"...how is it that I can go from point A to point B in my application or any area in the domain of information technology faster than before? Am I, or rather, the things I control, moving faster? Or is the space of information technology shrinking?" – Nathaniel Davis
One of the highlights of the year for me as editor of the *Bulletin* is the annual information architecture (IA) issue, which primarily comprises presentations from the ASIS&T IA Summit, in this case the 11th Summit, held in Phoenix in April. There are no formal proceedings for the summits, so it is our pleasure to make some of these presentations more widely available each year. This year’s guest editor, the *Bulletin*’s associate editor for IA, Thom Haller, has elected to return to the theme of definitions of IA and the nature of IA as a field. “What is Information Architecture? Practical Definitions and Useful Principles in Our Second Decade of Work and Study” includes four papers in addition to Thom’s introduction. In thoughtful presentations Dan Klyn, Andrea Resmini and Keith Instone, Nathaniel Davis and Dan Brown share with us a variety of views and approaches to contemplate.

I am also pleased to have a second feature article from Marko Rodriguez, this time with co-author Peter Neubauer. Marko is a specialist in graph databases and graph theory, and, as such, is very mathematically oriented in his work. But graph theory is an important science underlying a great deal of the current development in data structures and databases, for example, for RDF – the subject of Marko’s previous article – and Marko is fortunately willing to undertake the difficult task of making some of this material accessible. If you have wondered how graph representations of RDF are processed or what comes next in database development after the relational model, this article is for you. When I was teaching in the area of contemporary knowledge organization, I would have welcomed it wholeheartedly as a valuable reading, and I hope many of you will share that view.

We also have reports on two other significant ASIS&T activities. First, the IA Summit was not the only ASIS&T event held in Phoenix. The maiden voyage of the ASIS&T Research Data Access and Preservation Summit addressed a field of growing importance in information science and reflected the leading role that ASIS&T members have played in its development. We are very happy to have a summary of the summit prepared by Reagan Moore and William L. Anderson.

In the other significant event, also in April, the ASIS&T Information Professional Task Force “went international” as Nancy Roderer and Gail Bonath report, when Nancy Roderer lead a People to People delegation to China to “discuss the education and promotion of information professionals.” The location of some Chinese LIS schools within information institutions was one interesting feature of the Chinese experience that the delegation explored – along, of course, with quite a lot of excellent fun and food.

Finally, ASIS&T President Gary Marchionini reflects on his President’s Page about the nature of the Society as we approach our 75th anniversary in 2012, a precursor of things to come.
Over the course of this year’s activities I have had several opportunities to revisit fundamental themes that recur in all professional societies. What is the American Society for Information Science and Technology? Whom does it serve? What problems does it address? Work this year by the publications, membership and Annual Meeting program committees, as well as discussions in my presidential conference calls and at the various meetings, have each given rise to fundamental questions about mission, focus and participation.

A look at the “About ASIS&T” page (http://asist.org/about.html) on the ASIS&T website highlights our 73-year history as a society devoted to improving information access. My sense of what ASIS&T is today is somewhat broader in that for me, our mission is to study information – its genesis, propagation, use and effects over time. We tend to be particularly interested in information’s effects on people and institutions and in the technologies and techniques associated with it. I claim that ASIS&T is the world’s leading society in information science and a home for information science research and the practices that stem from research findings. As a professional society, we aim to promote the information professions through evidence-based research and innovative development. Thus, as researchers, we investigate problems that arise in any stages of the entire information life cycle; and as practitioners, we aim to apply the results of systematic investigations to invent effective and efficient solutions to those information problems. Some ASIS&T members focus more on research and some on practice, but all ASIS&T members are made of the same metal when it comes to understanding and solving information problems. As a result of the recent meeting of the program committee for the upcoming Annual Meeting, I am optimistic that we can forge a program that is well struck on both sides of this coin.

A second musing about the nature of ASIS&T arose as I participated in the LIDA conference in Zadar, Croatia, in May. I was struck by the level of excitement and engagement shown by European ASIS&T members. There was a doctoral forum run by the European ASIS&T Student Chapter (well organized by doctoral student Cathal Hoare from University College Cork, Ireland) that highlighted
research done in several Eastern European universities. I hope that there will be a follow-up doctoral forum next year. Students and faculty outside of North America recognize ASIS&T as the leading information science society and aspire to participate actively.

Additionally, as I read JASIST, I am struck by the international scope of coverage and authorship. Papers from non-North American authors dominate many issues of JASIST. For example the May and June 2010 issues have 30 full papers, only 13 of which are from authors from North American affiliations (and one of these includes an author from Asia). Looking back at May and June issues from 1996 there were nine full papers, four from North America and five from other parts of the world. If we go back another 20 years to the March-April and May-June 1976 issues we find eight of 10 papers from authors with North American affiliations, but two that are not. This quick perusal of the collection on my shelf (yes, paper), demonstrates the long-standing globalization of our primary research publication.

These observations of meetings and of our publications validate how international ASIS&T is in fact, if not in name. Indeed, ASIS&T is increasingly international in scope and participation. As the leading information science society, we attract a broad range of participants from around the world and have an obligation to advance research and practice related to information science globally.

As we head toward our 75th anniversary (2012) it is important that we reflect on our mission as the leading society in information science to balance discovery and application and leverage information technologies to expand global participation in ASIS&T.
Balloting for New ASIS&T Officers and Directors Underway

It’s election season again for the American Society for Information Science and Technology. Electronic balloting is now underway for the election of a president-elect, a treasurer and two directors-at-large. Voting ends September 30, 2010, with newly elected officers and directors taking their seats on the Board of Directors at the conclusion of the Annual Meeting in Pittsburgh.

Candidates for president-elect for 2011 and succession to the presidency in 2012 are Diane Sonnenwald and Howard Rosenbaum. Candidates for treasurer for the 2011-2013 administrative years are Vicki Gregory and Victor Rosenberg. Candidates for two director-at-large slots for three-year terms – 2011-2013 – are Katriina Byström, Linda Rudell-Betts, Marcia Lei Zeng and Elaine Toms.

Candidates for President-Elect

Howard Rosenbaum is associate dean and associate professor in the School of Library and Information Science (SLIS) at Indiana University, Bloomington. He serves as director of the master of information science program and co-director of the graduate certificate in information architecture program. He studies social informatics, online communities and the uses of social theory in information science. Rosenbaum has presented his work at ASIS&T and numerous other organizations. He teaches classes on digital entrepreneurship, information systems design and intellectual freedom. He has been recognized for excellence in teaching and for the innovative use of technology in education. He is the recipient of the Frederic Bachman Lieber Memorial Award for Teaching Excellence, IU (2005), a statewide MIRA Award for Technological Innovation in Education (2003) and the Indiana Partnership for Statewide Education Award for Innovation in Teaching with Technology (2002).

Diane H. Sonnenwald is head of school and professor at the School of Information and Library Studies at University College Dublin. She is also an adjunct professor in computer science at the University of North Carolina. She holds a Ph.D. from the School of Communication and Information at Rutgers University, M.S. from Montclair State University and B.A. from Muhlenberg College. She was a postdoctoral fellow at Risø National Labs (Denmark) and a Fulbright professor in Information Studies at the University of Tampere, Finland. Prior to joining academia she worked at Bellcore. Her research has been funded by the National Library of Medicine, National Institutes of Health, National Science Foundation and European Science Foundation, among others. She is the recipient of the ASIS-ISI Doctoral Dissertation Award, U.S. Army Research Laboratory Scientific Contribution Award, ALISE Research Methodology Best Paper Award and Bellcore Award of Excellence. She serves on the editorial boards of JASIS&T, JELIS, Information Research and the Journal of Library and Information Science (Taiwan).

Candidates for Treasurer

Vicki L. Gregory, professor at the School of Library and Information Science, University of South Florida, earned her doctorate at Rutgers University and holds an M.A. and M.L.S from the University of Alabama. Since becoming a member of ASIS&T in 1984, she has served as the Florida chapter president and on the national level has been a member of SIG/LAN and SIG/DL, serving as chair of SIG/LAN. She was elected Deputy SIG Cabinet Director and then SIG Cabinet Director. For the last three years she has served as ASIS&T...
treasurer, following four years a member of the Budget and Finance Committee. She also currently serves as chair of the ALA Committee on Accreditation.

**Victor Rosenberg** is associate professor in the School of Information at the University of Michigan. Formerly, he was chairman and CEO of Personal Bibliographic Software and the developer of ProCite and related software. For 15 years he was responsible for the financial health of a small business. He is the author of numerous papers, films and software applications. His interests include information retrieval, information policy and entrepreneurship. He taught at the University of California at Berkeley after getting his doctorate in library science from the University of Chicago and a masters degree in information science from Lehigh University.

**Candidates for Directors-at-large**

**Katriina Byström** is associate professor/reader at the Swedish School of Library and Information Science, University of Borås in Sweden. In teaching and research she focuses on task-based information seeking and retrieval in workplaces and on information architecture. She holds a Ph.D. from the University of Tampere, Finland, and is an active member of the academic community of LIS within teaching, research and administration. She has been the director of the Swedish School of Library and Information Science and served on numerous committees. She is co-founder and associate editor of the international *Journal of Information Architecture* and has a broad experience serving as a peer-reviewer in a number of high-standard journals. She has organized international academic events and has curriculum development experience. Furthermore, she works as a senior researcher in a private company. At present, she heads two research projects: Better Search Engine focusing on work task based search support and Better Web with focus on the development of digital information and communication milieus.

**Linda Rudell-Betts** has been an ASIS&T member since 1991. She has served the award-winning Los Angeles Chapter as publicity coordinator, secretary, chair and now continues as membership committee chair. Her ASIS&T national involvement includes Chapter Assembly committee advisor (1998 to 2002, 2008 to present), *Bulletin* Advisory Board (2001 to present), Annual Meeting planning committee member (2001 and 2003), Annual Meeting papers and posters reviewer (2005, 2006, 2008). She is a reference librarian at the Los Angeles Public Library, where she does her utmost to pass on her version of best practices in information retrieval to her patrons. Prior to her current public service position, she was an information science consultant, designing, compiling and managing information retrieval vocabularies for periodical databases, Internet software applications and records management systems. She received the ASIS&T Chapter Member-of-the-Year award in 1999. She holds an M.L.S. from UCLA and a B.A. from the University of Texas at Austin in French literature. To unwind from her day, she plays bassoon with a Santa Monica community band.

**Elaine Toms** is professor and Canada Research Chair in Management Informatics and runs the iLab at the Faculty of Management, Dalhousie University, Halifax, Nova Scotia, Canada. She was formerly an associate professor in the Faculty of Information Studies, University of Toronto, and at the School of Library and Information Studies (now School of Information Management) at Dalhousie University. She holds a B.A. in economic geography and education from Memorial University, St. John’s Newfoundland; M.L.S. from Dalhousie University; and a Ph.D. from the University of Western Ontario, London. Her research interests lie at the intersection of human computer interaction, information retrieval and the representation and presentation of information. Her work has been funded by NSERC, SSHRC, OCLC, Heritage Canada, Canada Foundation for Innovation and the Canada Research Chairs Program.
She was/is co-investigator with three Canadian national research networks. Her publications have appeared in journals such as the *International Journal of Human Computer Studies*, the *Journal of the American Society for Information Science and Technology* and *Information Processing & Management*, as well as in the proceedings of a number of national and international associations.

Marcia Lei Zeng, professor of library and information science at Kent State University, has been a member of the faculty since 1992. She holds a Ph.D. from the University of Pittsburgh and an M.A. from Wuhan University in China. She was a visiting associate professor at Columbia University during her sabbatical in 1999 and 2000. Her major research interests include semantic technologies and knowledge organization systems, metadata, database quality control and multilingual information processing. Her scholarly publications include more than 60 papers and three books. She has chaired and served on standards committees and working groups for several national and international library and information organizations. She is also a member of the executive board of the International Society for Knowledge Organization and the advisory board of the Dublin Core Metadata Initiative and an invited expert on the W3C Library Linked Data Incubator Group. She received the ASIS&T Doctoral Forum Award for outstanding doctoral research in 1992. She has served ASIS&T as chapter chair and officer, *JASIST* and conference proceedings referee, awards juror, SIG workshop and session organizer and ASIS&T Annual Meeting program committee member. She has been the chair of the ASIS&T Standards Committee and the voting representative for NISO since 2006.

### Additional Candidate Information

For more information about the candidates and to review their position statements defining their campaigns, please visit the ASIS&T website. All current ASIS&T members are eligible to vote for officers and directors. Full information on the electronic balloting process is available at the ASIS&T website at [www.asis.org/elections/](http://www.asis.org/elections/).
academic programs and student services for the graduate school at UWM.

