What is Cyberinfrastructure?

Craig A. Stewart*
812-855-4240
Pervasive Technology Institute & Research Technologies, UITS**
stewart@iu.edu

Stephen Simms
812-855-7211
Research Technologies, UITS & Pervasive Technology Institute
ssimms@indiana.edu

Beth Plale
812-855-4373
School of Informatics and Computing & Pervasive Technology Institute
plale@cs.indiana.edu

Matthew Link
812-855-633
Research Technologies, UITS & Pervasive Technology Institute
mrlink@indiana.edu

David Y. Hancock
812-855-4021
Research Technologies, UITS & Pervasive Technology Institute
dyhancoc@indiana.edu

Geoffrey C. Fox
812-856-7977
School of Informatics and Computing & Pervasive Technology Institute
gcf@indiana.edu

*All authors at Indiana University, Bloomington, IN 47405

**UITS is University Information Technology Services

ABSTRACT
Cyberinfrastructure is a word commonly used but lacking a single, precise definition. One recognizes intuitively the analogy with *infrastructure*, and the use of *cyber* to refer to thinking or computing – but what exactly is cyberinfrastructure as opposed to information technology infrastructure? Indiana University has developed one of the more widely cited definitions of cyberinfrastructure:

> **Cyberinfrastructure consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high performance networks to improve research productivity and enable breakthroughs not otherwise possible.**

A second definition, more inclusive of scholarship generally and educational activities, has also been published and is useful in describing cyberinfrastructure:

> **Cyberinfrastructure consists of computational systems, data and information management, advanced instruments, visualization environments, and people, all linked together by software and advanced networks to improve scholarly productivity and enable knowledge breakthroughs and discoveries not otherwise possible.**

In this paper, we describe the origin of the term cyberinfrastructure based on the history of the root word infrastructure, discuss several terms related to cyberinfrastructure, and provide several examples of cyberinfrastructure.

Categories and Subject Descriptors
C.2 COMPUTER-COMMUNICATION NETWORKS.

General Terms
Management, Performance, Design, Security, Human Factors, Standardization.

Keywords
Cyberinfrastructure, e-Science, Infrastructure, Open Science Grid, TeraGrid.

1. INTRODUCTION
The term 'cyberinfrastructure' was coined in the late 1990s, and its usage became widespread in 2003 with the publication of "Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure" by Atkins et al. [1].

The development of the term can be traced directly to foundational work in computer science that resulted in creation of computer grids in the 1990s. In 1995, several leading researchers and supercomputer centers collaborated to demonstrate I-WAY at the IEEE/ACM Supercomputing'95 conference, providing a glimpse of the scientific research opportunities that could be created by linking supercomputers and visualization systems via high-speed networks [2]. Noted researcher Ian Foster subsequently published several articles defining grids, starting with the following definition in 1998 [3]:

> **A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities.**

Definitions of various types and aspects of computer grids were subsequently expanded upon by Foster and others. This was the catalyst for the concept of cyberinfrastructure, which is very much a generalization of the concept of a computational grid. The 2003 report by Atkins et al., however, did not contain a simple and concise definition of the word cyberinfrastructure. As a result, a number of institutions have attempted to craft their own definitions. The purpose of this paper is to present the history of the word cyberinfrastructure, provide two useful working definitions of cyberinfrastructure (each useful in particular contexts), provide examples of cyberinfrastructure in use today, and explain several concepts related to cyberinfrastructure.

2. ORIGIN AND HISTORY

The word cyberinfrastructure is based on the word infrastructure, a relatively recent addition to the English language. A 1927 issue of Chambers's Journal, cited in a list of associated quotations in a 1989 edition of the Oxford English Dictionary, mentions that "The tunnels, bridges, culverts, and 'infrastructure' work generally of the Ax to Bourg-Madame line have been completed" [4]. This seems to be the first known occurrence of the word infrastructure in the English language.

The Merriam-Webster online dictionary defines infrastructure as [5]:

> **Main Entry: in·fra·struc·ture**
>
> **Pronunciation: \'in-frə-ˌstrək-chər, -(ˌ)frä-**
>
> **Function: noun**
>
> **Date: 1927**
>
> **1: the underlying foundation or basic framework (as of a system or organization)**
>
> **2: the permanent installations required for military purposes**
>
> **3: the system of public works of a country, state, or region; also: the resources (as personnel, buildings, or equipment) required for an activity.**

The Oxford Dictionary of English defines infrastructure as [6]:

> **infrastructure (n.)**
>
> **the basic physical and organizational structures and facilities (e.g. buildings, roads, power supplies) needed for the operation of a society or enterprise.**

The word *cyberinfrastructure*, clearly based on the word *infrastructure*, is most simply an infrastructure for knowledge. The first use of the term "cyber-infrastructure" seems to have been in a press briefing [7] about Presidential Decision Directive NSC-63 [8] by Richard Clarke and Jeffrey Hunker. While the Presidential Decision Directive speaks of "cyber-supported infrastructures," the press briefing text mentions "cyber-infrastructure" specifically.

