“The humanities and social sciences are concerned with the human experience. Sciences, too, deal with actions, processes and interactions. Information systems, therefore, are concerned with events, but can operate only on objects (bits, books, “documents”) – and events are not objects.”

Special Section begins on page 12.

### Annual Meeting Coverage

#### Special Section | Foundations of Information Science

**[12] Narratives, Facts and Events in the Foundations of Information Science**  
by Michael Buckland, Guest Editor

**[13] From Facts to Judgments: Theorizing History for Information Science**  
by Ryan Shaw

**[19] Facts and Frameworks in Paul Otlet’s and Julius Otto Kaiser’s Theories of Knowledge Organization**  
by Thomas M. Dousa

**[26] Events as a Structuring Device in Biographical Mark-up and Metadata**  
by Michael Buckland and Michele Renee Ramos

#### ASIS&T 2009 Award of Merit Speech

**[30] Facilitating and Enhancing Scientific Discovery**  
by Carol Tenopir

---

### Departments

[2] Editor’s Desktop

[3] President’s Page

*The ASIS&T Version of “Stimulus Funding”*
  by Donald Case, outgoing ASIS&T president

**ASIS&T Participation Agenda**
  by Gary Marchionini, incoming ASIS&T president

[7] Inside ASIS&T

[10] ASIS&T Information Professionals Task Force: October 2009 Update
  by Nancy Roderer

### IA Column

  by Thom Haller
I have just returned from the ASIS&T Annual Meeting in Vancouver. It was well attended and had a number of very interesting sessions – user information behavior and data curation seemed to be particularly hot topics. The poster sessions were also extensive and very popular – the coffee helped. And there was free wi-fi in the reception area! We look forward to possible additional joint meetings with our excellent hosts, the Canadian Association for Information Science.

The Society also announced at the meeting that next year’s Annual Meeting in Pittsburgh will be held in conjunction with the DC 2010, the Dublin Core Metadata Initiative Conference, and a summit Research Data Access and Preservation will be held in conjunction with the 2010 IA Summit in Phoenix in April. Also, Information Today has purchased the annual Search Engine Conference, and ASIS&T members will be able to get better prices for that event. The Board of Directors also announced that JASIST will begin to carry six literature review articles each year to replace the Annual Review of Information Science and Technology. Many authors consider it a more visible venue, and the journal will benefit from the increased usage that such articles generate.

Because of the timing of this issue, much reporting of the Annual Meeting will be postponed to our February/March issue, but we are able to begin our coverage of the meeting with a few special features in this issue. These items include President’s Pages from both the outgoing and incoming ASIS&T presidents; the speech given at the Awards Luncheon by Carol Tenopir, the recipient of this year’s Award of Merit; and an update on the work of the information professionals task force. Also, our special section in this issue is based on an Annual Meeting panel covering the role of facts and events in information retrieval.

Finally, Thom Haller, our new associate editor for information architecture, discusses what motivates him to pursue this profession.
Message from the Outgoing President

The ASIS&T Version of “Stimulus Funding”

by Donald Case

As I write this column there are renewed calls for national governments to spend money on public projects that promise to stimulate their economies. Even though some economic indicators have been improving, unemployment is up while most real estate and retail sales remain down – at least in North America and Europe. The thinking among some economic advisors is that more intervention is needed to jumpstart spending and thus stimulate job growth.

Within ASIS&T we have our own version of such “stimulus funding,” however it is not targeted towards economic activity but rather towards the long-term well-being of the Society and its membership. Given the healthy income stream that ASIS&T receives from its publishing contract, coupled with surpluses from previous years, the Board has proposed increased spending on a broad range of member benefits.

In my previous column I wrote about most of the objectives that the Board had identified in its annual retreat (and in earlier meetings). These include such actions as increasing the rate for funding of regional and student chapters; funding more scholarships for students to attend the Annual Meeting; broadening the “hardship” category so that more low-income members can continue membership benefits when they need them most and at a subsidized rate; and hiring a part-time “interaction manager” devoted to making ASIS&T more participatory by promoting social media (such as blogs, wikis and Second Life).

Donald Case has just completed his term as 2009 ASIS&T president and will now serve one year as past president. He is a professor in the University of Kentucky College of Communication and Information Studies. He can be reached at dcase<at>uky.edu.

Message from the Incoming President

ASIS&T Participation Agenda

by Gary Marchionini

ASIS&T is the leading society devoted to concepts and issues in information science, and I am honored and enthusiastic to serve as president for 2010. As ASIS&T enters the second decade of the 21st century we find new possibilities for expanding member participation in a myriad of information activities and for our society to participate more actively in the international discourse related to information concepts, issues and policies. I aim to focus my attention on increasing participation in ASIS&T on several fronts and in helping ASIS&T take on new leadership roles in information science as a vibrant field that supports a broad range of professionals.

This general mission will be operationalized on several fronts in the year ahead. Some will require new initiatives; others will require changes to our mode of operation and all will need the active participation and engagement of ASIS&T members. Today’s social media offer substantial new opportunities for broader and more frequent information exchange, and professional societies are poised to take advantage of Douglas Engelbart’s bootstrapping ideas – just as information professionals adopted his ideas of computational cognitive augmentation in the past. Participation takes many forms, but always entails action. In our frenetic world of ubiquitous information flows and viral trends, it is evident that even small actions by people at scale lead to significant effects. I ask that every ASIS&T
to keep in touch with members, generate content and stimulate professional activity and interaction within the Society.

Ah, but here is the rub: Doing all that we want to do requires a deficit budget for the fiscal year 2009-2010. Not only will new initiatives require financial outlays starting in 2010, but as I write this column the outlook is uncertain for the attendance at the Annual Meeting in Vancouver. Given the cutbacks in travel funds at many institutions in which our members work, coupled with an unusual location for the Annual Meeting, it is possible that the conference will not create the surplus that it usually does; non-East coast meetings tend to draw fewer members, and travel to this one is more expensive than usual (although well worth the price, in my opinion).

This led some of us into protracted discussion about what we could do to balance the budget even if the 2009 Annual Meeting were to suffer a small loss. We considered delaying certain expenses, such as the interaction manager post, but ultimately decided that all these initiatives were too urgently needed to delay any longer; like nearly all associations these days, our membership base continues to slowly erode; we need to improve our value to our members. Therefore we have proposed a modest deficit for this coming year.

Fortunately the latest word on advance registrations puts the Vancouver Annual Meeting on a track comparable to recent meetings in eastern U.S. cities, and there are early indications that we will get the additional Asian and European attendees that originally led the Board to advocate the British Columbia location. Internationalizing the Society continues to be a priority for us, and eventually that path will lead to an enriched, larger and more participatory Society. Another encouraging sign is a growing number of positions available and candidates seeking positions that were in the online system for the Annual Meeting.