**Joseph A. Busch**, 2001 ASIS&T president and founder of Taxonomy Strategies, has announced that he will move to the Washington, DC, area as a result of the merging of his company with McLean, Virginia-based Project Performance Corporation (PPC). Taxonomy Strategies, a leader in intranet, public website and e-commerce taxonomy applications, will join forces with PPC which specializes in taxonomy design, knowledge management and systems integration services.

**Joan K. Lippincott** of the Coalition for Networked Information will deliver the keynote speech at the 2nd International Symposium on Information Management in the Changing World, to be held from September 22-24, 2010, in Ankara, Turkey. She will talk about the impact of social networks and mobile devices in creating and maintaining personal information environments. She will stress that it is vital to understand user needs and behaviors if libraries, archives and museums are to be an integral part of users’ information habitats.

**José-Marie Griffiths**, 1993 ASIS&T president and most recently professor at the University of North Carolina at Chapel Hill, has been named vice president of academic affairs at Bryant University in Smithfield, Rhode Island. In her new position, she is responsible for all academic and research programs within the university’s colleges of business and arts and sciences. She will play a major role in the university’s upcoming strategic planning process.

**Jane Greenberg**, professor and director of the Metadata Research Center at the University of North Carolina at Chapel Hill, has been awarded the 2010 Jesse H. Shera Award for Distinguished Published Research by the Library Research Round Table of the American Library Association (ALA). Greenberg wrote the winning paper, “Theoretical Considerations of Lifecycle Modeling: An Analysis of the Dryad Repository Demonstrating Automatic Metadata Propagation, Inheritance and Value System Adoption.” The article explores lifecycle modeling for understanding metadata and reports the results of two extensive studies. The Shera award provides “recognition and monetary support for research employing exemplary research design and methods.”

**Barbara B. Moran**, professor, School of Information and Library Science at the University of North Carolina at Chapel Hill, has been appointed the first Louis Round Wilson Distinguished Professor effective July 1, 2010. The professorship is the result of a gift from the estate of the School’s founder, Louis Round Wilson, and matching funds from the state of North Carolina. Moran joined the faculty at SILS in 1981 and previously held the position of dean from 1990 to 1998. Moran has taught primarily in the areas of management and academic librarianship, and her research has focused on various aspects of management including leadership, organizational structures and career progression patterns. In addition, she continues to direct the school’s international programs where she has expanded the global outreach of SILS, most recently developing a new program to London, England.

**Richard E. Rubin**, director of Kent State University’s School of Library and Information Science since 1999, is now associate provost for extended education. In this role he will oversee the Office of Continuing and Distance Education, University Press and Summer Programs. Rubin had previously served as interim dean for Kent State’s College of Communication and Information, associate professor and assistant professor at Kent State. Before joining Kent State, Rubin served as visiting assistant professor for the Graduate School of Library and Information Science and research assistant for the Library Research Center of the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign.

**News from an Institutional Member**

The **University of North Texas Department of Library and Information Sciences** received a grant from the Robert and Ruby Priddy Charitable Trust to assist rural public libraries in Texas. Funding will support a three-year project, Promoting & Enhancing the Advancement of Rural Libraries (PEARL). The goal of the project is to enhance the role of public libraries in targeted rural communities in Texas. Five
The New England ASIS&T (NEASIS&T) Chapter announces that the 2010 NEASIS&T Best Paper in Information Science has been awarded to Danielle M. Cossarini, of Dalhousie University School of Information Management, for her paper entitled, “Communicating Scientific Information for Environmental Solutions: A Knowledge Management Perspective.” (A copy of Danielle’s paper is available on the NEASIS&T website at http://neasist.org/2010/06/04/2010-neasist-student-travel-award-competition-winner/). She will receive up to $750 to help defray the costs of attendance at the 2010 ASIS&T Annual Meeting in Pittsburgh.

News from ASIS&T Chapters

Patrice A. Clemson

PATRICE A. CLEMSON, longtime ASIS&T member and most recently on the faculty of the College of Information Sciences and Technology at Penn State University’s Beaver Campus, passed away on July 6. Born in Bethlehem, Pennsylvania, she began her college studies at Muhlenberg College where she earned a degree in political science. She later received her master’s and doctorate degrees from the University of Pittsburgh’s School of Information Sciences.

She spent most of her career as faculty or staff in academic institutions. She began at the New York University Graduate School of Business; subsequent positions were at the Penn State University Pattee Library, the University of Pittsburgh’s Graduate School of Public and International Affairs, Duquesne University and the University of Illinois Graduate School of Library and Information Science. She was also the University of Pittsburgh’s data administrator in the Computer Services and Systems Development department.

Patrice believed strongly that university faculty in her field should have practical experience in industry. Consequently, she also was employed for several years in the information services and data processing divisions of National Data Corporation and Pittsburgh National Corporation. She was also a talented fiber artist and weaver and an accomplished musician.

Patrice is survived by her brother Commander James F. Clemson, USN (Ret.), her daughter Meredith G. Hennon, her son William P. Hennon, and her husband John G. Hennon.

The International Calendar of Information Science Conferences (icisc.neasist.org/) is a nonprofit collaboration between the Special Interest Group/International Information Issues (SIG/III) and the European (ASIST/EC) and New England (NEASIST) chapters of the American Society for Information Science and Technology, with the additional support of Haworth Press.
The work of the ASIS&T Information Professional Task Force went international this year with Nancy Roderer’s leadership of a People to People delegation to China to discuss the education and promotion of information professionals. People to People (www.peopletopeople.com/) was founded by President Eisenhower in 1956 to foster direct interaction between citizens of the United States and other countries, and it sponsors a variety of programs for both students and professionals.

From April 17-26, 2010, 10 information professionals from the United States, Canada and Japan visited five libraries and information professional schools in Beijing and Shanghai. Members of the delegation included Nancy K. Roderer, director, Welch Medical Library, Johns Hopkins University; Gwen Alexander, dean, School of Library and Information Management, Emporia State University; Sandy Arbuthnot, web librarian, Toronto Public Library; Lori Beaudoin, public services librarian, United States Army Medical Research Institute for Infectious Diseases Library; Gail Bonath, librarian, Grinnell College; Pascal Calarco, head, Systems Department, University of Notre Dame; Anne S. Caputo, executive director, Dow Jones & Co., and president, Special Libraries Association; Rumi Graham, librarian, University of Lethbridge, Canada; Ann Prentice, professor emerita and former dean, College of Information Studies, University of Maryland, College Park; and Yukiko Sakai, associate manager, Keio University Library, Tokyo, Japan. Nancy Roderer and Ann Prentice are members of the Information Professional Task Force.

It was a special pleasure to have Anne Caputo, president of the Special Libraries Association, as a member of the delegation. Anne has been heavily involved with ASIS&T in the past and was able to contribute the perspective of SLA to the conversations. One relevant task that SLA has undertaken is the description of the competencies of information professionals – see the document, “Competencies for Information Professionals of the 21st Century,” available at www.sla.org/content/learn/members/competencies/index.cfm

The professional program for the delegation consisted of visits to three institutions in Beijing and two in Shanghai. The three in Beijing were the National Science Library, Chinese Academy of Sciences (NSLC); the Institute of Scientific and Technical Information of China (ISTIC); and Beijing Normal University and College of Information Science and Technology. The two in Shanghai were the Shanghai Municipal Archives and the East China Normal University Library.

Each of these institutions (with the exception of the Shanghai Municipal Archives) includes both a library and an information professional degree-granting program. The delegation had the opportunity to meet with both librarians and library school faculty at each institution.

NSLC is the library of the Chinese Academy of Sciences, which is the national academy for natural sciences in China. NSLC serves 30,000 researchers and 40,000 graduate students throughout China. The director of NSLC, Zhang Xiaolin, gave a presentation to the delegation titled, “Toward Knowledge-Based Information Professionals.” He detailed the National Science Library’s promotion of the embedded subject librarian, a change in service model, which their institution made in 2006. The embedded subject librarian has a research-based masters or doctoral degree in a
Subject librarians’ capabilities are not bound by the library’s resources.

Subject librarians share the users’ culture.

Subject librarians are capable of analyzing, developing and designing customized services.

Subject librarians’ capabilities are not bound by the library’s resources.

Next, we visited the Institute of Scientific and Technical Information of China (ISTIC). ISTIC is a nonprofit research center under the Ministry of Science and Technology, and it serves both the government and the general public. Dr. Zhao Zhiyun, deputy director of ISTIC, provided an overview of the institution for us. Like NSLC, ISTIC has a degree-granting graduate education program for information professionals. Ann Prentice, one of the members of our delegation, has taught in this program at ISTIC several times. While NSLC focuses heavily on service to users, ISTIC focuses on the provision of information services and collections. Besides publishing 19 science, technology and information science journals, they develop technologies and platforms for information such as Chinese vocabulary systems in subject areas, open-access journal platforms and information mining. ISTIC has purchased the archives of Nature and of the Institute of Physics for the entire country. These and other resources allowed them to provide document delivery for one billion items last year.

China’s model of graduate programs for information professionals attached to scientific institutes that are run by the government, as occurs at NSLC and ISTIC, is not found in the United States. ISTIC began to offer the M.A. degree in information science in 1984. They have approximately 35 students each year. NSLC began its graduate education programs in 1979 and began offering the M.L.S. and M.I.S. degrees in 1986. The Ph.D. program in library science began in 1993, and in 1996 NSLC began to offer the Ph.D. degree in information science in conjunction with Nanjing University. NSLC has about 150 students and 50 faculty members. The number of faculty is quite large as all faculty members are practicing librarians, supervisors and researchers as well as instructors. Both ISTIC and NSLC are government agencies with ISTIC being directly under the Ministry of Science and Technology. The Chinese Academy of Sciences, for which NSLC is the library, is the government institution for the management of China’s scientific research and was founded on November 1, 1949, one month after the founding of the People’s Republic of China. This involvement at the government level in information professional education extends to the mandating of a core curriculum at the undergraduate level.

Following our visits to the two scientific institutes, we observed information science programs at two universities: Beijing Normal University and East China Normal University (Shanghai). The administrative location of information science programs in Chinese universities varies widely. In Beijing the College of Information Science and Technology is part of the School of Management; at Shanghai the Department of Information is in the Business School. Although only a few institutions in China, including BNU and ECNU, offer the masters degree in library science, there are over 800 institutions in China that offer some sort of information-related degree. Most education in library and information science is at the undergraduate level. The Chinese Ministry of Education has established eight mandatory courses for library and information science curricula at the undergraduate level which include, among others, information technology, information resource management and information theories. Fifty percent of students in the graduate program have undergraduate degrees in library and information science; the other 50 percent have undergraduate degrees in the sciences or social sciences.

At each site we talked about possibilities for collaboration between the institutions and the People to People delegation. Possibilities that were discussed included participating in an ASIS&T panel in
October 2010, speaking at an SLA meeting in June 2011, participating in the Asian chapter of SLA or organizing an Asian chapter of ASIS&T, and developing faculty, student or librarian exchanges. Our Chinese colleagues were very interested in these suggestions.

We also visited the Bund Branch of the Shanghai Municipal Archives. While there we toured the “City Memory” exhibit which chronicles the development of the city of Shanghai from its beginning as a small fishing village in 1292 to the present time. We visited the public reading room and their electronic archives reading hall where about 3000 digitized films are available, along with other digital materials. This gave us the opportunity to find out about training for archives work in China. We learned that for the most part archives education is considered to be a different major than library and information science. Whereas public libraries in China are governed by a cultural ministry, the ministry of archives governs archives.

Of course our trip was not all work, and we did spend some time sightseeing. In Beijing we spent one day visiting Tian’an Men Square and the Forbidden City as well as the Badaling section of the Great Wall. In Shanghai we toured the Bund, the Shanghai Museum, Yu Garden and the Shanghai Urban Planning Exhibition Hall and attended a performance of the Shanghai Acrobatic Show. We were able to drive by the sites of the 2008 Olympics in Beijing and the 2010 World Expo in Shanghai. We also enjoyed many excellent meals, including one at the Quanjude Beijing Roast Duck Restaurant and a memorable meal at a private home in the Caoyang residential area.

Overall, the delegation found the trip very productive. We learned more about the education of information professionals in China, shared our goals for the promotion of information professionals and had positive responses to the idea of collaborating further. The state of information professional education in China seems to be on something of a parallel track with that of the United States, with a number of different kinds of programs aimed at different aspects of the information professions. Differences include a heavier government involvement and, perhaps, a greater concentration on technological aspects in China than in the United States. In addition to our interaction with the Chinese, we derived two additional bonuses from this trip. Because of the makeup of the delegation, we were able to discuss information professional education not only in China and the United States, but also in Japan and Canada. Through the participation of Anne Caputo, we learned more about SLA activity in this area and agreed on further collaboration between our respective organizations. At the completion of the trip, all participants agreed to continue talking. A first step in those ongoing conversations will be a workshop planned for the ASIS&T Annual Meeting in Pittsburgh; speakers there will include both People to People delegates and Professor Guoqiu Li from the Department of Information at ECNU.
What Is Information Architecture?
Practical Definitions and Useful Principles for our Second Decade of Study and Work

by Thom Haller, Bulletin associate editor for IA and special section guest editor

One of my former students sent me this cartoon. He called it “a little information architecture humor.”