Use of the term cyberinfrastructure within the National Science Foundation (NSF) gained momentum with efforts of Dr. Peter Freeman, Assistant Director of the NSF for Computer and Information Sciences and Engineering. Dr. Freemen wrote in his charge letter to Dr. Daniel Atkins a request for the blue-ribbon panel led by Atkins to make suggestions regarding high priority cyberinfrastructure investments by the NSF. The executive summary of the subsequent report includes the statement, "The newer term cyberinfrastructure refers to infrastructure based upon distributed computer, information and communication technology. If infrastructure is required for an industrial economy, then we could say that cyberinfrastructure is required for a knowledge economy." [1]

Figure 1 below is from Atkins et al. 2003 [1] and was used in that report as part of the explanation therein regarding cyberinfrastructure.

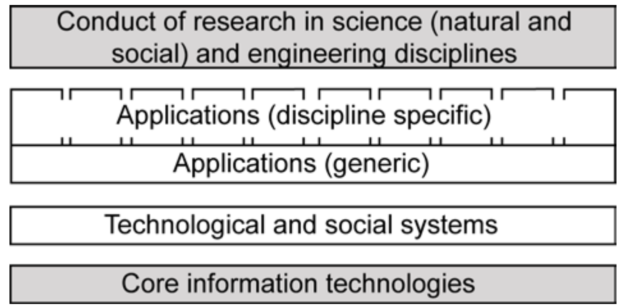

Figure 1. Diagram of relationship of different levels of systems involved in scientific research, from Atkins et al. [1].

Following the above figure in the report is the explanation "Cyberinfrastructure brings together many technologies (hardware, software, processing, storage, communication, etc.) to provide a coherent end-to-end functionality in support of applications; that is, at its heart cyberinfrastructure is a technological system. Many core technologies have themselves a system flavor, but we distinguish technological systems at the top level of hierarchy—where technology meets applications and uses—and observe that systems in this sense have special significance to both cyberinfrastructure and to applications."

While all of these explanations are helpful they do not constitute a definition that allows one to easily distinguish what is cyberinfrastructure from that which it is not.

3. CURRENT WORKING DEFINITIONS OF CYBERINFRASTRUCTURE

In 2005, Fran Berman, then director of the San Diego Supercomputer Center, led an NSF-funded workshop [9] that included the following definition:

> **The component parts of Cyberinfrastructure are human, software, hardware, instrument, and other resources coordinated so as to interoperate "end-to-end" and to support multiple users simultaneously. At scale, this complex structure will need to involve appropriate user incentive structures, effective organizational frameworks, policy and privacy constraints, and a wealth of other social mechanisms to ensure stability, performance, and usefulness.**

Several of the authors of this paper, still thinking that none of the definitions of cyberinfrastructure in existence as of 2007 characterized and distinguished cyberinfrastructure from something that one might simply label "information technology infrastructure," contributed to the development of the following definition [10]:

> **Cyberinfrastructure consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high performance networks to improve research productivity and enable breakthroughs not otherwise possible.**

In a joint report of the EDUCAUSE Campus Cyberinfrastructure Working Group and Coalition for Academic Scientific Computation [11], a broader definition of cyberinfrastructure was offered that explicitly moves beyond research activities to include scholarly activities more generally, inclusive of teaching and learning:

> **Cyberinfrastructure consists of computational systems, data and information management, advanced instruments, visualization environments, and people, all linked together by software and advanced networks to improve scholarly productivity and enable knowledge breakthroughs and discoveries not otherwise possible.**

This is a very modest edit of the 2007 IU definition provided above. The first definition is heavily research-centric and appropriate for use in cutting edge research; the second is more general and is likely to be better in contexts relating to education and scholarship more generally. We note that while we have rendered the word as cyberinfrastructure it is often set in print as CyberInfrastructure and abbreviated CI. Either usage seems fine, and CI seems a useful and easily grasped abbreviation.

The value of any definition lies in its ability to identify that which it calls out and to distinguish those entities that do not meet the definition. The above definitions of cyberinfrastructure are helpful in that they each have three distinct elements relating to technology, people, and effect. Cyberinfrastructure includes technological systems – computing, data storage, advanced instruments, etc. – but these components are not by themselves cyberinfrastructure. A supercomputer may be a component of cyberinfrastructure, but in isolation and by itself does not constitute cyberinfrastructure. By itself, a supercomputer is a supercomputer. The two definitions recommended above explicitly include people as critical elements of cyberinfrastructure and identify two characteristics intended to distinguish cyberinfrastructure from what one might more routinely consider 'information technology infrastructure.' Those two characteristics are: the linkage of technology elements by software and high performance networks into a larger system; and the use of cyberinfrastructure having as its outcome the improvement of research productivity and breakthroughs not otherwise possible. [This last point of emphasis in is drawn directly from emphasis repeatedly made in talks by Dr. Peter Freeman on the topic of cyberinfrastructure.]