We are not counting on the 2009 Annual Meeting to generate a surplus that would balance the projected deficit. Yet, we are confident that the Vancouver meeting, like the initiatives mentioned above, will create long-term benefits for the Society.

MARCHIONINI, continued

member commit to taking some new action for ASIS&T. The following categories of action serve as a beginning agenda, and I welcome your suggestions for expanding the agenda (www.asis.org/wiki/bits/ is one place to post ideas).

Membership

It will be challenging to grow membership, given the current economy. I see a two-prong strategy. First, I strongly believe that the growth trajectory is in international members, especially from Europe and Asia/Australia. We have an active European Chapter, and I hope that we can add several new chapters around the globe. Many of us have colleagues in these regions, and I encourage everyone to reach out and welcome them to ASIS&T. Second, I believe that personal social networks are powerful means to attract members and leaders. I encourage all ASIS&T members to recruit their own personal contacts to ASIS&T and to use social networking tools to maintain and extend ASIS&T connections.

Leadership

For ASIS&T to flourish we must continue to inspire and nourish young information scientists to make ASIS&T their primary professional organization. In addition to leadership within the society it is essential that ASIS&T be the lead society in defining and promoting information science as a field and profession. I challenge each ASIS&T member in the following ways:

- Promote ASIS&T in some tangible way. For example, as you Twitter, blog, or participate in other social networking activities, mention ASIS&T or specifically write about ASIS&T. Join and participate in ASIS&T groups on Facebook and other services. Put your ASIS&T affiliation on your web pages and mention your ASIS&T affiliation when writing position statements or other papers.

- Invite and mentor one or more potential ASIS&T leaders to join the society and help shape the future of our field.
We can focus on people we know who are intellectually curious, who may be isolated in their professional roles and who have not been ASIS&T members.

- Make ASIS&T synonymous with IS. Promote the ASIS&T brand via connecting ASIS&T to information research, education and outreach events and services. As you plan and attend events, bring handouts or contact ASIS&T headquarters for promotional items to share. Give trip reports about ASIS&T events or activities to colleagues at work.

- Foster information science as a field. For example, volunteer in K-16 settings to talk to students about information science. Share articles from ASIS&T publications at government and policy meetings.

- Forge cooperative ventures with other societies to advance national and international agendas. As you see opportunities for ASIS&T to participate in collaborative position statements or events, discuss ways to get ASIS&T leaders and members involved.

**Participation as a Continual Process**

ASIS&T members must think and act beyond the Annual Meeting and other formal event inflection points. This entails re-conceptualizing our formal events and increasing our ASIS&T participation on a more regular basis.

I believe we can restructure the Annual Meetings on several fronts to make them more interactive and ongoing. Some goals include

- Link the Annual Meeting and Summit(s), chapter events and SIG events. For example, add short “best of” summaries from the Annual Meeting at local events and add short summaries or reprises of presentations from chapter and SIG events at the Annual Meeting.

- Increase use of virtual participation. Add video contributions and social networking services before, during and after ASIS&T meetings. Invite participation via remote connections (e.g., Skype appearances by commentators or reporters).

- Broaden the use of special sessions or tracks that broaden participation at meetings. For example, add an industry track or topic tracks to meetings; add more spontaneous and inclusive sessions.

- Propose and organize new ASIS&T summits on important topics. For 2010, we will introduce a new **Summit on Research Data Access and Preservation**. The summit will take place in Phoenix, April 9-10, 2010, at the Hyatt Regency Hotel in conjunction with the IA Summit. Please check the ASIS&T website for details. It will be chaired by Reagan Moore and has an advisory board that includes William Anderson, Christine Borgman, Hsinchun Chen, Sayeed Choudhury, Michael Lesk, Gary Marchionini, William Michener, Art Pasquinelli, Sudha Ram and Stu Weibel. The summit will address three main questions: (1) What data access and preservation capabilities are required within and across research groups? (2) What technical solutions exist to meet these needs and how do they scale across domains? (3) What are the social contexts under which research communities assemble to share data?

To increase participation on a daily basis, the following activities are encouraged:

- Reinvigorate communication media (*Bulletin*, blog, website, Facebook) by following and adding your voice through comments and suggestions. Create your own blogs, groups and discussion forums to address information topics and issues that affect your career or interests.

- Increase roles in information science career development and education. Use ASIS&T publications and communications streams as continuing education for yourself and your colleagues. Participate in formal training...
and certification functions at schools and universities. Help develop the ASIS&T career placement services (for example, use the Annual Meeting placement service, post positions to the ASIS&T Jobline, help human resource departments understand the value of information science degrees).

- Leverage social networking to personally invite participation both physically and virtually.

- Make chapter and SIG events more broadly accessible. For example, put highlights of chapter and SIG events on websites, use newsfeeds or other streams to share and promote these events. Be sure that your local media know about ASIS&T events.

These ideas are a first step and I welcome additional ideas and improvements on these to extend participation in ASIS&T and participation by ASIS&T.
New Officers and Directors Join ASIS&T Board

Each year at the ASIS&T Annual Meeting, a new administrative year begins, and the first official order of business is the introduction of new faces to the ASIS&T Board of Directors. In November in Vancouver, the annual changing of the guard took place, and new officers and directors took their seats on the Board.

Each of the positions filled through the summer balloting process is for three-year terms; in addition, one unexpired seat on the Board was also filled. Those elected are Linda C. Smith, president-elect; and Karen Fisher and Prudence Dalrymple, three-year terms as directors-at-large, and France Bouthillier, to the two remaining years left vacant by the resignation of Peter Morville.

As the new members took their seats, Gary Marchionini, elected last year as president-elect assumed the presidency from Donald Case, who continues on the Board as past president for one year. Linda C. Smith is professor and associate dean for academic programs in the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, where she has been a member of the faculty since 1977.

With support from the Institute of Museum and Library Services, she has worked with colleagues to enhance curricular offerings in digital libraries and data curation and to build a collaborative distance education model through initiatives of the WISE (Web-based Information Science Education) consortium. She received the ASIS&T Outstanding Information Science Teacher Award in 1987 and has served ASIS&T in many capacities since joining in 1972: SIG and chapter officer and advisor, award jury member and chair, conference program committee member and chair, member of various standing committees, and editorial positions with both JASIST and ARIST.

