2002, http://bayanihancomputing.net/
[23] Search for Extraterrestrial Intelligence, http://setiathome.berkeley.edu/sah_about.php
[24] Survival Strategy for Securing International Competitiveness, Korea IT Times, 2007
[25] Ella Tallyn, et.al. Introducing e-Science to the Classroom, ttp://www.equator.ac.uk/var/uploads/Ela2004.pdf
[26] Prof. Iversen, Oxford University e-Science Open Day, 2004
[27] Henry Gardner, et. al, eScience curricula at two Australian universities, Australian National University and RMIT, Melbourne, Australia, 2004
[28] PREGINET is a nationwide broadband research and education network that interconnects academic, research and government institutions. It is the first government-led initiative to establish a National Research and Education Network (NREN) in the country. PREGINET utilizes existing infrastructure of the Telecommunications Office (TELOF) of the Department of Transportation and Communications (DOTC) as redundant links.
[29] David De Roure, et. al; A Future e-Science Infrastructure, 2001
[30] Usability Research Challenges in e-Science, UK e-Science Usability Task Force
[31] Tony Hey and Anne E. Trefethen, The UK e-Science Core Programme and the Grid, ICCS, Springer-Verlag Berlin Heidelberg, 2002