4. EXAMPLES OF CYBERINFRASTRUCTURE

Two well-known examples of cyberinfrastructure – the Open Science Grid and the TeraGrid – have been described in detail as examples of large-scale cyberinfrastructure [12, 13]. In order to further clarify the concept of cyberinfrastructure, we present several examples drawn from research and development at Indiana University.

4.1 Simulation of Gas Giant Planet Formation

It has been thought for centuries that planets form from a gaseous nebula surrounding a protostar as the star is forming. However, understanding the details of the formation process remains an outstanding problem in astrophysics.

Astrophysical simulations of gas giant planet formation are performed with a variety of numerical methods. Some of the codes in use today have been producing scientifically significant results for several years, or even decades. In order to produce such results, researchers must simulate millions of resolution elements for millions of time steps, capture and store output data, and rapidly and efficiently analyze this data. With the help of IU's Data Capacitor, geographically distributed supercomputers are employed to accomplish these many tasks [14].

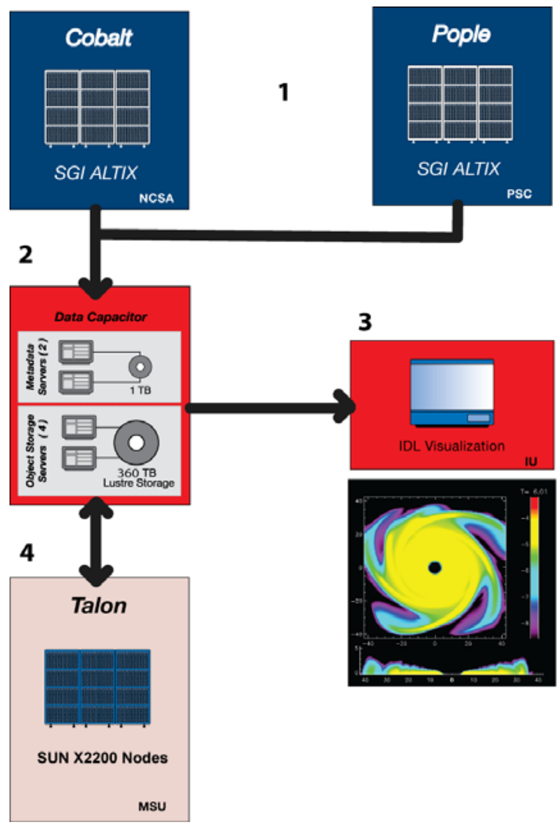

Figure 2. An astronomer (3) views the results of simulations occurring at PSC's Pople and NCSA's Cobalt supercomputers (1) as they are being written to IU's Data Capacitor (2). Further analysis of the data that have been written (2) takes place on MSU's Talon supercomputer.

Figure 2 shows the distinction between the systems that are elements of cyberinfrastructure and cyberinfrastructure as a system that includes humans as an integral part. The critical issue here is the astronomer (in this case Scott Michael of Indiana University) is able to interactively watching the output of simulations running remotely at the Pittsburgh Supercomputing Center. Data from the simulation are written to the Data Capacitor at IU as simulations are being performed, and the researcher can either allowing simulations to run to completion or cancel jobs when simulations produce nonsensical results. In the latter case the astronomer is able to adjust parameters and restart jobs. This is an excellent example of cyberinfrastructure in that it involves supercomputers, data storage systems, and visualization systems all linked by networks (in this case the TeraGrid network [15]) and middleware (in this case Globus and the Common TeraGrid Software Stack [16]), with a researcher as an integral component of the application execution. These advanced simulations of the formation of gas giant planets could not be done without supercomputers, and the linkage of visualization and the scientist increases productivity in ways that enhance research and the efficiency with which resources are used. The linkage does create a performance cost; the application uses supercomputer time 45% less effectively than it would otherwise [17]. However, by enabling interactive monitoring of the simulation as it develops, it is possible to stop simulations when it becomes clear that initial conditions are leading to uninteresting or nonsensical simulations. In such cases, were the simulations to run in batch mode to completion, the performance penalty would be 100% since the computing time would have been wasted.