Karen Fisher is professor in the University of Washington Information School and chair of the Information & Society Center (ISC). She teaches and conducts research on how people experience information as part of everyday life. Frequent author, she is currently studying the benefits of free access to computers and the Internet in public libraries, supported by the Institute of Museum and Library Services and the Bill & Melinda Gates Foundation. Karen holds Ph.D. and MLIS degrees from University of Western Ontario and a bachelor’s degree from Memorial University of Newfoundland. Previously, she chaired the UW iSchool’s Master of Library and Information Science program, was a visiting researcher at Microsoft Research and a visiting professor at Oslo University College, Norway. Her ASIS&T activities include leadership activities in SIG/USE, service on the search committee for the JASIST editor and on the program committees for two Annual Meetings.

Prudence Dalrymple is on the faculty of the College of Information Science and Technology at Drexel University where she also directs the Institute for Healthcare Informatics. She holds a doctorate from the University of Wisconsin-Madison and a master’s degree in health sciences informatics from Johns Hopkins University School of Medicine. She has also been on the faculty at the University

The International Calendar of Information Science Conferences (http://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.
of Illinois at Urbana-Champaign, was the director of accreditation at the American Library Association and served as dean of the Graduate School of Library and Information Science at Dominican University. Pru attended her first ASIS&T meeting in 1976 as a student and has been an active participant in conferences, committees and SIG/HFIS. She has served on several award juries and standing committees. She is a member of the ASIS&T Task Force on the Information Professional where she is a key participant in the task force’s activities.

France Bouthillier is associate professor and director of the school of information studies, McGill University. She received her Ph.D. from the University of Toronto, her masters in library and information sciences from the Université de Montréal, and she graduated from the University of Quebec in Education and Business Administration. Her teaching areas are related to the management of information services, business information and competitive intelligence. Her recent publications and conference presentations deal with the information needs of small businesses, knowledge management, management of information services, training of information professionals and competitive intelligence. She has worked on program committees for several ASIS&T Annual Meetings and as a reviewer for JASIST for a number of years.

ASIS&T 2009 Annual Meeting Coverage

As this issue of the *Bulletin of the American Society for Information Science and Technology* is being prepared for online access, ASIS&T members are at the 2009 Annual Meeting in Vancouver. Significant coverage of the meeting, including photographic montages, will be included in the February/March 2010 issue of the *Bulletin*, but in this issue we are able to offer a few peeks at the proceedings. First is this list of the winners of the 2009 ASIS&T Awards for which more details will be provided in the next issue. In addition, the articles in the feature package in this issue, including the Special Section as well as articles from the winners of both the ASIS&T Award of Merit and James M. Cretsos Leadership Award, are drawn from the Annual Meeting. Look for more Annual Meeting coverage in our next issue.

Award of Merit – Carol Tenopir

Watson Davis Award – Edie Rasmussen

Thomson Reuters Outstanding Information Science Teacher Award – Diane Kelly


James M. Cretsos Leadership Award – Cassidy R. Sugimoto

Pratt-Severn Best Student Research Paper Award – Katie O’Leary, University of British Columbia, for “Information Seeking in the Context of a Hobby: A Case Study of a Young Adult with Asperger’s Syndrome”

Thomson Reuters Doctoral Dissertation Proposal Scholarship – Heather Piwowar for *Foundational Studies for Measuring the Impact, Prevalence and Patterns of Publicly Shared Biomedical Research Data*

ProQuest/ASIS&T Doctoral Dissertation Award – Luanne Freund for *Exploiting Task-Document Relations in Support of Information Retrieval in the Workplace*

Thomson Reuters Citation Analysis Research Grant – Cassidy Sugimoto for *Measuring Interdisciplinarity: An Exploration of a Novel Metric Applied to ILS Dissertations*

Chapter-of-the-Year – New England Chapter of ASIS&T (NEASIS&T) and Potomac Valley Chapter (PVC)
**Inside ASIS&T**

**News about ASIS&T Chapters**

The **Central Ohio ASIS&T chapter**, with co-sponsorship provided by the **Northern Ohio Chapter of ASIS&T** and the Cleveland chapter of SLA, presented a live broadcast of *Geo Tagging Digital Photography for Learning, Research and Just Plain Fun!* from Kent State University. Presenter **Mace W. Mentch**, assessment specialist in instructional technology and academic computing for Case Western Reserve University, discussed the geo tagging of photographs. He also demonstrated geo tagging devices, such as GiSTEQ Phototrackr and Trails, and their integration with Google Maps, Google Earth and Flickr. Examples of using the technology in research applications and for display of walking routes and points of interest were also included.

**Ann Roll**, a student in the LEEP Online Distance Education Program at the University of Illinois at Urbana-Champaign, is the first-place winner in the annual **Los Angeles Chapter of ASIS&T (LACASIST)** Margaret McKinley Memorial Student Scholarship Essay Competition. Ms. Roll’s interests lie in the areas of image retrieval and information visualization. Runner-up honors go to **Margaret A. Driscoll**, MLIS candidate at San Jose State University, and **Donna LaFollette**, recent graduatueon of the LEEP program at the University of Illinois.

**ANNUAL MEETING COVERAGE, FROM PAGE 8**

**Student Chapter-of-the-Year** – University of Washington Information School Student Chapter

**Chapter Member of the Year** – **Nicole Henning**, New England Chapter of ASIS&T

**Chapter Event-of-the-Year** – **Introduction to Content Management Systems Workshop**, Carolinas Chapter of ASIS&T, and **Mobile Mania**, New England Chapter of ASIS&T

**Chapter Innovation-of-the-Year** – **New England Chapter of ASIS&T** for the “TED Talks Film Festival

**Chapter Publication-of-the-Year** – Potomac Valley Chapter for the **PVC Blog**

**SIG-of-the-Year** – SIG/International Information Issues (SIG/III)

**SIG Member-of-the-Year** – **Ingrid Hsieh-Yee**, SIG/ED


**News from ASIS&T SIGs**

The newly established history fund awards initiated by SIG/HFIS and overseen by the ASIS&T History of Information Science and Technology Fund Advisory Board were awarded for the first time this year.

The 2009 Research Grant Award was given to **Charles Meadow** to research the historical background of the digital divide. Meadow is professor emeritus, Faculty of Information, University of Toronto, and the author of many books and articles relating to information science.


Both researchers are expected to present their papers at the 2010 ASIS&T Annual Meeting.

**Denise A. D. Bedford** has been appointed the Goodyear Professor in the Information Architecture and Knowledge Management (IAKM) program, School of Library and Information Science at Kent State University. The professorship was made possible by the generous support of the Goodyear Tire and Rubber Company. She had been senior information officer at the World Bank since 1997.