The drawing points to the real-life challenges faced by those of us trying to structure information so others can find it, use it and appreciate the experience. We know organizations make many choices about what information to present and how to present it. And we know, as the drawing illustrates, that they often fail. Probably along the way, people and politics have influenced choices that differ from what site users want. Along the way, possibilities are lost because no one thought fully about information structure and use. Typically, in a site like the one depicted above, no one has thought through the information architecture.

Early this summer I attended a professional seminar given by someone I respect. He is well versed in applying taxonomic thinking in a business environment. His organization has garnered a significant following. During his presentation, he offered the worst definition of information architecture I had heard in years. (I won’t provide it here. But I can note that he showed us the slide in his PowerPoint slide deck and said, “This is my definition, but I’m not here to talk about definitions.” He then continued with his presentation. I wondered, “So why did you include this in your deck?”)

Last week, I attended a local “meet-up” where agency professionals presented their ideas on developing content. When the primary presenter referred to information architecture, he scoffed and said, “That’s so 1996.” I was taken aback.

“Why would he say 1996?” I wondered. I felt surprised to hear the professional label squished into a small time frame. I was not surprised to hear someone use the term with casual disdain; I often hear the term minimized. (For example, I frequently hear the term used as a noun: “We just need an information architecture, and then we’ll build this site.” Or, as a 2010 IA Summit attendee noted, “To my colleagues, IA just means working with metadata.”)

As a longtime IA/UX instructor, I typically offer big-picture perspectives such as those offered by Richard Saul Wurman when he introduced the term, saying, “Information architects make the complex clear.” I also direct students to different definitions, including this definition provided by the IA Institute:

- The structural design of shared information environment
- The art and science of organizing and labeling websites, intranets, online communities and software to support usability and findability
- An emerging community of practice focused on bringing principles of design and architecture to the digital landscape

Thom Haller, the Bulletin’s associate editor for information architecture, is a speaker, writer, user advocate and teacher of principles of performance-based information architecture and usability. Since 1998, Thom has taught classes on architecting usable web/Intranet sites. As a teacher, Thom enables students to structure information so people can find it, use it and appreciate the experience. He can be reached at thom<at>thomhaller.com; thomhaller<at>twitter.com.
But as the institute will tell you, this definition is broad. And then humans come along and look at broad definitions and face challenges with what all the words mean. My students look at that definition at the beginning of class and wonder, “What does structural design mean? What is a “shared information environment”? By the end of the class, we have constructed our own definitions and elevator speeches (but it’s taken us several weeks).

That’s why I like the definition of information architecture that Dan Klyn presents in this issue. I like its real world tone. I like how it tells the story and helps us understand when information architecture is needed and used. I also like Dan Brown’s straightforward definition (provided as an assumption to the principles he provides in this issue): “information architecture is the practice of designing structures.”

Why define? That question is asked often. I recall attending graduate school parties where my colleagues tried to define communication. Their efforts weren’t particularly successful. Communication scholarship languished as communication study became associated with football scholarships and minimal academic rigor. Ever think of rhetoric as “meaningless bombast”? There’s a story of a field of study now diminished in the eyes of the academy and the business community!

As people who believe information structure matters, we have a responsibility to define our work and make the invisible visible. This issue of the Bulletin takes some of the voices from the recent ASIS&T Information Architecture Summit. We believe this content can solidify the understanding of what information architecture is (and is not). We hope these ideas can encourage discussion, inspire research and propel action.

A Peek into Summit Conversations

The promotional literature for the 11th annual IA Summit suggested “the best conversations happen in the hall.” That may be true, but many worthwhile ideas emerged from the conference sessions. You will find many of these conversations online at the Boxes and Arrows site (www.boxesandarrows.com).

Among the presentations online are observations, challenges and perspectives offered in a conversation led by user experience professional Nick Finck. Nick synthesized this feedback on a blog post (an open discussion among information architects, http://ow.ly/2jhJ4) The presentation synopsis asked, “Are we all a part of ‘the looming fate of a fragmented industry struggling to stay alive’ … or might we instead see a unified, respectful community (with a clear vision, a clear goal, a true value and a solid message)? I found value in many of the observations, especially words from Richard Saul Wurman, who reaffirmed his belief in understanding.

I don’t think in terms of wireframes…I don’t think in terms of information architecture, I think in terms of understanding. I’m interested in understanding things that interest me and I am interested in the systemic rigor of being able to explain. I’m interested in communicating with another human being …and I am doing it with rigor and responsibility, and that is what information architecture is. Ultimately, it’s not the modality, not the particular technology… it’s just making yourself understandable. [Synthesis from Wurman’s observation during session. www.nickfinck.com/blog/entry/an_open_discussion_among_information_architects/]

In this Issue

In this issue we highlight a few of the ideas presented at the 2010 Summit and ideas recently discussed online.


Beyond a need for definition lies a need for communication. In their article for this issue, “Research and Practice in Information Architecture: Current Relationships,” Keith Instone and Andrea Resmini place this need at the heart of our work: “We believe a large village of practitioners is needed to raise a discipline, and at the heart of the village is the need to communicate.” In this article they call for a community-wide research agenda and a need for practice-led research.
In his article “Information Architecture, Black Holes and Discipline: On Developing a Framework for a Practice of Information Architecture,” Nathaniel Davis wonders about our challenges to “strategically tackle information with greater precision in more complex situations.” He describes the challenge of increasing information and offers a framework from which we can group and assess IA practice.

Dan Brown closes the conversation with structure to move the conversation forward. He offers eight fundamentals to support information architects and other user experience designers as they meet real-world demands. He concludes by writing:

A serious theoretical framework establishes a place for interpretation, reconsideration and debate. A theory of information architecture would escalate our conversations, taking us beyond “what do we do” and even “how do we do it” and into something far more interesting: “How do we do it better?”

What Next?

There’s our question: How do we do this better? How can we focus our research with our practice? How can we move from casual observation toward practice-led research? How can we become confident in our understanding of information architecture so we can attend to our job of creating understanding? Explore this issue. Read different perceptions of definition, identity and roles. Afterward, forward me your thoughts, definitions and priorities for an IA research agenda.

This conversation continues…
A friend of mine who works at a storied and successful advertising and PR firm that’s undertaking a reconfiguration to become an “integrated agency” asked me about job descriptions for information architects. He was looking for a way to advocate for a position called “information architect” in this firm and was asking for a job description that would make sense to the VPs and C-levels. I didn’t come up with anything good after Googling for two minutes so I decided to craft my own:

The role of the information architect is to help discover and articulate the “why” of the project and to work directly with the client and horizontally across vertically oriented disciplines within the agency to ensure consistency and continuity of meaning in the processes, products and metrics the team uses to measure its outcomes.

In the early stages of the project, the information architect assists in the formulation of strategy and uses simple pictures and sometimes complex linguistic structures to assist the team and the client in explaining and understanding the nature of the client’s business and the nature of the experiences that consumers will have with the client’s brand, products and collateral.

In the middle stages of the project, the information architect assists the project team in identifying the “what” and then subsequently the “how” of the project and in prioritizing, selecting and arraying specific tactics in ways that protect the continuity, consistency and purposefulness of the resulting user experiences.

During the implementation stage of the project, the information architect listens to and collaborates with the people who build, test and deploy the solution in order to capitalize on the unique insights that emerge during implementation, maximizing opportunities to enhance the solution as it is being built.

Once a solution is deployed, the information architect assists in the measurement of performance and of users’ experience and in the collection of diverse insights to continuously improve the performance and maintainability of the solution.

Upon re-reading the fictional job description I’d just written, I quickly realized that I’d done it wrong. I’d forgotten the cardinal rule of user-centered design. Instead of writing copy that would make sense to and persuade the executives at the agency, I wrote the copy as if I were the one who needed to be convinced.

There Are No Such Things as Information Architects Anymore

Before starting in on a user-centered rewrite of this job description, this time with the explicit purpose of persuading advertising executives that they should create a position called IA, I paused to reflect on what these folks might find on Google if they did a quick search for information architect. Would they find the video from Jesse James Garrett’s closing plenary at the ASIS&T Information Architecture Summit in 2008 where he declared...
“there is no such thing as information architects - there are only and have only ever been user experience (UX) designers?” Or would any of them find the episode of Jeffery Zeldman’s Big Web Show where he introduced that day’s guest as “what we used to call IA?” Probably not. It’s even less likely that their Google search would connect them with footage from the 2009 ASIS&T Information Architecture Summit, where Richard Saul Wurman reiterated his original conception of IA as an “understanding business” which is concerned with “making the complex clear.”

No, I think the cacophony and near-impenetrability of the first page of results for searches on information architecture and advertising would – thankfully – combine to protect these executives from the harder-hitting rhetoric in the war of words that would seem to pit IA against user experience. In my rewrite of this job description, I decided that there were just two external reference points for “IA” that would matter to the audience I was writing for. I needed to focus on what other folks in their social network talk and think about IA and UX, and what other agencies say and think about IA and UX. And in both of these categories, it’s my belief that IA has, in the words of the Mad Hatter in Tim Burton’s remake of Alice in Wonderland, “lost its muchness.”

“You Used To Be Much Muchier”

There’s a combination of anecdotal and Google Trends evidence to support my assertion that the folks who run this advertising firm would perceive IA (if they perceived it at all) as something less than user experience. From a simple mathematical standpoint, the number of references to user experience that advertising executives encounter over a given week or month will surely dwarf the number of references to and encounters with information architecture.

How do these executives’ peers and competitors think these terms relate to roles and job titles? The spread between listings for “user experience” compared to listings for “information architect” (both in aggregate as well as when narrowed by the advertising industry) follows a similarly top-heavy pattern.

“The Lunatic of One Idea in a World of Ideas”

If you’ve not already detected it, I will here state plainly that I’m biased when it comes to matters having to do with the role and discipline of IA. You could accurately characterize me, in the way Wallace Stevens described a tragic character in one of his poems, as “the lunatic of one idea/in a world of ideas.” I teach the IA course at the University of Michigan School of Information – a job that I love and that I inherited from my mentor Peter Morville, who’s the co-author of the bestselling “Polar Bear” O’Reilly book on IA. I’m also a devoted disciple of Richard Saul Wurman, the lapsed architect and former protégé of Lou Kahn who invented information architecture as chairman of the American Institute of Architects annual conference in Philadelphia in 1976. Morville and Wurman have widely diverging ideas about IA, but they’ve both articulated an expansive vision for the role and the discipline of IA.

At the other end of the spectrum, beginning with a diagram in 2000 called “The Elements of User Experience” and then reinforced in 2003 an influential book of the same title, Jesse James Garrett cast a far narrower vision for the role of IA, allowing it only half of a slice in a six-slice array that visually catalogs the constituent members of an overarching concept called user experience.

I place myself at the far end of the Wurman/Morville side of the continuum, and while I acknowledge that the task of writing a job description for “user experience designer” would be a much easier sell in the integrated digital
agency world, I believe that this word integration is the key to persuading my friend’s bosses that “information architect” is what they want for an integrated agency. Not to the exclusion of “user experience” or “user experience designer,” but as the starting point for delivering on the idea of integration.

Integration, Like Information Architecture, Is Horizontal

So as you may have already predicted, I never did the re-write. My original copy is what I ended up sending, and the deciding factor for me was the particular context of integration. There’s a fair amount of cynicism at this agency and probably many other “traditional” agencies that are transitioning to the “integrated” model – skepticism that true integration is possible in a context where work has always taken place in a silo world of departments and handoffs. This unease is why an expansive, Wurman/Morville sort of IA seemed to me to have the best chance at connecting with the executives who are shaping their agency for an integrated world where print and broadcast and PR and web are united by a common choreography that creates and sustains success for the client and profits for the agency.

Wurman’s idea of an information architect is particularly well-suited to the integrated advertising agency because, as he envisions it, the IA works independent of specific tools and modalities, and the IA’s work obtains in all communications media. It’s also always (for Wurman, at least) been posited as the role for a single person. Not that you wouldn’t or couldn’t have Wurmanite IAs working in teams, in a business climate where significantly more jobs are being shed than filled and created, arguing for the creation of a role for a single IA with an expansive horizontal authority is perhaps a more plausible proposition than hiring in and creating roles for the “big umbrella” UX professionals whose expertise has nevertheless been winnowed into Garrettesian buckets of IxD, IA, visual designer, information designer, content specialist and so forth.