4.2 Linked Environments for Atmospheric Discovery

The challenges facing society today, in the form of food and water shortages, dependence on foreign oil, and climate change are complex problems requiring solutions that cut across traditional academic departments and discipline boundaries. Research solutions often involve analyzing the boundaries between two or more systems, such as the atmosphere and the ocean, requiring access to more data than needed in studies of either system in isolation. For example, an agricultural researcher might investigate computational methods for determining how frequently farmers need to apply fungicide to minimize potato blight. Applying fungicide only when necessary reduces both the farmers' costs and the environmental impact from fungicide entering the water system through surface runoff. To determine the ideal application frequency, the researcher must incorporate weather forecast data into her predictive models; however, customized weather forecast data products are difficult for a non-meteorologist to use. This is where cyberinfrastructure enters the picture. Cyberinfrastructure can be used to bridge the gap between sophisticated data products and information from one domain and their application in another. The use of the word cyberinfrastructure here means the middleware and layers of tools that sit on top of the computing systems, data storage systems and computer networks to create and deliver information to a researcher that is targeted and formatted in a way that is immediately useful. In this example the agriculture researcher needs a 36-hour weather forecast for each day of the growing season over a 500 square mile region of North Dakota where the forecast provides more detail near the surface and less detail in the higher levels of the atmosphere. Cyberinfrastructure can deliver this specific information, clearly, without extraneous data.

The LEAD II Science Gateway emerged from the NSF-funded Linked Environments for Atmospheric Discovery project [18]. LEAD II supports data search and discovery by crawling Internet weather data sources every five minutes and indexing the data for use in analysis. It pioneered the community account as a way for researchers to execute computationally intensive analysis and modeling tools on TeraGrid resources. Users are provided private data storage for data they create in the gateway, and the data are organized through a metadata catalog [19].

LEAD II provides tools for designing and conducting computational analysis. A user composes a computation by connecting tasks together into a directed graph, usually through a graphical interface (Figure 4). The task graph is then handed to a workflow engine that runs it. The workflows that LEAD II runs are computationally intense, requiring hundreds of CPUs to complete in a few hours runtime and producing gigabytes of data as a result.

When the researcher needs to add new analysis tasks or new versions of a model, she adds these through a web tool [20] that "wraps" the application in a web service wrapper so that it can be invoked through the LEAD II gateway.

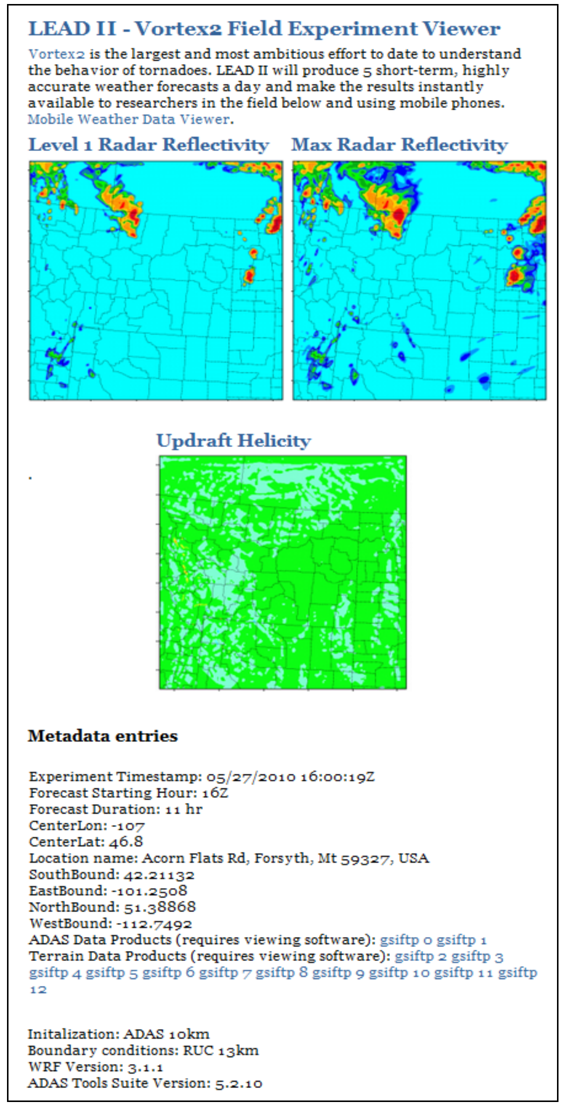

Figure 3. Sample weather forecast created through the LEAD II Vortex2 Field Experiment Viewer.

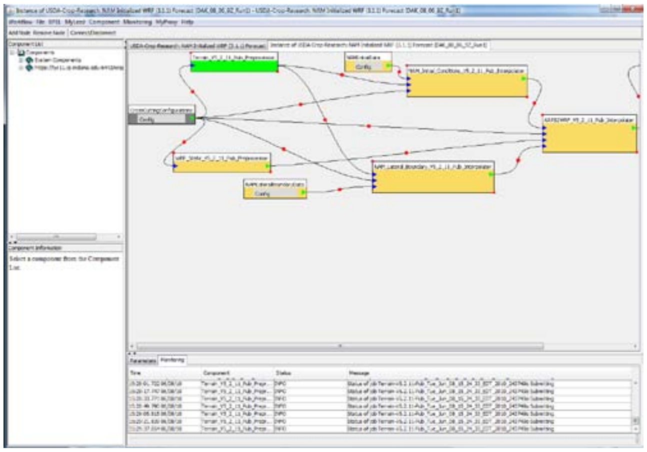

Figure 4. LEAD II graphical workflow interface.