**News about ASIS&T Members**


**SIG Member-of-the-Year** – **Ingrid Hsieh-Yee**, SIG/ED


**News about ASIS&T Members**

**Denise A. D. Bedford** has been appointed the Goodyear Professor in the Information Architecture and Knowledge Management (IAKM) program, School of Library and Information Science at Kent State University. The professorship was made possible by the generous support of the Goodyear Tire and Rubber Company. She had been senior information officer at the World Bank since 1997.

**Student Chapter-of-the-Year** – University of Washington Information School Student Chapter

**Chapter Member of the Year** – **Nicole Henning**, New England Chapter of ASIS&T

**Chapter Event-of-the-Year** – **Introduction to Content Management Systems Workshop**, Carolinas Chapter of ASIS&T, and **Mobile Mania**, New England Chapter of ASIS&T

**Chapter Innovation-of-the-Year** – **New England Chapter of ASIS&T** for the “TED Talks Film Festival

**Chapter Publication-of-the-Year** – Potomac Valley Chapter for the **PVC Blog**

**SIG-of-the-Year** – SIG/International Information Issues (SIG/III)

**SIG Member-of-the-Year** – **Ingrid Hsieh-Yee**, SIG/ED


**News about ASIS&T Members**

**Denise A. D. Bedford** has been appointed the Goodyear Professor in the Information Architecture and Knowledge Management (IAKM) program, School of Library and Information Science at Kent State University. The professorship was made possible by the generous support of the Goodyear Tire and Rubber Company. She had been senior information officer at the World Bank since 1997.
From the first White Paper circulated in 2007 to the current time, the key question for the ASIS&T Information Professionals Task Force has been, “How can we identify current information education programs so that the stakeholder communities will have information relevant to their needs?” Answers to the question must come from all the groups directly affected by those education programs, including those who design and manage information programs, prospective students who seek out programs to fit their needs, prospective employers who are looking for graduates with the right portfolio of knowledge, skills and understandings to do the kinds of work needed in their areas of the information world and professional societies that want to be certain that they are providing professionals with the kinds of support needed in their careers. To-date most of the planning and design of programs has been under the purview of the faculty of academic institutions. While this structure assures that the programs meet the academic requirements of their institutions, they may not have had the level of input from other stakeholders that might make the programs more directly useful to students, employers and professional societies. An aim of the task force is to broaden the opportunity for stakeholder input in order to assure that information programs meet the needs of those who rely on them to support their career objectives and meet the needs of the information organizations they will serve.

Since 2007 the task force has

- commissioned a study to provide information on the number, location and particular thrust of existing graduate programs in information.
- identified more than 900 graduate degree granting programs in information
- identified 60 key information organizations representing stakeholders and invited them to join our efforts
- convened a meeting of stakeholders at the Brookings Institute on September 9, 2008, to discuss ways in which information related to the existence of information programs could be made available to stakeholders.

Documentation on each of the above activities is available at www.asis.org/news.html.

In February 2009 the Task Force, in response to comments and suggestions by stakeholders developed a proposed solution to respond to the expressed need to provide information on available graduate information programs. An identification database is proposed that would identify programs thus making it easier for prospective students, employers and society at large to understand what is and what is not included under the umbrella phrase information professionals. Program identification would vouch for content but would make no statement about quality. The database will provide information on the content of programs, their geographic location and other information pertinent to decision making. The database will provide those responsible for academic programs a way of seeing what is out there and how one’s own program relates, or not, to other programs in the region, in terms of disciplinary content, research focus and other relevant data elements. It will also provide professional associations the opportunity to look at the data and identify those areas of the information professions that are not represented in their membership, which can serve as a useful recruiting tool and a means of providing support to practicing professionals. The database is a way to describe the information professions based on the content of existing educational programs. A factor analysis can serve to provide a benchmark of the content of existing programs. The database will be kept up to date and will be a means of
showcasing the evolving nature of the field as demonstrated by its academic programs.

Five task forces will work on specific aspects of database development:

1. A database framework task force will conduct a factor analysis of a representative sample of the 900+ information programs identified to-date to identify types of programs and develop a framework for describing them.

2. A database content task force will develop a list of specifications for information to be included about each academic program and will define the information items required to describe programs.

3. A database specification task force will develop specifications for a program database to include the information specified by the database content task force. Particular attention will be paid to building a front end that will support social networking activities.

4. A promotional task force will develop specifications for a promotional campaign to inform stakeholders of the database, how it will be of benefit, how it is used and how it can be accessed.

5. An administrative mechanisms task force will investigate the administrative mechanisms that will be required to put the identification program and database management in operation.

At each step of the process, stakeholders will be invited to participate in the work of the task forces and will be informed regularly of progress.

The task force continues to bring professional issues to information organizations in general and to ASIS&T in particular. Two task force sessions were held at the Annual Meeting in Vancouver. The first of the sessions was Stakeholder’s Views on Information Education. It included presentations by Rachel Elkington, a recent iSchool graduate now working in an interaction design firm; Cynthia Fugate, associate director for research and instruction at the University of Washington Libraries; Deanna Morrow Hall, president of Corporate Information Resources, Inc., and Mark Greene, director of the American Heritage Center at the University of Wyoming, all representing professional societies. This presentation provided the opportunity for stakeholders to comment on the extent to which information about information programs has been available to them and how it might be more useful to their decision-making processes. It also provided a venue for educators to interact directly with stakeholders to discuss the above issues.

A second task force panel at the Vancouver meeting, System of Professions: Testing the Boundaries, addressed the information professions from a theoretical perspective. It included three speakers, Marcia Bates, University of California, Los Angeles; Prudence Dalrymple, Drexel University; and Cassidy Sugimoto, a doctoral student at the University of North Carolina, Chapel Hill. It was moderated by José-Marie Griffiths, also from the University of North Carolina, Chapel Hill. Drawing upon their published research and professional experience, the speakers addressed various aspects of the information professions: the structure(s) of the discipline and the profession(s), their evolving relationships to one another and their intellectual heritage and future.

The work of the task force continues. The proposal described above is being discussed with potential funding agencies, and we hope to begin some version of the work within the next six months. As task force chair, I will be leading a People-to-People delegation to China to take this topic international, talking with a spectrum of information-related education programs in that country. Groups represented at our 2009 meeting have been invited to participate in the delegation, and we are seeking funding so that task force members can participate as well.