There may currently be no such thing as information architects in these “integrated” advertising agencies. But I believe that if you posit the idea of an information architect within an integrated agency context with the requisite degree of muchness and horizontalism, the VPs and C-levels will quickly realize the value, and you can call it whatever you need to.
It takes a village of practitioners to raise a discipline.
— Nathaniel Davis. “Practice of Information Architecture” [1]

What is the current relationship of research and practice in information architecture? That was the question posed in a “back of a napkin” collective-thinking exercise introduced by Keith Instone and Andrea Resmini in “Bridging IA Research and Practice,” a session they co-hosted at the 11th ASIS&T IA Summit in Phoenix, Arizona [2]. For additional ideas from participants, please see the list and description of other “napkin sketches” at the end of the article.

During the session, Melissa Weaver crafted the following sketch.

Weaver’s drawing represents the “now” in IA: the academy teaches students, and they, in turn, become practitioners who apply what they have learned. It is a one-way relationship.

The session was meant as an opportunity to re-start a dialogue between practitioners and researchers in the field of IA, with the primary goal being to offer some common ground to discuss the different approaches to core IA issues, facilitate understanding of the different views and interests at play, highlight both the strengths and weaknesses of the instances academia and business present, and identify the added value this diversity brings to the conversation.

Secondarily, the session served as a test bed to verify the feasibility to extend and expand the conversation to the 12th IA Summit in 2011. The following are among the questions raised:

- Should a pre-conference consortium be part of a tentative answer to bridge the gap?
- Should IA-related publications, such as the Journal of Information Architecture, be somewhat more involved in this specific conversation? How?

“Beyond expectations” is how the organizers summarized the workshop. More than 30 participants from both camps (research and practice) exchanged ideas. To uncover these ideas, the group facilitators used rapid sketching – attendees drew sketches on real paper napkins, which were collected, explained and discussed. A number of very interesting points were raised in these brief, open exchanges. The following major themes, statements and questions [3] emerged:

1. Research and practice in IA are fractured, with very weak connections in between (and this gap is not unique: many disciplines face this challenge).

2. There is a need for an IA research agenda (or if there is one, it needs to be common knowledge). Or, the opposite: there is no such thing as IA research by itself: there is HCI, IS or LIS research that IAs care about.

3. A huge percentage of time, people and resources in the field are devoted to IA practice only. Few resources are committed to scientific IA research.
4. Many people confuse “project research” that is done in a specific context (and is extremely useful for that project, but hard to generalize to anything else) with “scientific research” which proposes a hypothesis and tests it with data.

5. Resources like Google scholar provide an easy means to find existing research, but making sense of the research and applying it is very difficult. There is clearly an opportunity for a “Carl Sagan like” figure to bridge the gap from research “awareness” to “consumability” for practitioners (with its related question whether there is a business model waiting to happen here).

6. IA research needs to be part of an aggregation that looks across disciplines and present boundaries to really be effective for practitioners.

7. Transferring knowledge is possible. BJ Fogg of Stanford Persuasive Technology Lab and of “Persuasive Technology” book fame [4] was introduced as an example of a researcher who spent a lot of time explaining his research to a company.

Detecting Patterns

Most of these points might seem to stem from plain common sense, and certainly some of them have been long discussed in the community at large. Nonetheless, the picture they paint is definitely interesting. A few common patterns can be detected:

- Current state of research and practice
- Challenges and opportunities
- Roles and responsibilities
- Necessary next steps.

All seven sections point to one single undercurrent that touches on all points: communication. Research and practice have to bridge their distance and move from Melissa Weaver’s “now” to the another image she illustrated during the workshop: her “ideal” environment – one where communication is bi-directional.
no common language, it may be easy to decide there is no specific IA research – the second major theme in our list above.

We certainly have good publications that try to spread seminal research ideas in IA as well as enlightened or useful best practices from the field. The *Bulletin of the American Society for Information Science and Technology* [8] is one of them. The *Journal of Information Architecture* [9] is another, more strictly connected to standard academic communication. We also have plenty of good international professional online magazines such as *Boxes & Arrows* [10] or *Johnny Holland* [11] and plenty of local ones such as the Swiss *UXMagazine* [12] and the Italian *Trovabile* [13]. You probably know a few more in German, Dutch or Japanese.

Challenges and opportunities. Again, it seems that the missing piece is communication, awareness of being part of a global conversation incorporating both research and practice. Speaking generally, practitioners can only be interested in research insofar as it has a direct connection to whatever project they have at hand. They are also typically not interested in furthering research upon reaching project completion. Academia has been trying to find ways to overcome this dichotomy. Action research, for example, tries to overcome this short-circuit by having researchers *de facto* embedded inside the problem solving and production process.

Professionally oriented research methodologies could help this conversation along. Maybe it is time to reconsider the relationship between IA (among many others) and user centered design, assign this latter its rightful place as one of the useful methodologies and tools in our toolbox, and think about ways action research, co-design or practice-led research could help us move the conversation further.

Very recently, for example, in “Maturing a practice” [14], Hobbs, Fenn and Resmini suggest that having the village of practitioners producing research might be difficult, but not impossible. Practice-led research (PLR) could provide one such way to kick-start a virtuous process. Practice-led research is mostly used in fields where the artistic, personal or design-related component of the field or discipline is largely dominant. Much of the research in the field is conducted by people who are already active in it and who, for the purpose, wear two different hats, researching into their own practice.

PLR would, for instance, allow a single research discussion to develop where otherwise dual conversations on research and on practice would exist. Hobbs, Fenn and Resmini conclude:

> It is our suggestion that a broad uptake of PLR by practicing user experience designers and information architects could assist in the generation of knowledge and discourse and that this in turn could considerably assist in the maturing of the practice towards a discipline.

Should information architects learn more about practice-led research? This is one possible path, one we will be exploring during the next academic year. But is it the only path? Certainly, creating a robust discipline of information architecture (and user experience design) would require more than the active uptake of one single research method.

Questions relating research and practice are not new. Researchers have definitively been thinking and researching into the IA problem space [15] (major themes 2, 3 and 4), but as Resmini, Byström and Madsen argue in their article “IA Growing Roots” [16], connections between the different efforts are still thin and stretched, to the point of losing a great deal of momentum just because, as the attendees at the IA Summit remarked, there is no collective agenda and practitioners and academia stay on separated, non-communicating sides. We are back at communication again.

Roles and responsibilities. Several of the points identified in our conversations at the IA Summit touch on one more question that is central to the development of a common conversation between research and practice in IA: What are the respective roles and responsibilities? What do practitioners have to do to advance IA? What do researchers have to do? Many through the years, and most recently Jeff Parks [17,18], have outlined some pretty interesting ideas on the discipline and ways the community at large should engage in conversation around frameworks and methodologies. Peter Morville has worked to push the boundaries of the framing of the field itself. Jesse James Garrett has (famously) presented views on our role.

Other practitioner/writers such as Andrew Hinton [19, 20] and Christian Crumlish [21] have addressed the roles and responsibilities of the information
architect and have provided thoughtful articulations of the issues relating to their identities as IAs and the field at large. This kind of reflection on the role is scarce in IA research. A rather large number of information architecture research papers [15] work around the discipline, its status, its boundaries, its scope and goals, but they fail to address the information architect as an individual who is involved in IA research.

Possibly one of the most interesting ones has been put forth by Dan Klyn, who in a post on his blog *Wildly Appropriate*, titled “Job description: information architect” [22] offered his personal view of what the role of the information architect is in the field of practice. A version of this post, including his five-point definition, appears as the first article in this special section.

**Looking Forward**

To look forward, we return to the results of the information architecture session. If we exclude the very obvious – a willingness or necessity to keep the conversation going – we reach two conclusions: (1) Communication of the respective goals, agendas and advancements between research and practice is lacking, and (2) because of the way the field and the larger IA community has evolved, IA resources in the community are heavily skewed towards doing.

We certainly acknowledge that research has produced quite a number of papers: positioning papers, where opinions are voiced or general trends identified, and scientific papers on specific IA subjects. We believe it’s time to turn research questions into practical ones and rigorously test some of the assumptions of the field. Second, practice has produced 15 more years of artifacts. It’s time these are discussed, criticized and explained in a way that benefits the field at large and that produces factual and theoretical knowledge to be reused.

So to look forward we see the need for communication and for available resources. These requirements are really two faces of the same coin. We believe a large village of practitioners is needed to raise a discipline, and at the heart of the village is the need to communicate.

To communicate, we need channels. One central issue that should not be dismissed is the fact that most IA-specific conferences and events are targeted at professionals, both in the way they are run and organized and in the way they stand in the face of academic relevance. This issue directly connects to a few of the observations in the list, namely to those that point out that there is no IA research proper and that there is a need for building long-standing bridges that stretch into the future. While we do think that there is definitely IA-specific research, unique perspectives can become lost when communicated in non-IA-specific channels.

To communicate we need a common language. Perhaps we need some middle-ground individual or organization, such as the Information Architecture Institute (http://ia institute.org/) to act as a bridge or as a facilitator. As we disseminate information and cross-pollinate with other ideas, we create business opportunities for publishers such as O’Reilly or Rosenfeld Media.

Can we offer a specific strategic response? Perhaps not. But we can offer a few thoughts on the key actionable points that can push the conversation forward:

1. Build long-term relationships between researchers and practitioners. Student engagement is one such long-term strategy since students are currently in academia (close to where research happens). But students will someday be the leaders of the community of practice. Sponsor students to participate in relevant IA conferences.

2. Build up common channels and meeting points such as the Research and Practice session at the 11th ASIS&T IA Summit. Academic interests. Reintroduce academic space in major IA conferences to make them again relevant to researchers. Include representatives from the *Journal of Information Architecture* and other publications.

3. Actively promote published communication channels (*Bulletin of the American Society for Information Science and Technology*, *Boxes & Arrows*, *Journal of Information Architecture* and other professional magazines) to both sides in a way that helps mature them and their audiences.

4. Disseminate IA-specific conversations, clearly identifiable as such, in related communities, conferences and meetings, such as the human-
computer interaction conferences and the ASIS&T Summits. IA needs to be visible and legitimate in more than one way.

5. And finally work toward a pre-conference meet-up consortium at the 12th ASIS&T IA Summit to move this discussion forward.

IA-specific research could assist in creating scientifically driven, scientifically based knowledge. This research can support practitioners and help move IA from a conversation within a community of practice to a discipline. Bridging researchers and practitioners to work for the common good seems not only inevitable but also essential to foster growth and help the field mature into a discipline. Contributing to one single conversation will benefit both communities. And it’s the only way to move the discipline forward.

Additional Napkin Sketches from “Bridging IA Research and Practice”

Here are some additional perspectives gathered on the back of the napkin at the session, “Bridging IA Research and Practice” a session co-hosted by the authors at the 11th ASIS&T IA Summit in Phoenix, Arizona.

**Sketch A:** [www.flickr.com/photos/resmini/4511076085/in/set-72157623664715269/](www.flickr.com/photos/resmini/4511076085/in/set-72157623664715269/)

What’s the problem? Where is the research agenda? What are the high-level problems that need to be solved? What are the big questions that research needs to help us answer? A research agenda sets this direction. [Ref: No. 2 in the list of themes, statements and questions.]

**Sketch B:** [www.flickr.com/photos/resmini/4511077501/in/set-72157623664715269/](www.flickr.com/photos/resmini/4511077501/in/set-72157623664715269/)

One huge circle that has many cracks. That is the “practice” circle. One small circle off to the left: that is the “research” circle. The practice circle dominates and the research circle barely exists [point #3]. Very far apart and no connections. [Ref: No. 1 in the list of themes, statements and questions.] Also, the practices themselves are fractured.

**Sketch C:** [www.flickr.com/photos/resmini/4511078061/in/set-72157623664715269/](www.flickr.com/photos/resmini/4511078061/in/set-72157623664715269/)

(Front of the napkin): This is “Now.” Academy (research, teach) leads to Students (learn) leads Practice (apply). Arrows go in 1 direction only.

**Sketch D:** [www.flickr.com/photos/resmini/4511078527/in/set-72157623664715269/](www.flickr.com/photos/resmini/4511078527/in/set-72157623664715269/)

(Back of napkin): “Ideally.” The 3 Academy, Student, Practice for a triangle with exchanges across each side [Ref: No. 8 in the list of themes, statements and questions.].

**Sketch E:** [www.flickr.com/photos/resmini/4511719030/in/set-72157623664715269/](www.flickr.com/photos/resmini/4511719030/in/set-72157623664715269/)

“Research. Practice. Cannot do both. Cannot be in both places.”