LEAD is also an excellent example of how cyberinfrastructure opens opportunities for education and student-led research activities. LEAD has been used repeatedly in the WxChallenge collegiate weather forecasting competition [21]. Since academic year 2006-2007, thousands of students have used the LEAD science gateway to access supercomputers and forecast weather events. This has led to advancement of education in weather and climate as well as expansion on the number of undergraduate students who have personally used a supercomputer before receiving a baccalaureate degree.

4.3 Polar Grid

Figure 6. Polar Grid equipment in the field.

Scientists report that in the past decade the world's polar ice caps have been steadily and dramatically decreasing. In 2008, scientists revealed that a vast ice sheet in Greenland is melting faster than originally believed and Canada reported that a 19-square mile piece of the 4,500-year old Markham Ice Shelf detached, reducing the ice sheet's size significantly. Polar Grid is a cyberinfrastructure that helps in the global race to understand ice sheet melting. It is a cutting edge, high performance computing infrastructure developed by IU and used by polar scientists working for the Center for Remote Sensing of Ice Sheets (CReSIS) [22]. Traditionally, data collected during polar research expeditions could not be extensively analyzed until expedition scientists returned to their home labs at the end of a season of data collection. Scientists could not evaluate data accuracy and quality until an expedition had ended, and experiments could not be repeated or expanded until returning to the field the following year.

Polar Grid advances polar science research by allowing scientists to begin data analysis while still in the field. This early analysis allows scientists to pinpoint data collection errors and identify areas of interest for further data collection. Scientists can then adjust experiments accordingly, collecting more accurate and useful data before leaving the field. Polar Grid equipment is used to transfer, process, visualize, and store ice sheet data collected from air and surface radar and sensors. Data containing information about an ice sheet at a given point in time is precious; if lost, it cannot be recreated or replaced. For this reason, Polar Grid equipment not only analyzes data, but also provides secure and reliable backup and data storage. Field location equipment includes computational clusters, data storage and networking equipment, servers, and laptops. Equipment is transported and protected from extreme cold, wind, and snow in specially-designed ruggedized cases. The combination of this innovative, field-deployed equipment and satellite linkage of computers and people collecting data on the polar ice sheets with researchers in the US is an excellent example of cyberinfrastructure as a system involving people and enabling innovations and discoveries not otherwise possible. The ability of Polar Grid to enable data analysis and collection in the field and its transmission back to the US has already improved the day-in, day-out collection of precious data about the condition of ice sheets. Analyzing and

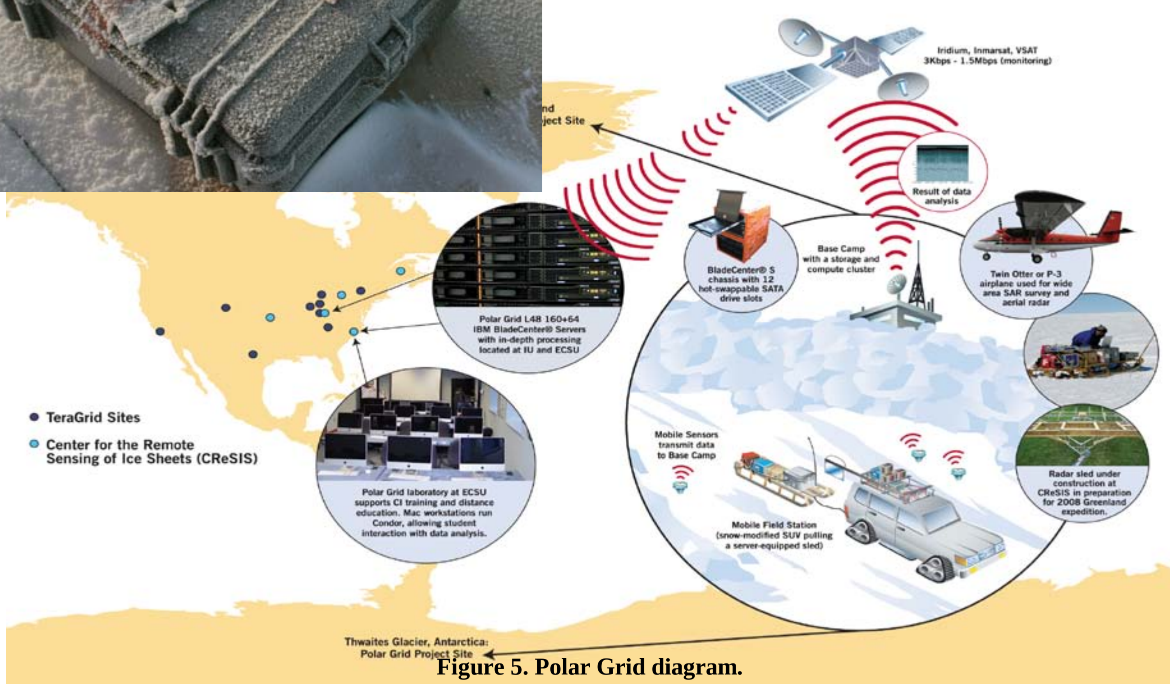

Figure 5. Polar Grid diagram.

understanding synthetic aperture radar images of ice sheets in the field on a real-time, continuous basis was simply not possible prior to Polar Grid.