The Information Professionals Task Force will continue to provide opportunities for dialogue among the various stakeholder groups with the objectives of making the academic programs in information as useful as possible to all involved and of making the path to an information professional career more apparent to all. The task force welcomes your thoughts and your participation.
The humanities and social sciences are concerned with the human experience. Sciences, too, deal with actions, processes and interactions. Information systems, therefore, are concerned with events, but can operate only on objects (bits, books, “documents”) – and events are not objects. Suzanne Briet wrote that “a document is evidence in support of a fact,” but facts (like data) have no meaning absent a narrative explanation. The three papers in this section explore the role of events and facts in information organization and retrieval and are based on their authors’ presentations at the 2009 ASIS&T Annual Meeting in Vancouver in a panel with the same title sponsored by the Special Interest Group/History and Foundations of Information Science (SIG/HFIS).

In “From Facts to Judgments: Theorizing History for Information Science,” Ryan Shaw, a doctoral candidate at the University of California, Berkeley, discusses the past as idealized images of people, places, events and ideas. Greatly expanded access to historical information through digitization has led to projects to extract facts from such resources in order to present history succinctly in databases. Shaw discusses the limitations of approaches that lift facts from their narrative context in the historical accounts. He advocates systems that enable us to see and retrieve historical events as bundles or colligations of narratives.

Thomas Dousa, in “Facts and Frameworks in Paul Otlet’s and Julius Otto Kaiser’s Theories of Knowledge Organization,” traces the origins of the idea that information units – or facts – can be extracted from documents and (re)organized within the frameworks of knowledge organization systems (KOSs). Otlet and Kaiser, who were both pioneers in knowledge organization in the late 19th and early 20th centuries, held nearly identical views about the analysis of documents into aggregates of facts, but key differences in their methodological and ideological outlooks resulted in vastly divergent narratives of knowledge organization and starkly different KOSs. Otlet developed a universal KOS: the UDC; Kaiser’s approach was particularist, creating different narratives for specific communities – a tension that is all too familiar to contemporary practitioners. Dousa is a doctoral student at the University of Illinois, Champaign-Urbana.

Finally, Michael Buckland and Michele Ramos in “Events as a Structuring Device in Biographical Mark-up and Metadata” report on the rationale for using events to structure biographical data for markup. Events are seen as arbitrarily defined actions suitably framed by the four facets of what, where, when and who. The paper summarizes the problems and solutions for each of these categories.
There is increasing interest in representing the past as a database of historical facts. Drawing upon the increasing availability of digitized historical texts and advances in text mining and semi-structured databases, these projects seem set to fulfill Paul Otlet’s dream of extracting the factual content from texts and making it available to answer queries about the past.

For example, the academic project DBpedia [1] aims to extract facts from Wikipedia infoboxes – the sections found on certain categories of Wikipedia articles that present basic facts using a standardized template. For example, the “Historical Event” infobox template includes fields for the event’s preferred name, alternate names, date, location, participants and result, as well as a representative image. On a Wikipedia article such as “French Revolution,” these values are presented as an HTML table. After parsing and processing by the DBpedia project’s algorithms, these values are transformed into a standardized data model for representing subject-predicate-object expressions (for example, “French_Revolution: location: France”).

Freebase is a commercial service that similarly parses Wikipedia infoboxes into structured data, but which has the more ambitious goal of integrating this data with all other available public domain data and furthermore providing interfaces for editing and adding to it.

Projects like DBpedia and Freebase mine historical facts primarily from the riches of Wikipedia, hoping that the collaborative effort there will trickle into their own projects. Other projects aim to mine the web at large for historical knowledge. Bruce Robertson calls the web “a historian’s fantasy” and envisions sophisticated tools for organizing and querying not only encyclopedia articles but also digitized archives, scholarly editions, journal articles and blog posts [2]. Pursuing a similar vision, digital historian Dan Cohen and programmer Simon Kornblith have developed a system called H-Bot [3] that parses Google search results to answer historical questions. And Google itself has begun incorporating timeline results – historical facts mined from the web – into its search results, as a search for a historical personage or event will usually show. All of these systems might be considered descendants of IBM’s “Professor RAMAC” computer program, which at the 1958 World’s Fair impressed audiences with its ability to answer historical questions using its “stack of 50 fast spinning disks” on which were stored “the principal historical events of the world from the birth of Christ to the launching of Sputnik I” [4].

H-Bot claims to demonstrate the “automatic discovery of historical knowledge” by using Google search results to answer simple factual questions such as “When was George Washington born?” or “Who was Lao-Tse?” The program uses simple sentence parsing techniques to transform questions into keyword searches that are passed to Google. It then uses statistical techniques to extract frequently repeated information such as dates and names from the returned web pages. (In practice, however, H-Bot usually draws its answers from a handful of online reference sources such as Wikipedia or Wordnet.) With its focus on simple factoids, its creators admit that H-Bot “offers an impoverished view of the past.” But they divert attention from this critique by implicitly equating such fact-finding with historical knowledge and by promising greater things to come (a favorite technique of AI researchers for decades). Most tellingly, however, they argue that the profusion of historical evidence surviving from the recent past poses a problem for historians, who must adopt the same techniques for managing information overload that scientists use.
This rhetoric should be familiar to information scientists. From Paul Otlet to Vannevar Bush to the National Science Foundation (NSF)-funded projects of today, much research in information management and retrieval has focused on the needs of scientists and specifically on how to help scientists avoid reading. The problem as it has been framed by this research is that scientists must stay current with ever more scientific literature in a finite amount of time. This dilemma has led to a focus on text summarization, filtering of irrelevant information and extraction of key facts from explanatory narrative. With the advent of large-scale corpuses of digitized texts, these techniques are now being proffered for the humanities as well. The call for proposals from a recent Digging into Data program funded by the National Endowment for the Humanities (NEH), NSF and others to develop data-mining tools and techniques for humanist scholars exemplifies how the problem of too much text is being framed (emphasis added by this author): “Now that scholars have access to huge repositories of digitized data – far more than they could read in a lifetime – what does that mean for research?”

Databases of historical facts are in part a response to this perceived problem: when there are too many histories to be read, boil them down to bare facts that can be subjected to powerful selection and retrieval mechanisms.

Such approaches are better suited to the sciences, given that scientists are assumed to be engaged in a cumulative research effort in which later researchers build upon the work of earlier researchers. This ideal of cumulative research effort requires that the complexity of earlier work be distilled down to reusable conclusions or facts. Bruno Latour in his Science in Action [5, 1-17] famously describes this process of fact production as “black-boxing.” A black box is a metaphor for a mechanical or computational component that is used to fulfill a functional requirement without knowledge of its internal implementation. While it may be possible to know how the black box works, all that is relevant to its users is that it produces expected outputs from given inputs. While scientists can in principle go back and recreate the experiments of their predecessors, efficient cumulative research requires that most of the time they simply trust that the facts they inherit work as advertised.