**Sketch F:** [www.flickr.com/photos/resmini/4511719466/in/set-72157623664715269/](www.flickr.com/photos/resmini/4511719466/in/set-72157623664715269/)

“Research in...Cognitive psychology...Library science...HCI...Computer science...Learning sciences...Social psychology...etc.” They all lead to “Practice in Information architecture” and there appears to be a lot of money to be made. [Ref: No. 6 in the list of themes, statements and questions.]

**Sketch G:** [www.flickr.com/photos/resmini/4511720490/in/set-72157623664715269/](www.flickr.com/photos/resmini/4511720490/in/set-72157623664715269/)

“Theory” at the top of a tower, “Practice” at the bottom. An ivory tower? How does one make it up the stairs? Is that a technology component on the right?
Resources Mentioned in the Article


[5] This situation has been widely discussed, for example by P. Morville in his books and on his blog [6] and by A. Resmini and L. Rosati [7] in their approach to information architectures as ubiquitous ecologies.


[12] UXMagazine: www.uxmagazine.it/


[21] Christian Crumlish centered his salute to the 4th Italian IA Summit around this very issue. A barely discernible video of some 5 minutes of his talk is available at www.youtube.com/watch?v=Kkh1Hg0wMCG.

Have you ever wondered how long it takes for a field of practitioners to produce the understanding that enables the flowing rivers of discipline – long enough for academic enlightenment? How does an industry actually achieve the longevity and relevancy seen in science and law; in dentistry and medicine; in masonry or carpentry? Even music and the theatrical and fine arts possess histories of understanding that have been taught and passed on throughout generations.

All of these examples are well known institutions of discipline that are centuries in the making – if not longer. But, for the field of information architecture for the web – one that was given a name only 10 years ago and has only been probing for less than 20 years – the expectation to render a discipline is upon it.

In a world of information technology – where one year is like 100 – academic institutions and the business community want instantaneous answers for thinking about the discipline. Academic institutions that seek to train the ways of information architecture request the substance of our work. The business community continues to demand justification and quantitative proof of the value of information architecture in order to make practical cases for it at the tables of business strategy.

Just as product-based technology markets are maturing faster, the field of information architecture is burdened with the same expectation. And the pressure – to communicate to the world that the field of information architecture is a necessary business commodity and has a perspective that can be understood and taught – is growing.

To supply the demand placed upon it, the field of information architecture must deny the more natural inertial forces of organic evolution over time; because the time we think we have is not the same as was enjoyed by earlier disciplines and other pioneering industries. And worse, time may be running out.

To meet the new kind of demand of the information age, the IA field at large must proceed with the clear intent to cultivate discipline rather than to happen upon it.

By adopting an actionable plan to obtain discipline in the practice of information architecture, the field will flourish. If not, while a functional interest within organizations might be all that remains – augmented through other established disciplines and abstract mathematical algorithms – the distinct professional field of information architecture may soon cease to exist.

Things Are Rarely What They Seem

Growing up as a child in the 70s and 80s, the propaganda about the world around me was that everything was pretty much figured out. Previous human enlightenment had raced away to produce industry upon industry teeming with established disciplines – offering springs of knowledge that would satisfy the occupational interests of future generations. Our forefathers had created and lived through the evolutions of social and technological enlightenment, the industrial age and the inventions of mechanization – we stood on the shoulders of giants. Well, that was the high-level view of reality, at least.
It just so happened that underneath the thick superficial crust of human accomplishment and occupational constancy was a dynamic sub-level of human ingenuity within the scientific community that was making new discoveries. Its most socially disruptive invention, the invention of the computer processing unit (CPU), could be compared to an “event horizon” in astrophysics – the CPU being a black hole-device that (for good or bad) increases the velocity and distorts time for anything that gets in its path.

Mankind’s relationship with the CPU – through the use of software applications – is pushing ever closer to exponential increases in performance as the people that rely on it are continually finding new requirements for greater efficiency. Whole societies are getting sucked in. Now, millions of corporate networks are outnumbered by tens of millions of social networks. And both are pulsating with information. The “space” that this information occupies within the realm of information technology accommodates new networks that seem to spring out of nothingness in cosmic fashion. It appears that space is expanding. But technologists seem to find ways to traverse this space with greater efficiencies.

When this happens, I am skeptical of what seems to be the obvious, but I must question how it is that I can go from point A to point B in my application or any area in the domain of information technology faster than before? Am I, or rather, the things I control, moving faster? Or is the space of information technology shrinking? Has a mile become an inch? What dynamic change are we really experiencing?

Change and Time

What doesn’t seem to change as quickly in the evolving domain of information technology are the people that enable its existence. The applications and components we use seem to flow with great efficiency, but our individual ability to process and take action against newer computed interactions has not changed. As the information around people becomes more complex, it is becoming more complex for the people to use it. Instead of using more of the information we are creating, we are using less of the information we create. And instead of understanding more information we come across, we understand less of it. Consumption is outweighing utility.

On a grand scale, as people continue to use information at exponentially increasing rates, information is slowly stagnating. If nothing is done to efficiently release information into some broader dynamic flow of efficiency, the information that now defines much of our IT-based experience will become unintelligible, inaccessible and completely unresponsive.

In the domain of information technology, time is dependent and defined by events happening. If there are no events, time cannot be measured. When time is measured, it’s done so in context to the objects that make up the measured experience – such as a targeted objective and the time it takes for a person to execute the objective.

In our work to prepare and present information, co-dependence has emerged. This co-dependence is among four factors: the CPU, the applications...
that require them, the information that feeds the applications and the people
that depend on all the others working without failure. Let’s refer to this
relationship as the QC – short for Quartet Compression.

Because of the QC, time is not something perceived directly; it’s the
intangible gap between interactions through which people individually
perceive. Time exists because of these spatially defined interactions – and
vice versa. Hence, space and time in the domain of IT are co-dependent – so
much so that they are one in the same.

More specifically, under the Quartet Compression of Technology, a
domain driven by information technology reveals a co-dependence that
defines a unique contextual space-time relationship as shown in the equation
below, where, \( spIT \) represents the space-time of information technology:

\[
spIT = (CPU \leftrightarrow Application \leftrightarrow Information \leftrightarrow People)
\]

Because everyone holds their own digital position in the space of IT, they
also hold their own unique position in time, and this is where things get tricky!

As Einstein observed, “change is constant.” Unfortunately, this constant
flux in the space-time of information technology creates a concern, because
things do not appear to be gravitating towards a sustainable equilibrium of
beneficial activity for all; and it appears that space-
time is literally running out for some.

As the end looms over an unsuspecting civilization of
information pacifists, data
junkies, information hoarders
and early adopters, people
can be observed having advancing or receding experiences of time. Some
can be seen interacting with ease and rapidity while others demonstrate
frustration over menial tasks that move them to completely disengage from
active conversation.

Hope

As the years have passed, individuals from various backgrounds – who
are forming communities with labels such as information architecture,
interaction design, content strategy, user experience design and usability
engineering – have managed to understand enough of the complexity of the
co-dependence (QC) to slow the pace at which civilization is approaching
the point of no return. But their efforts are drowned out by the swooshing
winds of occupational confusion.

Fortunately, civilization is still far enough from the event horizon of an
information black hole that it is generally unaware of this dire circumstance.
Nor do we typically know that without the intentional probing into the
architecture of the QC, our descent into the abyss may have happened a
decade or more sooner. This previous success leaves room for hope as a
loose network of communication and community has encouraged the idea
of interrupting the continuum of the QC by targeting the information
component and controlling it through a “managed architecture.”

The main approach – inspired, in part, by the discipline of library science
and commonly referred to as information architecture – has been partially
successful and has displayed areas of reliability. However, the broad
communities of interests that surround information architecture inefficiently
communicate. We see no clear or widely understood paths explaining how to
strategically tackle information with greater precision and in more complex
situations. This challenge alone could be the undoing of the fragile
emergence of the IA profession.

A Call to Action: Engage in a Discipline of Practice

As time paradoxically and rapidly progresses to the point where it may
slow everything to a screeching halt for some, those dependent on information
technology don’t have centuries to wait for a discipline that tames the threat
of information. Exercising a proactive perspective around the practice of information architecture in the domain of information technology and building consensus are critical for reversing the hand of time that is in the grips of the QC.

If broader consensus on the nature of information that partially constitutes the QC can be reached and effectively communicated, then more people can be systematically trained to architect it, and curators and business owners of CPU-driven technologies and interfaces can be educated on the inherent risks and responsibility they have to manage information in a sustainable manner.

To accomplish this effort, the future of a discrete discipline of information architecture rests on those practitioners who choose not to sit on the proverbial fence of occupational identity. UX designers, interaction designers, content strategists, user researcher and testers are all necessary and fine. But, the world needs information architects.

A call to be an information architect will not mean that one is an expert, but that one is in the pursuit of expertise in cracking the challenges presented by information as it is presented in context to a computing interface and other media. An information architect may as well mean researcher, because there is much to do and discover to truly prepare the field with the artifacts necessary for higher learning.

Hence, to reveal what will be the discipline of information architecture, we must first exercise a discipline of practice with the rigor of research.

At the ASIS&T 2010 IA Summit, I gave a presentation titled, “The Practice of Information Architecture: It Takes a Village of Practitioners to Raise a Discipline.” The presentation introduced a formal definition of practice and additionally argued that the practice of information architecture must include an understanding of how businesses are modeled for efficiency.

Since the presentation, I’ve expounded on the practice definition to create what I now refer to as the DSIA Practice Framework. The framework, which originated from theoretical and practical insights from my own brand of practice-led research (PLR), described on my website www.methodbrain.com, is meant to serve as a guide for practitioners of information architecture. The framework is rooted in the following definition of practice from my ASIS&T presentation:

Practice is the collective behavior of intentional empirical probing around an area of interest, whereby the contribution of documentation of discovery enables consensus that builds and reinforces discipline around such behaviors.

The italicized words in this definition highlight the key behaviors of practice that inform a basic framework by which the pursuit of one’s IA practice and discipline may be groomed and assessed:

<table>
<thead>
<tr>
<th>Primary Behaviors of Practice</th>
<th>Organizational Assessment</th>
<th>Individual Practitioner Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention</td>
<td>Is your IA organization able to articulate its functional objective within its parent organizational business model?</td>
<td>Do you understand the work product of your IA organization? Does your intention align with the IA organization?</td>
</tr>
<tr>
<td>Areas of Interest</td>
<td>Does your organization explore subject matter that fills gaps in its processes and offer value to your IA organizational model?</td>
<td>Are you pursuing core area/s of interest? Do your interests complement others in your immediate group? Do you understand your role?</td>
</tr>
<tr>
<td>Contribution</td>
<td>Is your organization fostering a forum for building knowledge that is in line with the strategy of the IA organization?</td>
<td>Have you set a goal to contribute your insights to further a collective knowledge and shared awareness within your IA organization?</td>
</tr>
<tr>
<td>Documentation</td>
<td>Is your organization archiving and providing supportive communications around its collective discoveries?</td>
<td>Are you allotting time to record your experiences – good or bad – in order the apply learning in the future?</td>
</tr>
<tr>
<td>Consensus</td>
<td>Is your organization refining its insights and identifying a path whereby knowledge can become accepted as “best practice”?</td>
<td>Are you growing in understanding? Are you testing/validation your assumptions and/or those of your peers?</td>
</tr>
<tr>
<td>Discipline</td>
<td>Does your group focus on understanding the system and efficiency of its functional responsibility, to the business organization, in order to repeat it.</td>
<td>Are you building upon and referring to formal systems (discipline, theory) to help guide what you do?</td>
</tr>
</tbody>
</table>
Closing Thoughts

In many cases it takes a lifetime to mature a discipline. In the domain of information technology, a lifetime is measured in market cycles that are driven by consumers with an insatiable appetite for information. Further, a disturbing amount of dormant information may be collecting in the corners of the IT universe – fueling an apocalyptic-size hole of inefficiency.

In the midst of this threat, information architects have an opportunity to begin an era of methodical practice that enables a discipline as well as the transference of acquired knowledge to address the real qualitative effects of information. The DSIA Practice Framework offers an approach that should be explored and vetted for its usefulness in the pursuit of maturing one’s information architecture practice.

When information architecture is widely practiced with rigor, the field may soon demonstrate the theories and practical methods to handily reduce the gravity of the growing black hole of information that quietly disrupts the experience of time and effort by users of computing interfaces.

By documenting our successes and failures through a practice framework, the information architecture discipline will become more tangible with a legacy worth noting in the history of information technology. I look forward to the story that will be told.
Eight Principles of Information Architecture

by Dan Brown

Information architecture, a field in relative infancy and constantly rediscovering itself, does not yet have a well-established theory that drives the design of structures for websites. You can pick up two books on graphic design and see the same topics covered in each. Why is there no such agreement for books on information architecture? Indeed, information architecture has yet to normalize, and the constant new demands on our work make progress toward that goal challenging.