4.4 FutureGrid

FutureGrid [23] represents a very interesting type of cyberinfrastructure. The project aims to be a "cyberinfrastructure for computational science" by creating a hardware, network, and software environment in which researchers can perform and replicate experimental research, including performance analysis, in computational and computer science. In this case humans are interacting with cyberinfrastructure to create new software for cyberinfrastructure of the future.

5. USAGE OF TERM

Since the use of the term cyberinfrastructure in the report by Atkins et al. this term has been used often. Citeseerx [24] provides a list of 373 publications in response to a search on the word cyberinfrastructure. Most recently, the NSF has announced a major new initiative called the "Cyberinfrastructure for 21st century research" that will help define NSF strategies relative to NSF-funded research and cyberinfrastructure [25]. While the word cyberinfrastructure has clearly gained currency in academic circles, there remain questions about how the term will be used and how the things we now refer to with the word 'cyberinfrastructure' will develop over time. Usage within the NSF has focused on cutting-edge capabilities. In such usage it is accepted that in the effort to enable research that would not otherwise be possible, cyberinfrastructure will sometimes produce innovative results and sometimes analyses will fail. However, as noted in the report "Understanding Infrastructure: Dynamics, Tensions, and Design – report of a workshop on history & theory of infrastructure: lessons for new scientific cyberinfrastructures" [26] there is a tension between desire for new capability and desire for reliability. In the excitement of exploration, one might well accept a level of uncertainty about whether or not cyberinfrastructure will work properly that one would not so gladly tolerate in the function of traditional infrastructure components such as plumbing or electricity. At the same time, there is value in the general and functional focus of the definitions we present here for cyberinfrastructure. By being focused on the general function of a system of technology, direct involvement of people, and innovation as an outcome, this definition avoids the current mix of confusion and hype one sees in cloud computing, in which some clearly important concepts and new ideas become mixed with a proliferation of "{Insert most any noun} as a service."

6. OTHER DEFINITIONS AND RELATED TERMS

There are several terms related to cyberinfrastructure that merit mention.

A *Science Gateway* is a particular type of cyberinfrastructure – an interactive interface that provides end-to-end support for a particular scientific workflow or set of tasks used by a virtual organization or scientific community. Science Gateways are defined as "a community-developed set of tools, applications, and data that is integrated via a portal or a suite of applications, usually in a graphical user interface, that is further customized to meet the needs of a targeted community" [27]. According to Gannon et al. [28] these kinds of cyberinfrastructures have five common components:

1. Data search and discovery.
2. Security.
3. User private data storage.
4. Tools for designing and conducting computational analysis, that is, "workflow" tools.
5. Tracking data provenance.

LEAD II is an excellent example of a Science Gateway. Another example is the NanoHUB project, which provides computational tools for nanotechnology and is led by Purdue University [29]. The role of Science Gateways in expanding usage of cyberinfrastructure is discussed at some length in [30].

A related concept is *e-Science*, a European term defined by Malcolm Atkinson as "denot[ing] the systematic development of research methods that exploit advanced computational thinking" [31]. e-Science is sometimes described as equivalent to cyberinfrastructure, but this seems not exactly right. e-Science has a sense of being more about cyber-enabled science and somewhat less about the underlying infrastructure. This sense is also indicated by a separate definition of e-Infrastructure on the e-Science web page as "denot[ing] the digital equipment, software, tools, portals, deployments, operational teams, support services and training that provide computational services to researchers" [31]. This definition is very close to the definitions of cyberinfrastructure given here.

Another related term is *cybersecurity*, defined in the Merriam-Webster dictionary as [32]:

> **Main Entry: cy·ber·se·cu·ri·ty**
>
> **Pronunciation: \-si-ˌkyu̇r-ə-tē**
>
> **Function: *noun***
>
> **Date: 1994**
>
> **measures taken to protect a computer or computer system (as on the Internet) against unauthorized access or attack**

Cybersecurity is thus very general – applying to computers and computer systems generally, not just those systems that would be included in the definition of cyberinfrastructure given in this report.