Yet as Louis Mink [6] points out, work in history does not produce “detachable” conclusions of this kind. It is rare for historians to simply accept an earlier account of some historical subject. Re-examination of primary evidence is the rule. Mink argues that this practice is not common because historiography is less developed than the sciences, but instead is due to a difference in the nature of the conclusions that historians produce. Rather than simply producing facts about the past, the historian aims to produce what Mink calls “synoptic judgments” of some complex of actions and events in the past.

A synoptic judgment is an interpretive act in which one moves from seeing that a series of things happened (the facts about the past) to seeing those happenings as a synthetic whole. Once she has reached such a judgment herself after immersion in the historical evidence, the historian’s task is to lead others through the interpretive process via the medium of a written text. Through the techniques of narrative representation, the author of a historical text aims to show past actions and events as a coherent whole when seen from a certain perspective. The exhibition of this whole as represented by the thick description of the historian’s narrative is the conclusion and as such cannot be detached from that narrative. In other words, the historian’s conclusions inhere in the structure and organization of her narrative. Even when the historian summarizes her narrative in separate statements, Mink argues, these statements are not detached conclusions but simply reminders to the reader of how the historian has ordered and organized her true conclusion, the narrative itself.

Databases of historical facts purport to help us answer questions about the past, and in a narrow sense they do that. But few of these systems take us further than the initial “Hey, neat!” reaction inspired by Professor RAMAC. The problem is not that the facts are wrong – Cohen and Rosenzweig [3] show that, on the contrary, they can be quite accurate by most standards – or that they are incomplete (though they certainly are). Nor is insufficiently advanced technology to blame – even if we were able to perfectly extract historical facts from texts, disambiguating every name and indexing each fact in the absolute grid of time and space, we would still face this problem. The problem is that systems like this are grounded in an impoverished conception of how we represent the historical past, a conception that focuses on atomic facts rather than synoptic judgments.
The problem is an old one. It can seem obvious that what we need to understand the past are facts about the past, and that a perfect history is thus one that identifies and enumerates “everything that happened” in terms of such facts. Yet upon careful consideration these notions are quite problematic. Philosophers have often hypothesized the idea of a complete historical database in order to demonstrate what the problems are. Arthur Danto imagines an “Ideal Chronicle” with descriptions of “absolutely everything that happened,” in the order it happened, thus providing the “whole map of the Past” [7, 148-181]. Danto argues that even such a complete database of the past would not obviate the need for historiography, since the role of historians is not simply to recount factual data about the past, but to represent the significance of episodes in the past from the perspectives of the present. These perspectives (and thus our criteria for significance) are constantly changing. It is this kind of change, not simply the discovery or refutation of historical evidence (addition or deletion of facts from the database) that results in new historiographical conclusions. Or, as Mink puts it, even if we could “sit before a screen and directly review the past in its minutest details,” we would still need some imaginative representation of the past to help us make sense of it all in light of our current historical situation, and it is the role of history to develop such imaginative representations and not simply gather the detailed facts.

Acts of synoptic judgment produce imaginative representations that are articulated as historical narratives. Once historians have developed narratives that relate sets of facts under some synthesizing ideas, they usually label these narratives with phrases like “The Renaissance” or “The French Revolution.” The philosopher of history W. H. Walsh calls this process colligation, appropriating the philosopher of science William Whewell’s term for “the binding together or connection of a number of isolated facts by a suitable general conception or hypothesis” [8, 59-64]. Walsh supports a hermeneutic conception of historical method in which the goal of historians is to imaginatively and empathetically reconstruct the experiences and thoughts of people in the past. Accordingly, Walsh’s notion of colligation initially depended upon happenings being intrinsically connected by virtue of being intentions or consequences of some past actor’s plans. A “suitable general conception” for binding together facts was one that illuminated those facts as being part of a (conscious or unconscious) policy guiding the behavior of people in the past. Later Walsh expanded his notion of colligation to include any case in which some set of events as is interpreted as a connected process or development, whether or not such a policy could be discerned [9].

Even though concepts are of primary interest in library and information studies, colligatory concepts have been mostly overlooked. Even the most sophisticated theoretical discussions of concepts in the literature tend to equate concepts with classes or categories. For example in his recent survey of concept theories Hjørland [10, p. 1522] asserts that “[c]oncepts are dynamically constructed and collectively negotiated meanings that classify the world according to interests and theories” (emphasis added). This preoccupation with classification is perhaps understandable in light of the aforementioned focus on scientific domains. The sciences seek to abstract away from unique individuals to generalized classes that can be related by laws. While historians do generalize, they also – arguably primarily – seek to assemble descriptions of unique past events into connected and coherent but no less unique representations. Concepts like “The Renaissance” colligate rather than classify.

The most fully developed theory of colligation to date has been developed by Frank Ankersmit, who in his Narrative Logic: A Semantic Analysis of Historian’s Language [11] seeks to explain how colligatory concepts – which he calls “narrative substances” – are constructed from sets of statements expressing facts. A narrative substance is a point of view from which to regard the past, articulated by means of a specific historical narrative. Ankersmit contends that each individual historiographical narrative constructs a narrative substance so that, for example, there are as many “Renaissances” as there are narratives on the subject, since each narrative articulates a specific point of view. So when we speak generally about “The Renaissance,” we are really talking about a whole family or type of narrative substances that have been given the same name.

Furthermore, Ankersmit claims that when we define such types, we do so extensionally rather than intensionally. An intensional definition of a type is one that defines some necessary and sufficient conditions for belonging to the type. For example, one might define a mug intensionally as “a type of cup
made of glass or ceramic and having a handle large enough to accommodate a whole hand.” An extensional definition of a type, on the other hand, enumerates the members of a set of individuals considered to be instances of that type. An extensional definition of *mug* would collect all the world’s individual coffee mugs and beer steins and so on and thereby declare “these are mugs.”

Ankersmit argues that one can define types of narrative substances extensionally by clustering narrative substances that contain overlapping sets of statements. He proposes a thought experiment in which a giant matrix is constructed. Along one axis of the matrix are aligned all the declarative statements made about the past that have actually appeared in some text or another. Along the other axis are aligned all the narrative substances constructed by means of those statements. Each cell in the matrix is then filled with a “0” or a “1” indicating whether or not the corresponding statement was used to help construct the corresponding narrative substance. Given such a matrix, we could then try to identify types of narrative substances by grouping together narrative substances with similar patterns of 0s and 1s, in much the same way that we might identify types of drinking vessels by looking for similar shapes or handles or materials. Ankersmit posits that we will observe that “certain classificatory patterns automatically appear.” These clusters in “narrative space” reflect the fact that historians write in response to other historians and construct their narrative substances by distinguishing them from those that came before (which implies a significant degree of overlap).