I’d expect such a theoretical framework to have a set of principles – guidelines based in universal truths that provide a sketch of what makes any information architecture good. White space, typography and color theory provide a set of principles for graphic design. Where are the ones for information architecture?

While the industry has yet to settle on a standard theory, there are a handful of principles I keep coming back to – principles that guide my design decisions but don’t dictate a particular approach. They help me zero-in or refine a concept by providing a stable and reliable worldview. Fairly general in nature, they apply to most situations and have sufficient room for interpretation so I can fine-tune them for individual projects.

I first codified these principles when a client of mine faced internal pressure to justify the design we presented. She knew the direction was sound and rested on several iterations, validated requirements and existing user research. She worried, however, that her stakeholders wouldn’t be convinced. I put together a one-page summary of these principles to establish a theoretical foundation for her internal discussions.

These principles make an assumption: information architecture is the practice of designing structures. These principles help guide the design of structures, but they presume the following:

- The information architect’s primary focus is the structure itself and secondarily the user interface representing the structure on screen (I make site maps and flow charts)
- The information architect has a good understanding of how people want to relate to the content and functionality contained in the structure (I’ve done my research)
- The information architect has a good understanding of the range of content and functionality to be supported by the structure (I’ve inventoried the content)

Got those? Good. Let’s look at the principles. Here’s a preview:

1. The principle of objects – Treat content as a living, breathing thing, with a lifecycle, behaviors and attributes.
2. The principle of choices – Create pages that offer meaningful choices to users, keeping the range of choices available focused on a particular task.
3. The principle of disclosure – Show only enough information to help people understand what kinds of information they’ll find as they dig deeper.
4. The principle of exemplars – Describe the contents of categories by showing examples of the contents.
5. The principle of front doors – Assume at least half of the website’s visitors will come through some page other than the home page.
6. The principle of multiple classification – Offer users several different classification schemes to browse the site’s content.
7. The principle of focused navigation – Don’t mix apples and oranges in your navigation scheme.

8. The principle of growth – Assume the content you have today is a small fraction of the content you will have tomorrow.

The Principle of Objects

Treat content as a living, breathing thing with a lifecycle, behaviors and attributes.

Where it comes from. Thinking of content as objects comes from object-oriented programming, where a computer program is broken up into discrete, logical chunks. Each chunk has methods, behaviors that the chunk of code can perform, and properties, pieces of information attached to the object. Objects are really templates, so the methods and properties provide a framework for thinking about all instances of a particular object.

So, when I say “object,” I mean that website content has

- a consistent and recognizable internal structure (such as the different ingredients of a recipe) and
- a discrete set of behaviors (such as how recipes can have variations or can become more or less important).

How I use it. When starting a new project, I identify all the content types – common structures that will be used throughout the site. For marketing sites, I might identify content types such as products or services or industries. But that’s boring.

Let’s take a recipe site. I say “recipe” and you already have a model in your head about what kinds of information it will include: ingredients, quantities, process, servings, cuisine, dish type, preparation duration, dietary information and so forth. This structure gives me different ways to sort, expose, classify and connect content of this type together. We can create relationships to recipes that are complementary. Or we can connect this recipe to different versions of the same dish. Recipes, therefore, have a predictable structure that enables building relationships. Got that? There’s more.

You can even think of recipes as having behaviors. Recipes can “reproduce” when people comment on them with their own twist. A recipe’s importance can recede or intensify in different seasons (strong emphasis on matzoh ball soup in mid-Spring, for example). This kind of content has pretty predictable ways in which it behaves.

The Principle of Choices

Create pages that offer meaningful choices to users, keeping the range of choices available focused on a particular task.

Where it comes from. The Paradox of Choice is a book by Barry Schwartz [1] that came out in 2005. In brief, the book’s message is that a greater number of options can make it more difficult for people to make a decision. More options means more cognitive effort, and more effort can sometimes mean more anxiety. People think they like having a lot of options, but they really do not.

How I use it. Corporate intranets are ripe for long lists of content. The organization publishes policies, all on a particular topic (benefits, say) and rarely makes the editorial effort necessary to keep the lists lean and focused. The cost is obvious: Users spend more time sifting through lists to find what they need. They might therefore avoid looking through the lists altogether, never finding what they need, or they might use alternate channels, like picking up the phone, defeating the purpose of the intranet.

In designing information hierarchies, I tend to spread things out. That is, I tend to make shorter lists of choices, at least at the upper levels of the hierarchy.

The Principle of Disclosure

Show only enough information to help people understand what kinds of information they’ll find as they dig deeper.

Where it comes from. Progressive disclosure is a common design concept that builds on the ideas that we can only process so much information at once, but that we can use what we have to anticipate what’s to come. I like the explanation in Universal Principles of Design, a book by William Lidwell, Kritina Holden and Jill Butler [2, p. 154]: “Information presented to a person who is not interested or ready to process it is effectively noise.”
How I use it. The main way progressive disclosure influences thinking about a design is that I have to think about content in terms of layers. A site presents different layers of the same content in different areas of the site.

Let’s go back to the recipes. A cooking site can’t display a recipe in its entirety on every page of the site; that would be ridiculous. Instead, categories of recipes should show much less information about the recipe, but the right information. If I’m looking at a list of seasonal springtime dishes, just showing me the cuisine of each dish won’t give me much useful information. Instead, other pieces of information – a list of recipe names, a picture of the dish and maybe a couple of the key ingredients – help me learn just enough about the recipe to decide if I want to click further ahead.

The Principle of Exemplars
Describe the contents of categories by showing examples of the contents.

Where it comes from. Cognitive science has long studied how people categorize things. The field looks at what it means to hold a concept in your head. Ultimately, psychologists have discovered that our brains represent categories as networks of good examples.

How I use it. Whenever I display a category name, I provide a few examples of the content that lives in that category. This practice is perhaps less relevant on recipe websites and more relevant on corporate intranets. For the last intranet that I designed, I included a list of the main categories on the home page. Adjacent to each category name appeared a list of two or three items that best represent that category:

- Forms (W-4, equipment request, expense report)
- Policies (vacation, work from home, parental leave)

These links provide quick access to the most commonly used forms, but more importantly offer some insight as to what content lives inside the category.

Netflix has a great example. If you are a regular customer, the Netflix home page reveals a range of esoteric categories. Here are two of the categories on my Netflix home page:

- Emotional Dramas
- Understated Suspenseful TV Shows

Instead of providing a description of these categories, Netflix does three things for each category:

- Shows me movies or shows that I’ve already seen from these categories, providing a rationale for presenting them to me.
- Shows me four videos I have NOT seen.
- Provides a link to more movies or shows in this category.

By displaying six videos in the category, Netflix does a better job of helping me understand the category’s contents than any straight description.

The Principle of Front Doors
Assume at least half of the website’s visitors will come through some page other than the home page.

Where it comes from. This notion has become commonplace since even the larger sites find that most of their traffic comes through a side door, not the home page. These direct links are the power of search.

How I use it. The principle yields at least two practical pieces of advice. First, a destination page has to help users understand what else they can find on the site. By “destination page” I mean a page where users would land to consume content: a video on YouTube, a photo on Flickr or an article on WashingtonPost.com. Users coming to these pages through an outside search engine may find what they’re looking for, but that page also may have the responsibility to tell visitors where they are and what else they can do while they’re here.

Landing on a recipe for matzoh ball soup, a visitor to our recipe site should see the recipe itself, of course, and additional content for that recipe that adds some value (customer comments or reviews). However, the page also must take users to related recipes (like kugel) and perhaps provide a taste of what other kinds of recipes they can find on the site.

The second piece of advice coming from this principle is that the home page does not have to do everything. Good home pages exemplify the breadth of information on the site, but do not attempt to reveal every last
piece of content. They do not try to expose every category and provide a path to every nook and cranny. They instead focus on helping new users understand what the site is all about.

**The Principle of Multiple Classification**

Offer users several different classification schemes to browse the site’s content.

*Where it comes from.* Good information architecture acknowledges that people have different ways of looking for information. Even narrow audiences represent a diversity of motivations and mental models (how we imagine information space). A classification scheme attempts to provide simple ways for finding information across this range. A classification scheme describes what labels you will use to categorize the website’s content.

*How I use it.* Corporate intranets offer a ripe opportunity to use multiple classification schemes. Having worked on a few of these, I found a consistent pattern between organizations. People definitely use topics (like, vacations, for example) to find content on an intranet, but they also use content type (like, policy or form). They might also use department (HR, for example) but for some departments that is more meaningful and generally tied to the topic than for others. In short, then, either topic or content type frequently constitutes the starting point for finding information on an intranet. Navigation mechanisms incorporate both of these classification schemes in ways that let users employ them independently but also combine them as necessary.

This principle is a double-edged sword. Providing multiple ways to find content benefits users, but providing too many ways can overwhelm and distract them.

**The Principle of Focused Navigation**

Don’t mix apples and oranges in your navigation scheme.

*Where it comes from.* Lots of design teams use phrases like “global navigation” or “right rail” or “left menu bar” to refer to a menu that lets users browse content, but do so with its location on the page. Where a menu appears should not constitute its definition.

As for where it came from, it’s all me. Working with design teams in organizations that use phrases like “global nav,” I realized that the label makes it difficult for those teams to understand what the navigation menu is for. A navigation menu loses its purpose when its name comes from its location on the template.

*How I use it.* Designing navigation means establishing a strategy for finding content on the website. That strategy can entail several different navigation mechanisms – menus providing access to the content in different ways. On a content-focused website I designed recently, I incorporated four navigation mechanisms:

- **Topic navigation:** The site’s main navigation, which constituted the six main topic areas.
- **Timely navigation:** A short menu providing links to subtopics that were especially relevant or seasonal.
- **Signpost navigation:** A menu appearing on interior pages to reveal how the article was classified and giving users an opportunity to explore those categories.
- **Marketing navigation:** A short menu appearing adjacent to the topic navigation providing links to services offered by the organization.

**The Principle of Growth**

Assume the content you have today is a small fraction of the content you will have tomorrow.

*Where it comes from.* Of all the principles on this list, this one is at once both the most self-evident and the most confounding. We all know where it comes from: the web today has more content than it did last year and the year before and the year before that. The convenience of publishing content combined with inexpensive storage means people put stuff up, but don’t take stuff down.

Designing for the exponential proliferation of content is challenging. Designing a structure that can accommodate twice as much content as it can today is like building an armoire with twice as many drawers. The digital storage of information comes with (apparently) limitless space, but the
presentation of and access to that information must conform, on some levels, with the limitations imposed by physical space. Why? The people who browse the stacks at a library are the same people looking for content online: we have the same cognitive and ergonomic requirements.

**How I use it.** Unfortunately, no hard and fast rules come from this principle. It is at once inevitable (there will be more!) and impossible to predict (how much more? what more?). This is the crux of the challenge: a website can grow in several different directions at once:

- It can add more content to existing categories: More articles under a topic.
- It can add different types of content to existing categories: Videos in a topic that’s been dominated by text.
- It can create new categories: A brand new topic.

The approach I use depends on the previous principle – focused navigation. By assuming that the information space you design requires several different navigation mechanisms, you can accommodate different types of growth within each one. Anticipate various kinds of growth your site might endure and identify how each type of navigation deals with it.

For example, in a navigation mechanism that just deals with topics, the information architect can establish very broad top-level categories, making it easier to accommodate new topics as sub-categories. This menu doesn’t have to deal with any other forms of growth, just the addition of new topics.

To deal with new content types, the information architect can design the topic page anticipating new forms of content – video, presentations, photo galleries, podcasts – and design the page to accommodate them. Again, this page shouldn’t try to deal with any other forms of growth, just the possibility of new content types.

By splitting up the responsibilities of accommodating growth, a website can do so gracefully and hopefully with minimal fuss.

**Universal Principles of Information Architecture**

As Lidwell, Holden and Butler say in the introduction to their book, “The use of well-established design principles increases the probability that a design will be successful.” [2, p.13]

The key term here is “well-established.” Information architecture, though perhaps in its infancy compared to other design fields like building architecture and graphic design, can still establish a set of self-evident truths derived from other fields of design, experience and testing. These principles can constitute the foundation of a theory of information architecture, a framework for informing design decisions about the structures of websites (among other things).

What separates these principles from, say, those described in the Lidwell, Holden and Butler book is that, for now, these are my principles. I derived them from my work either organically or as adaptations of principles from other fields validated through experience. They’re the principles I bring with me to every new project, a psychic history of prior challenges, debates and lessons learned. They serve as the starting point for project-specific principles, a fertile ground for growing guidelines geared especially for a particular client, set of design challenges or project objective.