7. CONCLUSIONS

The term *cyberinfrastructure* has become heavily used within the US advanced research and advanced information technology community. We believe there are two definitions that are useable and useful, depending on context:

> **Cyberinfrastructure consists of computing systems, data storage systems, advanced instruments and data repositories, visualization environments, and people, all linked together by software and high performance networks to improve research productivity and enable breakthroughs not otherwise possible.**
>
> **Cyberinfrastructure consists of computational systems, data and information management, advanced**

instruments, visualization environments, and people, all linked together by software and advanced networks to improve scholarly productivity and enable knowledge breakthroughs and discoveries not otherwise possible.

Any definition is useful to the extent that it identifies something important and distinguishes it from other things. The definitions above, supplemented by the examples presented here, will we hope help further explain this important concept and aid the development of new cyberinfrastructure technology, features, and functionality that will advance research, education, and workforce development in the US and throughout the world.

8. ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation through grants OCI-0116050, OCI-0338618, OCI-0451237, OCI-0535258, OCI-0504075, CNS-0521433, CNS-0521433, OCI-0721656, OCI-072176, CNS-0723054, OCI-0910812, AGS-0331480, OCI-1002526, and OCI-0928542. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Stewart wishes to thank particularly Thomas Sterling for discussions of the concept of cyberinfrastructure and the importance of discriminating power of such a definition; that discussion was the origin of this paper. Stewart acknowledges the helpful discussions of the CASC/EDUCAUSE joint workshop including the organizers (Patrick Dreher, Vijay Agarwala, Stan Ahalt, Guy Almes, Sue Fratkin, Thomas Hauser, Jan Odegard, and Jim Pepin) and all participants which led to one of the two definitions of cyberinfrastructure discussed in this paper and which informed the creation of this paper as well.

9. REFERENCES

[1] Atkins, D.E., K.K. Droegemeier, S.I. Feldman, H. Garcia-Molina, M.L. Klein, D.G. Messerschmitt, P. Messina, J.P. Ostriker, and M.H. Wright. Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. 2003. Available from: http://www.nsf.gov/od/oci/reports/toc.jsp [cited 7 Jun 2010]

[2] Korab, H. and M.D. Brown. (eds) 1995. Virtual Environments and Distributed Computing at SC'95: GII Testbed and HPC Challenge Applications on the I-WAY. SUUPERCOMPUTING '95, December 3-8 1995, San Diego, CA. Published by ACM, 1515 Broadway, New York, NY 10036-5701

[3] Foster, I. and C. Kesselman, eds. The Grid: Blueprint for a New Computing Infrastructure. 1998, Morgan Kaufmann.

[4] "infrastructure." In: Oxford English Dictionary (second edition). 1989. Available from: http://dictionary.oed.com/cgi/entry/50116561?single=1&query_type=word&queryword=infrastructure&first=1&max_to_show=10 [cited 7 Jun 2010]

[5] "infrastructure." In: Merriam-Webster Online Dictionary. Available from: http://www.merriam-webster.com/dictionary/infrastructure [cited 7 Jun 2010]

[6] "infrastructure noun." In: The Oxford Dictionary of English (revised edition). C. Soanes and A. Stevenson, ed. Oxford University Press. 2005. Available from: http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t140.e38373 [cited 7 Jun 2010]

[7] Clarke, R. and J. Hunker. Press Briefing by Richard Clarke, National Coordinator for Security, Infrastructure Protection, and Counter-Terrorism; and Jeffrey Hunker, Director of the Critical Infrastructure Assurance Office. 1998. Available from: http://www.fas.org/irp/news/1998/05/980522-wh3.htm [cited 7 Jun 2010]

[8] Clinton, B. Presidential Decision Directive/NSC-63. 1998. Available from: http://www.fas.org/irp/offdocs/pdd/pdd-63.htm [cited 7 Jun 2010]

[9] Berman, F. and H. Brady. Workshop on Cyberinfrastructure for the Social and Behavioral Sciences: FINAL REPORT. [PDF] 2005. Available from: http://www.sdsc.edu/about/director/pubs/SBE/reports/SBE-CISE-FINAL.pdf [cited 8 June 2010]

[10] Pervasive Technology Institute at Indiana University. IU Cyberinfrastructure News. March 2007. Available from: http://pti.iu.edu/ci/iu-cyberinfrastructure-news-march-2007 [cited 8 June 2010]

[11] EDUCAUSE Campus Cyberinfrastructure Working Group and Coalition for Academic Scientific Computation. Developing a Coherent Cyberinfrastructure from Local Campus to National Facilities: Challenges and Strategies. February 2009. Available from: http://www.educause.edu/Resources/DevelopingaCoherentCyberinfras/169441 [cited 8 Jun 2010]

[12] Avery, P., "Open science grid: Building and sustaining general cyberinfrastructure using a collaborative approach." First Monday, 2007. 12(6). Available from: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1866/1749 [cited 9 Jun 2010]

[13] Simms, S.C., C.A. Stewart, and S.D. McCaulay. "Cyberinfrastructure resources for U.S. Scholarship: the TeraGrid." In: Proceedings of the 36th annual ACM SIGUCCS conference on User services. pp. 341-344. 2008. Portland, OR, USA: ACM. Available from: http://portal.acm.org/citation.cfm?id=1450057 [cited 8 Jun 2010]

[14] Walgenbach, J., S.C. Simms, J.P. Miller, and K. Westneat, Enabling Lustre WAN for Production Use on the TeraGrid: A Lightweight UID Mapping Scheme (accepted), in TeraGrid'10. 2010: Pittsburgh, PA.