As Ankersmit points out, such an extensional procedure for identifying types can never be precise. Depending on how we interpret similar patterns there will be many possible groupings into types. Moreover, for any given interpretation of similarity, there will always be boundary cases that could belong to more than one cluster. At best, extensional typification can identify regularities in how we have chosen to conceptualize reality, but it cannot tell us anything about reality itself. In other words, looking at written history this way tells us something not about the reality of the past, but about the contours of the concepts developed by historians over time. “The Renaissance,” “The Cold War,” “The French Revolution” and “9/11” are not objectively existing entities in the past, but are names of types of stories we tell to understand the past.

Finally, Ankersmit argues that an extensional procedure like this is the only way to identify types of narrative substances. The alternative would be to intensionally identify types in terms of logical definitions based on attributes of the things being classified, the way we define the type *mammal* as “warm-blooded,” “vertebrate” and “having hair or fur.” But this intensional identification is precisely what we cannot do for narrative substances. There is no logical definition, no core set of properties both necessary and sufficient for making a particular narrative a narrative about “The French Revolution.” While it’s easy to identify statements that would not appear in any narrative of the French Revolution – for example that the storming of the Bastille occurred in 1967 in Tokyo, Japan – we cannot identify statements that *must* appear in such stories or by virtue of which we *must* consider a given narrative to be a narrative about the French Revolution. Given all the narratives that have ever been written about the French Revolution, we may not be able to identify a single statement that appears in every one. Thus we cannot and do not identify types of narrative substances intensionally. Decisions about what “The Cold War” is can only be justified pragmatically, not logically.

With the large-scale digitization of books it may become possible to investigate Ankersmit’s theory by analyzing the full texts of historical narratives. Before undertaking such a project we must address some methodological problems. First is the issue of how to identify statements about the past. Ankersmit’s matrix involves statements (also known as *propositions*) about the past, not the sentences that express these statements. (Ankersmit contends (p. 19) that, “states of affairs in the past can be unambiguously described by means of constative statements.”) We cannot conflate statements with the sentences that appear in historical texts, as it is unlikely that any two texts will contain precisely the same sentence. So we must find a way to move from the sentences that appear in texts to the propositions those sentences express, a problem that has occupied linguists and philosophers of language since Bertrand Russell. Fortunately this is an area where information extraction technology might show its worth, as such technology is precisely concerned with transforming sentences into logical propositions. Despite well-known problems with propositional theories of language [12, 70-74] and the fact that information extraction technologies are plagued by errors, their output still might be usable for exploring Ankersmit’s theory.
The second problem, as Ankersmit points out, is that a given historical text constructs multiple narrative substances, and it is difficult to determine exactly which statements are being used to construct which narrative substances. Indeed, Ankersmit argues (p.104) that in order to identify narrative substances reliably we must, “compare historiographical topics studied and discussed by generations of historians.” Fortunately the catalogers who maintain the Library of Congress Subject Headings (LCSH) [13] have done this for us, creating subject headings for historiographical topics and assigning them to historical texts. Using the LCSH is not ideal for investigating narrative substances, however. Catalogers usually do not identify more than a couple of the narrative substances constructed in a given text. Since they are concerned with characterizing whole books, they will not identify historical narratives in books that are not primarily works of historiography. And since the goal is to collocate texts, there is no effort to distinguish differences among the narrative substances constructed by different texts. In essence, what catalogers have done is group various narrative substances into types a priori.

Notwithstanding these problems, it is plausible that one could use a historiographical subject heading to obtain a set of texts within which authors have constructed comparable narrative substances. Ankersmit suggests that we can refer to narrative substances with terminology such as “Renaissance,” “Renaissance2,” “Renaissance3,” etc. where the subscript n indicates that we are referring to the specific representation of the Renaissance constructed in the historical text n. Likewise, we could posit that a set of N texts found under the LCSH for the “Renaissance” constitute a set of narrative substances “Renaissance,” “Renaissance2,” “Renaissance3,”... “Renaissance,n.” We could then compare the propositions made in these texts to each other to build a model of the extension of the type “Renaissance.” This model could provide a basis for highlighting differences among the individual narratives constructed by the different texts. Though such a model would enable us to examine the structure of types already identified by the LCSH, it could not reveal new types not yet identified there. But we have to start somewhere.

Whether or not initial attempts to automatically discern and model them prove successful, it is time information science paid closer attention to colligatory concepts. In the digital environment, traditional components of the scholarly apparatus such as term lists, classification and categorization schemes, and thesauri are evolving into generalized semantic tools for enumerating and disambiguating concepts and mapping the relations among them [14]. Though in the past such tools have mainly been used for indexing and retrieval, in an era of full-text search I believe we will see other applications move to the fore. Specifically, semantic tools that map a given conceptual domain can be integrated into reading and writing environments to help users contextualize some fragment of interest. Given such applications, “fuzzy” concepts reflecting particular interpretive stances are at least as important as traditional categorical concepts. Ankersmit and his predecessors’ theory of colligation provides grounds for investigation of such concepts. Done properly, such an investigation may lead to new tools (or new ways of using existing tools) for research that avoids reducing conceptions of the past to bare facts.

ACKNOWLEDGEMENTS – Support from the Institute of Museum and Library Services through award LG-06-06-0037-06 “Bringing Lives to Light: Biography in Context” and from the Advancing Knowledge grant PK-50027-07 “Context and Relationships: Ireland and Irish Studies,” jointly funded by the National Endowment for the Humanities and the Institute of Museum and Library Services, is gratefully acknowledged.
Resources Mentioned in the Article

Website links
DBpedia: http://dbpedia.org
Digging into Data: www.diggingintodata.org/
Freebase: http://www.freebase.com
Google: www.google.com
H-Bot: http://chnm.gmu.edu/tools/h-bot
Wikipedia: www.wikipedia.org
Wordnet: http://wordnet.princeton.edu

Other Resources
Facts and Frameworks in Paul Otlet’s and Julius Otto Kaiser’s Theories of Knowledge Organization

by Thomas M. Dousa

The late 19th and early 20th centuries witnessed numerous developments in the domain of classificatory and indexing activities known today as knowledge organization (KO). Among the most striking of these was the emergence of the idea that documents could be decomposed not only into smaller bibliographical units (as, for example, a periodical into articles or a book into chapters), but also into yet smaller information units (such as, for example, the concepts or facts discussed in discrete passages within a text) and that, once identified, these information units could be reconfigured in new arrangements that would facilitate their retrieval [1, p. 223; 2, pp. 221–222]. This idea, which I term information analysis, would have a long and influential career in information science (IS) and continues to influence IS theory and practice to this day [see sidebar].