I hope you disagreed with them. At least one of them. A serious theoretical framework establishes a place for interpretation, reconsideration and debate. A theory of information architecture would escalate our conversations, taking us beyond “what do we do” and even “how do we do it” and into something far more interesting: “How do we do it better?”

---

**Resources Mentioned in the Article**


A graph is a data structure composed of dots (i.e., vertices) and lines (i.e., edges). The dots and lines of a graph can be organized into intricate arrangements. A graph’s ability to denote objects and their relationships to one another allows for a surprisingly large number of things to be modeled as graphs. From the dependencies that link software packages to the wood beams that provide the framing to a house, most anything has a corresponding graph representation. However, just because it is possible to represent something as a graph does not necessarily mean that its graph representation will be useful. If a modeler can leverage the plethora of tools and algorithms that store and process graphs, then such a mapping is worthwhile. This article explores the world of graphs in computing and exposes situations in which graphical models are beneficial.

The Bits and Pieces of the Dots and Lines

A model is a representation of some aspect of reality. Many models can be thought of as a collection of objects, such as people or concepts, and the relationships that exist between them, such as friendships or subclasses. Such objects and relations form a network. Graphically, an object in a network can be denoted by a dot, and a relationship can be denoted by a line. A structure formed by dots and lines is known as a graph – the mathematical term for a network [1]. The most common type of graph is the simple graph. An example is diagrammed in Figure 1. In a simple graph there are a set of vertices (dots) and a set of edges (lines), where edges are undirected and connect two unique vertices (that is, there are no loops), and no two edges exist between the same pair of vertices.

Despite the title of this article, dots and lines are not the only components in a graph modeler’s toolkit. There are many more bits and pieces in the world of graphs. In practice, rarely are vertices and edges the only data...
FIGURE 2. There are numerous types of graphs. Many of the formalisms described can be mixed and matched in order to provide the modeler the expressivity necessary to capture the essential features of a domain.

<table>
<thead>
<tr>
<th>Graph Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Half-edge graph</td>
<td>A unary edge (i.e., an edge that “connects” one vertex). It has limited practical application and is primarily discussed in mathematics.</td>
</tr>
<tr>
<td>Multi-graph</td>
<td>There are many situations in which it is desirable to have multiple edges between the same two vertices.</td>
</tr>
<tr>
<td>Simple graph</td>
<td>The prototypical graph, where an edge connects two vertices and no loops are allowed.</td>
</tr>
<tr>
<td>Weighted graph</td>
<td>Used to represent strength of ties or transition probabilities.</td>
</tr>
<tr>
<td>Vertex-labeled graph</td>
<td>Most every graph makes use of labeled vertices (e.g., an identifier).</td>
</tr>
<tr>
<td>Semantic graph</td>
<td>Used to model cognitive structures such as the relationship between concepts and the instances of those concepts [2]</td>
</tr>
<tr>
<td>Undirected graph</td>
<td>The typical graph that is used when the relationship is symmetric (e.g., friendship).</td>
</tr>
<tr>
<td>Directed graph</td>
<td>Orders the vertices of an edge to denote edge orientation.</td>
</tr>
<tr>
<td>Hypergraph</td>
<td>Generalizes a binary edge whereby an edge connects an arbitrary number of vertices [3].</td>
</tr>
<tr>
<td>Resource description framework graph</td>
<td>A graph standard developed by the World Wide Web consortium that denotes vertices and edges by uniform resource identifiers [4].</td>
</tr>
<tr>
<td>Edge-attributed graph</td>
<td>Used in applications where it is desirable to append non-relational metadata to an edge.</td>
</tr>
<tr>
<td>Pseudo graph</td>
<td>Used to denote a reflexive relationship.</td>
</tr>
</tbody>
</table>

Note that in many cases, these bits and pieces can be used in combination with one another, that is, they are not necessarily mutually exclusive. The list presented is not the complete space of all graph types, nor are the terms generally accepted in all domains. Many of these structures have been rediscovered in different domains and under different names. The important point is that there are numerous graph types and, consequently, there are systems and algorithms that exist to store and process them.

A common graph type supported by most graph systems is the directed, labeled, attributed, multi-graph – also known as a “property graph.” Graphs of this form allow for the representation of labeled vertices, labeled edges and attribute metadata (properties) for both vertices and edges. The property graph is common because modelers can express other types of graphs by simply abandoning or adding particular bits and pieces. For example, not allowing loops or multiple edges between two vertices generates a simple graph. Not allowing vertex/edge attributes generates a standard semantic graph. Restricting the vertex/edge labels to Uniform Resource Identifiers (URIs) generates a Resource Description Framework (RDF) graph (allowing for a few additional technicalities). Adding a weight attribute to an edge generates a weighted graph. The various graph types and the morphisms

-contained within a graph. For instance, sometimes it is useful to have a name associated with a vertex or a weight and direction associated with an edge. From primitive dots and lines various bits and pieces can be added to yield a more flexible, more expressive graph. Figure 2 diagrams a collection of different graph types while a short summary of each graph type illustrated is provided below:

- **Half-edge graph**: a unary edge (i.e., an edge that “connects” one vertex). It has limited practical application and is primarily discussed in mathematics.
- **Multi-graph**: There are many situations in which it is desirable to have multiple edges between the same two vertices.
- **Simple graph**: the prototypical graph, where an edge connects two vertices and no loops are allowed.
- **Weighted graph**: used to represent strength of ties or transition probabilities.
- **Vertex-labeled graph**: Most every graph makes use of labeled vertices (e.g., an identifier).
- **Semantic graph**: used to model cognitive structures such as the relationship between concepts and the instances of those concepts [2].
- **Vertex-attributed**: used in applications where it is desirable to append non-relational metadata to a vertex.
- **Edge-labeled graph**: used to denote the way in which two vertices are related (e.g., friendships, kinships, etc.).
- **Directed graph**: orders the vertices of an edge to denote edge orientation.
- **Hypergraph**: generalizes a binary edge whereby an edge connects an arbitrary number of vertices [3].
- **Undirected graph**: the typical graph that is used when the relationship is symmetric (e.g., friendship).
- **Resource description framework graph**: a graph standard developed by the World Wide Web consortium that denotes vertices and edges by uniform resource identifiers [4].
- **Edge-attributed graph**: used in applications where it is desirable to append non-relational metadata to an edge.
- **Pseudo graph**: used to denote a reflexive relationship.
that yield one graph type from another are diagrammed in Figure 3. Note the location of the property graph within this diagram. Finally, while it is possible to model a hypergraph in a property graph, it comes at the expense of using vertices in the property graph to denote both vertices and edges in the hypergraph. For this reason, there are specialized hypergraph systems, such as HyperGraphDB. For the remainder of this article, the more common property graph and its supporting technologies are discussed.

![Property Graph Diagram](image)

**FIGURE 3.** The property graph is a convenient structure because it contains most of the bits and pieces used in graph modeling. Simple morphisms of the property graph yield other common graph structures. Thus, graph systems that support the property graph data model also, implicitly support other graph types.

Preserving Dots and Lines

The computer science community has recently seen an explosion of database technologies. For decades, the relational database of Codd’s relational algebra has been the primary storage and query mechanism for large data sets [5]. However, with the continued growth of data and an increasingly variegated application landscape, new databases have emerged. In this space, no database is seen as the single solution to all problems. Instead, each database attempts to solve a particular data management issue. Itemized below are short descriptions of recent database types:

- **Document database:** These databases have the “document” as their atomic entity. Such objects are semi-structured and usually represented in XML (Extended Markup Language) or JSON (JavaScript Object Notation). A document can be retrieved by means of pattern matching a query document (that is, a semi-populated document) against all the documents contained in the database. The benefit of this model is that these databases scale horizontally with relative ease. This ability is due to the fact that documents lack references between one another. The drawback is that data is not interrelated and thus, cross-database analyses are costly. For many web applications the document database is a very suitable solution that supports data scale and a convenient symmetry between the document structure and processing languages such as those that natively support XML and/or JSON. Examples of such databases include MongoDB and CouchDB.

- **Key/value store:** This family of databases is focused on scaling large amounts of data over a large number of machines and, in turn, supporting heavy read/write loads. Most of the databases in this class were inspired by Amazon’s Dynamo [6]. A popular open-source key/value store is Tokyo Cabinet.

- **Triple/quad store:** Triple/quad stores were developed to support the demands of the Semantic Web/Web of Data/Linked Data community. These databases are optimized for storing and querying data represented according to the Resource Description Framework (RDF) [4]. Typical use cases include description logic reasoning [7] and SPARQL-based graph pattern matching [8]. AllegroGraph is a high-performance quad store with a large suite of extensions and features.

- **Column store:** Most column stores are modeled after Google’s BigTable database [9]. A big table is a sparse, distributed, persistent multi-dimensional sorted map. The map is indexed by a row key, column key, and a time-stamp. Real-world services implemented with BigTable include GoogleAnalytics and GoogleEarth. Cassandra is a popular open-source column store.
**Graph database**: Graph databases are optimized for the efficient processing of dense, interrelated datasets. In these databases, the atomic entity is the graph as a whole. The typical data model is the property graph. By supporting the interrelation of data, graph databases allow for fast traversals along the edges between vertices [10]. A popular graph database of this form is Neo4j.

There are numerous databases in this growing space that were not mentioned. Moreover, there are other database types not mentioned. It is out of the scope of this article to explore this space in depth. The interested reader is directed to related discussions, blog posts and presentations that are made freely available on the Internet. Of particular relevance to this article are the graph database and the property graph data model. Figure 4 diagrams a property graph containing people, their articles and a university. In this particular domain model, each vertex has a name property and a type property. Edges denote both a directionality and a relationship type (that is, an edge label). Moreover, it is possible to also include properties on an edge to further refine the way in which two vertices are related (for example, Josh started attending RPI in 2007).

A consequence of the flexibility of a graph is that other related data can be represented as graph structures along with the domain model. A typical case of the use of such graph extensions is the endogenous index. An index is usually a tree-structure that allows for the fast look-up of elements within a collection. If there were no indices into a collection, then to determine if a particular element had a particular property, each element in the collection would have to be examined. Assuming that the cost of a linear scan of this kind is O(n), where n is the number of elements, an index provides the ability to partition the elements into increasingly fine-grained bins and thus to reduce the lookup cost to O(log₂ n) in most cases. While an index creates more data (the tree structure), it makes up for this cost by greatly increasing the speed of element retrieval. Figure 5 shows a name-property index over the example graph diagrammed in Figure 4. Together, the domain model and the index of the domain model are seen as a single atomic entity. Searching for an element and moving between elements are accomplished by a unified framework: the graph traversal.

**Jumping from Dot to Dot**

The first aspect of using a graph is creating a graph. Once a graph has
been created, it can be subjected to algorithms that quantify aspects of its structure, alter its structure or solve problems that are a function of its structure. At the root of any of these algorithms is the graph traversal [10]. A graph traversal is a walk along the elements of a graph – from vertex to edge to vertex, etc. As this walk proceeds, aspects of the graph can be saved or manipulated and in general, an algorithm can be computed. In principle, any of the data models and databases presented in the previous section (and including typical relational databases) can be used to represent and process a graph. However, when traversing a graph is the ultimate use case for a graph data set, then a graph database is the optimal solution, an import point. A graph database is optimized for graph traversals because elements (vertices and edges) maintain direct references to their adjacent elements. It is this design choice that makes traversing a graph structure within a graph database fast and efficient.

To get a better understanding of how graph traversals work, the examples in this section will be expressed in terms of a graph programming language called Gremlin. In Gremlin, moving over vertices and edges is analogous in many ways to moving through the directory structure of a local file system. To demonstrate, a naive friend-of-a-friend query is represented as follows:

```
./outE[@label='friend']/inV/outE[@label='friend']/inV
```

Reading from left to right, this expression states

Start at the root vertex (., that is, the vertex to evaluate the expression on). Traverse to all the outgoing edges of the root vertex (/outE).
Filter out all edges that are not labeled “friend” (@label='friend').
For all those friend-labeled edges, go to their incoming/head vertices (/inV).
For all the friends of the root vertex, get their outgoing edges (/outE).
Filter out all edges that are not labeled “friend” (@label='friend').
For all those friend-labeled edges, go to their incoming/head vertices (/inV).