[15] TeraGrid. Home page. Available from: https://www.teragrid.org/ [cited 8 Jun 2010]

[16] TeraGrid. CTSS 4 TeraGrid Core Integration Implementation. Available from: http://www.teragridforum.org/mediawiki/index.php?title=CTSS_4_TeraGrid_Core_Integration_Implementation [cited 30 Apr 2010]

[17] Michael, S., S. Simms, Breckenridge, Smith, and M. Link, A Compelling Case for a Centralized Filesystem on the TeraGrid (accepted), in TeraGrid'10. 2010: Pittsburgh, PA.

[18] Droegemeier, K.K., D. Gannon, D. Reed, B. Plale, J. Alameda, T. Baltzer, K. Brewster, R. Clark, B. Domenico, S. Graves, E. Joseph, D. Murray, R. Ramachandran, M. Ramamurthy, L. Ramakrishnan, J.A. Rushing, D. Weber, R. Wilhelmson, A. Wilson, M. Xue, and S. Yalda, "Service-Oriented Environments for Dynamically Interacting with Mesoscale Weather." Computing in Science and Engg., 2005. 7(6): p. 12-29.

[19] Jensen, S. and B. Plale. "Schema-Independent and Schema-Friendly Scientific Metadata Management." In: Proceedings 4th Int'l IEEE Conference on e-Science. pp. 428-429. 2008: IEEE Computer Society. Available from: http://doi.ieeecomputersociety.org/10.1109/eScience.2008.114 [cited 8 Jun 2010]

[20] Kandaswamy, G., L. Fang, Y. Huang, S. Shirasuna, S. Marru, and D. Gannon, "Building Web Services for Scientific Grid Applications." IBM Journal of Research and Development, 2006. 50(2/3): p. 249-260.

[21] University of Oklahoma. WxChallenge. Available from: http://www.wxchallenge.com/ [cited 9 Jun 2010]

[22] Center for Remote Sensing of Ice Sheets (CReSIS). Home page. Available from: https://www.cresis.ku.edu/ [cited 9 Jun 2010]

[23] FutureGrid. Home page. Available from: http://futuregrid.org/ [cited 21 July 2010]

[24] CiteSeer. Home page. Available from: http://citeseerx.ist.psu.edu/ [cited 9 Jun 2010]

[25] National Science Foundation. Dear Colleague Letter: Cyberinfrastructure Framework for 21st Century Science and Engineering (CF21). 2010. Available from: http://www.nsf.gov/pubs/2010/nsf10015/nsf10015.jsp?org=OCI [cited 9 Jun 2010]

[26] Edwards, P.N., S.J. Jackson, G.C. Bowker, and C.P. Knobel. Understanding Infrastructure: Dynamics, Tensions, and Design. January 2007. Available from: http://www.si.umich.edu/cyber-infrastructure/UnderstandingInfrastructure_FinalReport25jan07.pdf [cited 19 Jan 2010]

[27] TeraGrid. Science Gateways Home. Available from: https://www.teragrid.org/web/science-gateways/ [cited 9 Jun 2010]

[28] Gannon, D., B. Plale, M. Christie, Y. Huang, S. Jensen, N. Liu, S. Marru, S.L. Pallickara, S. Perera, S. Shirasunam, Y. Simmhan, A. Slominski, Y. Sun, and N. Vijayakumar, Building Grid Portals for e-Science: A Services Oriented Architecture, ed. L. Grandinetti, Ed. . 2007, Amsterdam: ISO Press.

[29] Network for Computational Nanotechnology. Simulation, Education, and Community for Nanotechnology. Available from: http://nanohub.org/ [cited 8 Jun 2010]

[30] Zimmerman, A. and T.A. Finholt, Growing an infrastructure: the role of gateway organizations in cultivating new communities of users, in Proceedings of the 2007 international ACM conference on Supporting group work. 2007, ACM: Sanibel Island, Florida, USA. p. 239-248.

[31] Research Councils UK. e-Science. Available from: http://www.rcuk.ac.uk/escience/default.htm [cited 8 Jun 2010]

[32] "cybersecurity." In: Merriam-Webster Online Dictionary. Available from: http://www.merriam-webster.com/dictionary/cybersecurity [cited 9 Jun 2010]