At the end of this expression, the resultant vertices are the friends of the friends of the root vertex. Figure 6a diagrams the traversal, where the grey vertices are the returned vertices. This example is naive because in many cases, it is important to retrieve the root vertex’s friends of friends that are not also its friends. In such situations, the traverser must remember if a located friend-of-friend is not already a friend. In order to calculate the friend-of-a-friend, the friends must be determined first. Therefore, it is possible to save this information for later use. This idea is diagrammed in Figure 6b and the Gremlin expression is presented below, where the variable $x$ references the friends of the root vertex.

```
./outE[@label='friend']/inV[g:assign('$x')]/
outE[@label='friend']/inV[g:except($x)]
```

An important aspect of working with property graphs is that the edges are typed/labeled. The standard suites of graph algorithms found in most graph/network-theory textbooks are not immediately useful for property graphs [11] because most graph algorithms have been developed for unlabeled graphs. When vertices can be related in many different ways and vertices can represent various types of objects, the meaning of the rankings, paths and other features returned by standard graph algorithms are ambiguous. However, by interpreting a path through a graph as an edge, it is possible to express standard graph algorithms on property graphs [12]. The previously presented Gremlin expression followed a path from the root vertex to its friends’ friends.
This path can be considered a virtual (that is, inferred, derived) edge. From the perspective of this expression, a new implicit graph is created over the graph’s vertices that only contains edges labeled “friend-of-a-friend.” This idea is diagrammed in Figure 7. As such, this virtual graph is equivalent to an unlabeled graph because all edges having the same meaning. Therefore, all the standard graph algorithms can be meaningfully applied to this derived graph – for example, the shortest path between person A and person B through their friends of friends. The benefit of edge-labeled graphs such as property graphs is that there are as many types of rankings, scorings and so forth as there are types of paths that exist between the elements of the graph.

Conclusion

The concept of a graph was introduced in the late 19th century. During the many decades that followed, the world of graphs was primarily left to the toiling of mathematicians. In the last few decades, the sociology, physics and computer science communities introduced a suite of algorithms and insightful realizations about the nature of graphs found in the real world. Moreover, the increasingly large volume of data made available by the Internet has yielded datasets that reflect the graphs found in our technological and social systems. To satiate the need to handle and process these large-scale graphs, graph databases have come to the forefront. To make use of the graphs beyond simply representing their explicit structure, graph traversal frameworks and algorithms have been developed in order to shape graphs by driving the evolution of the entities that they model – for example, humans and their relationships to one another and the objects of their world [13].

Resources Mentioned in the Article

### Resources Mentioned in the Article, continued


### Systems and Languages Mentioned in the Article

- **AllegroGraph**: [www.franz.com/agraph/allegrograph/](http://www.franz.com/agraph/allegrograph/)
- **CouchDB**: [http://couchdb.apache.org/](http://couchdb.apache.org/)
- **HyperGraphDB**: [www.kobrix.com/hgdb.jsp](http://www.kobrix.com/hgdb.jsp)
- **JSON**: [www.json.org](http://www.json.org)
- **MongoDB**: [www.mongodb.org/](http://www.mongodb.org/)
- **Neo4j**: [http://neo4j.org/](http://neo4j.org/)
- **XML**: [www.w3.org/XML](http://www.w3.org/XML)
The American Society for Information Science and Technology sponsored a Research Data Access and Preservation Summit in Phoenix, Arizona, on April 9-10, 2010. The Summit was chaired by Christine Borgman, professor in the School of Information and Library Science at the University of California, Los Angeles, and by Reagan Moore and Gary Marchionini, 2010 ASIS&T president. Moore and Marchionini are professors in the School of Information and Library Science at the University of North Carolina at Chapel Hill. The primary goal of the Summit was to build a broad perspective on the needs of scientific data management communities, the technology approaches going into production data management systems in support of research and the legal and social implications of sharing data. Objectives for the Summit included the characterization of large-scale data management needs, the identification of representative data management systems and the identification of interoperability practices.

An example of an emerging data management environment is the Australian effort to build a national data management infrastructure that supports all Australian researchers. In the United States, similar National Science Foundation funded efforts are exemplified by the DataNet projects. Infrastructure is being developed to support a wide range of science and engineering disciplines and all of the phases of the scientific data life cycle. By sharing data management ideas across research projects, hearing the actual mechanisms being used to implement institutional repositories and learning the social imperatives behind formation of shared collections, the Summit hoped to build a context within which a national data management infrastructure can be defined.

The Summit had 75 participants from academic libraries, federal repositories and private companies. Most of the participants were from the United States, but participants also came from Australia, Canada, the Netherlands and the United Kingdom. An advisory committee collaborated on the design of the Summit, proposing topics and perspectives for organizing the meeting. The advisory committee members are listed in the accompanying box.

Advisory Committee Members

William Anderson, Praxis101 and University of Texas at Austin
Christine Borgman, University of California, Los Angeles
Hsinchun Chen, University of Arizona
Sayeed Choudhury, Johns Hopkins University
Michael Lesk, Rutgers University
Gary Marchionini, Johns Hopkins University
Reagan Moore, University of North Carolina at Chapel Hill
William Michener, University of New Mexico
Art Pasquinelli, Oracle
Sudha Ram, University of Arizona
Stu Weibel, OCLC

The Summit was organized around a progressive presentation of the challenges facing groups that manage scientific data. To provide as much expertise as possible, six panels were created, each one focused on a specific topic. The expectation was that
each panel would define basic concepts, present the state of the art and identify dominant research questions. Each panel would then respond to questions from the Summit participants. A key motivation of this report is to track the issues that resonated with Summit attendees and identify new issues to indicate areas where a future conference might place emphasis.

**Summit Program: Panel Topics and Discussion**

The Summit panels were organized around six aspects of scientific data management. The program with links to the associated presentation files is found here: [http://www.asis.org/Conferences/RDAP10/RDAP10Program.html](http://www.asis.org/Conferences/RDAP10/RDAP10Program.html)

A summary of the topics and discussion follows.

- **Panel 1: Data Life Cycle Management**
  - Mark McFarland, Texas Digital Library
  - Erin O’Meara, Carolina Digital Repository
  - Stacy Kowalczyk, Indiana University
  - Cynthia Ghering, Michigan State University
  
  The panel experts were from universities implementing digital libraries or institutional repositories. Since most of the participants were from academic institutions, this representation provided a firm grounding in the digital data management approaches being pursued by university libraries. One long-term goal for this group is the integration of scientific data collections into reference collections housed within institutional repositories.

- **Panel 2: Promoting Re-use of Scientific Data Collections**
  - Peter Wittenburg, Max Planck Society, Institute for Psycholinguistics
  - Sudha Ram, iPlant Collaborative
  - Roy Williams, International Virtual Observatory Alliance
  - John Harrison, Sustaining Heritage Access through Multivalent Archiving
  - Jonathan Crabtree, Odum Social Science Institute
  
  The panel experts were from national-scale research projects that face the challenges of organizing scientific data collections for use by multiple researchers. This application requires ingest of observational data, experimental data or simulation output and the organization of the data into sharable collections, as well as the publication of the data for use by the broader discipline and the preservation of the data as reference collection for use by future researchers. At each stage of the data life cycle, a broader community re-purposes the data for their specific uses. This perspective emphasizes social implications of use, technical challenges for managing massive amounts of data and expectations for how the data will be accessed and used.

- **Panel 3: Large-Scale Data Management Challenges**
  - John Graybeal, Ocean Observatories Initiative
  - Ken Gallup, Renaissance Computing Institute
  - Laurin Herr, CineGrid
  - Philip Maechling, Southern California Earthquake Center
  
  The panel experts were from national-scale projects faced with massive data collections (petabytes to hundreds of petabytes of data), management of sensor data streams and management of highly distributed data. Since every data management environment needs to distribute data across multiple locations to minimize risk of data loss, the distributed data management approaches were expected to point to technologies that would be beneficial to everyone. The panelists represented projects that are assembling some of the largest academic data collections being created.

- **Panel 4: DataNet Federation**
  - Sayeed Choudhury, Data Conservancy
  - William Michener, DataOne
  - Reagan Moore, Data Grids
  
  The panel experts included funded projects from the National Science Foundation (NSF) DataNet solicitation and current data grid software technology developers.
NSF is pursuing development of infrastructure to support digital libraries, large scientific data collections, the full data life cycle and long-term sustainability. The expectation is that technology developed within the DataNet partners will be useful to institutional repositories and digital libraries, as well as large-scale data grids and preservation environments. The issue of interoperability is critical for linking data management solutions across all stages of the data life cycle. Data grids provide infrastructure for shared collections that span multiple storage environments.

- **Panel 5: Developing Assessment Criteria**
  Mark Conrad, National Archives and Records Administration
  Jane Greenberg, University of North Carolina at Chapel Hill and the Dryad Project
  John Graybeal, Ocean Observatories Initiative
  Steve Richard, Arizona Geological Society

  The panel experts represented national projects that are developing domain-specific descriptions for collection context, defining the properties that the collection should maintain and proposing criteria to validate the collection properties. Each community is tackling these challenges on its own – developing specific semantics, ontologies, policies and assessment criteria. A long-range goal is the identification of common properties that all collections should manage, common policies and shared semantics.

- **Panel 6: Legal and Social Implications of Shared Collections**
  Ann Zimmerman, University of Michigan
  Melissa Cragin, University of Illinois
  Noshir Contractor, Northwestern University

  The panel experts represented projects that are assessing motivations for sharing data, identifying the criteria under which collection sharing takes place and considering development of incentives for sharing data.

This panel described the motivations for the construction of institutional repositories as well as those for disciplines to build shared collections. The social aspects of sharing strongly influence the success of attempts to re-purpose collections for new uses and develop reference collections for future researchers. A central question is the identification of incentives for researchers to share their data.

Two additional sessions were held to allow technology developers to present live demonstrations of their systems and give tutorials on use of the software. A poster and demonstration session was held on Friday evening, and a tutorial session was held on Saturday afternoon. The Friday evening session provided demonstrations of the LStore high-performance distributed data management system (Alan Tackett), the Sustainable Heritage through Multivalent Archiving (SHAMAN) preservation technology (John Harrison), the Fedora digital library middleware (Brad McLean) and the iRODS integrated Rule-Oriented Data System (Reagan Moore). The Saturday session included tutorials on Fedora (Brad McLean) and iRODS (Reagan Moore).

**Summary Conclusions, Hypotheses and Questions**

**Conclusions**
1. Institutional repositories are emerging within academia with the goal of building digital reference collections.
2. National-scale research projects are also assembling scientific data collections that represent the digital holdings for science and engineering disciplines.
3. The scale of digital collections is increasing beyond the capacity of institutional repositories, with major holdings housed in federal repositories.
4. The environments of institutional repositories, discipline-specific collections and federal repositories need to interoperate.
5. Initiatives are starting to drive formation of national-scale data infrastructure (NSF DataNet, NSF Teragrid, DOE Open Science Grid, DOE Earth Systems Grid, NOAA CLASS system, NASA DAACs). These systems need to interoperate.
6. Standards are being developed for digital data description (provenance, context), digital data representation (structure, processing mechanisms), digital repository trustworthiness (assessment criteria). These standards need to interoperate.
7. An understanding is being developed of the social motivations to encourage data sharing, formation of shared collections, collaborative research. The creation of a shared collection is a social process, requiring consensus on the properties that the digital objects within the shared collection will possess, the policies that will be used to manage the desired properties, the procedures that will be executed to enforce the properties and the assessment criteria that will be used to validate the original intent for forming the shared collection.

Hypotheses and other questions about the futures of scientific data collections
1. Will there be convergence among the multiple types of data management applications (data grid, digital library, persistent archive, data processing pipeline)? Can data management applications be characterized by the procedures and policies that control the collection?
2. Given that a social consensus drives the formation of a shared collection, what kind of mechanisms will emerge for communities to develop a consensus on collection sharing?
3. Given the massive size of emerging collections (tens to hundreds of petabytes), what sorts of computing and storage integration are essential? Will massive collections force the movement of the application to the data, instead of moving the data to the application?
4. Does the federation of collections drive long-term sustainability between institutions and the re-purposing of collections for use by new communities? What sorts of social agreements will emerge defining policies controlling re-use of collections? Can long-term sustainability be turned into a policy on collection re-use by requiring that the original use be sustained as an alternate set of procedures and policies within the new environment?
5. What forms of distributed data will become common across all data management applications to minimize risk of data loss, to ensure experts have a local copy and to facilitate data analyses?
6. Is a national data grid feasible? Could such a grid support research initiatives across federal agencies?
7. What kinds of federation of federal repositories are feasible? Can a unifying data management infrastructure support all federal activities?
8. How can social networks be formed to promote sharing and maintenance of data collections? What new mechanisms for social interaction are needed to promote development of consensus and agreements by communities?
9. Given that the multiple stages of the data life cycle correspond to re-purposing of a collection for use by broader communities, what are the social mechanisms for formation and maintenance of these broader communities?
10. Will collections become the standard for organizing data instead of file systems? The context provided by collections is essential for understanding how to discover, browse, access and manipulate data.

Participant Commentary: Summit Tweets
Summit participants used Twitter and the hashtag #rdap10 to share information about presentations and discussions. An online notebook of all the tweets so tagged is found here: www.twapperkeeper.com/hashtag/rdap10