The notion of information analysis may be traced back to two pioneers of KO, Paul Otlet (1868–1944) and Julius Otto Kaiser (1868–1927). The Belgian Otlet occupies a prominent place in the history of IS. A lawyer by training and a bibliographer by vocation, he made a number of fundamental contributions, both theoretical and practical, to the development of IS. Originator of the European documentalist movement, initiator or charter member of a number of organizations devoted to international cooperation in bibliography and (co-)creator of the Universal Decimal Classification (UDC), he developed a comprehensive vision of how information was to be organized and mobilized that prefigured, in many respects, the hypertextual systems that are in vogue today [3].

Kaiser, by contrast, holds a somewhat lower profile in the annals of IS. Born in Germany but spending most of his life in Great Britain and the United States, he served as an indexer and bibliographer for various commercial institutions, industrial enterprises and technical libraries. Today he is best known for his method of “systematic indexing,” which is considered

The Enduring Legacy of Information Analysis in IS

The idea that documents can be decomposed into smaller information units to which direct access might be provided has been historically influential in IS theory and practice, as the following three examples indicate:

- In the first half of the 20th century information analysis served as a theoretical fracture point in the professional cleavage between general librarianship and documentation/special librarianship. Whereas general librarians focused on providing subject access to books by means of subject headings that characterized the contents of the document as a whole, documentalists and special librarians sought to provide intellectual access to specific information with in documents through detailed indexing of all pertinent information units that they contained. Information analysis thus became a professional marker of documentation and special librarianship.

- In the 1950s and 1960s, information retrieval (IR) theorists drew a distinction between “document retrieval systems” and “fact retrieval systems.” The former, were intended to retrieve, in response to a user’s query, all documents that might contain information pertinent to answering that query, while the latter were to lead the user directly to specific pieces of information – facts – embedded within the documents being searched that would answer his or her question. The idea of information analysis clearly provided the theoretical impetus for fact retrieval (aka question-answering) systems.

- From the 1990s through today, the extraction of information units from digital documents has become a focal point of research within IS. XML-based markup languages are being used to identify and isolate chunks of information within digital documents, and systems are being designed to retrieve and collocate these pieces of information. Text-mining techniques are being developed to identify, retrieve and collocate information units from digital documents with unstructured text. The idea of information analysis is the basis of such work.
to be a forerunner of facet analysis [4]. Although Otlet and Kaiser worked independently of one another, at the turn of the 20th century, each developed a well-articulated account of information analysis that deeply informed his theory of KO.

In this article, I sketch Otlet’s and Kaiser’s ideas about information analysis and compare the types of knowledge organization systems (KOSs) that they constructed on the basis of these ideas. As we shall see, Otlet and Kaiser held very similar views about the possibility – and desirability – of disaggregating documents into information units and organizing the latter into indexed information files. Both men also agreed on the technological means to implement their information-analytic approach. Despite these points of convergence Otlet and Kaiser envisioned vastly different kinds of contextual frameworks for organizing the informational units that they sought to disengage from their documentary trappings and, as a consequence, they developed contrasting theories of what the scope and inner articulation of a KOS should be. The striking differences between the types of KOSs that Otlet and Kaiser constructed on the basis of their views of information analysis can be explained, in large measure, by reference to differences in their personal philosophies as well as to the different professional cultures within which they operated.

From Knowledge to Documents to Facts: Otlet and Kaiser on Information Analysis

At the root of Otlet’s and Kaiser’s notions of information analysis lay their views of the relationship of documents to knowledge. According to Otlet, knowledge encompasses “[e]verything which we know about objects in the external world or from our own thinking, physical objects (natural or artificial), non-physical objects (laws, thoughts, sentiments)” and its primary elements are “facts” and “ideas” [5, p. 73]. Knowledge, in his view, is recorded in the form of documents, wherein authors embed facts and ideas into conceptual structures that reflect their personal understanding of the phenomena that they are discussing or representing. Otlet held that “every document is an exposition of data, facts, and ideas … more or less well ordered, clearly formulated, strongly stylized” [6, p. 97]. According to Otlet, users of documents seek “recourse to documents in order to extract facts and information from them for the acquisition of knowledge, for study or for scientific research” [7, p. 105]; thus, if the new field of documentation was to support this “documentary method” of research, it would have to offer resources for providing access to the individual facts and ideas ensconced within documents.

When we turn to Kaiser, we see that a similar story prevails. In his view, documents – or, to use his preferred term, pieces of literature – constitute “descriptive record[s],” set down in language, of what individual human beings “observe” in and “reason out” about the world [8, § 52–53]. Kaiser held that “[t]he subjects of our observing and reasoning are things in general, real or imaginary, and the conditions attaching to them” [8, § 52]: the things he called processes and their conditions, processes. If knowledge can be reduced to knowledge of concretes and processes, and literature represents an encoding of knowledge into written language, it stands to reason that it must be possible to decompose a document into statements about concretes and processes. This reasoning was precisely the view of Kaiser, who put it thus: “from the standpoint of knowledge literature is confined to the description of concretes and of the conditions attaching to them; … for our purposes literature may be analyzed into terms of concretes and terms of processes” [8, § 298]. Because, Kaiser argued, men of affairs are interested in the information contained in documents rather than in the documents themselves, “we must try to dissociate information from literature” and so render it more accessible to its potential users [8, § 83].

In addition to agreeing on the possibility and desirability of information analysis, Otlet and Kaiser held similar views about what it would achieve. For Otlet, the goal of the documentalist – who, in his view, should also be a subject specialist – is to identify information units within the document and create individual records for each one. Once individual information units have been separated from their original bibliographic contexts, they can “be set out in a quite analytical way” into “encyclopedic repertories” – index files in which units of information will be organized in such a way as “to link together materials and elements scattered in all relevant publications” [5, pp. 84, 83]. According to Otlet, such repertories will comprise “inventories of facts, catalogues of ideas, and the nomenclature of systems and of theories” [5, p. 83]. Just as bibliographic repertories provide access
to (information about) documents, so will encyclopedic repertories provide access to information itself.

Kaiser’s view on the matter was virtually identical. In his view, the role of the indexer in a special library was to analyze documents, decompose them into individual informational units, record those units and organize them “on a uniform plan applicable to all the information incorporated in the index” [8, § 295]. In this way, he wrote, “[w]e take our literature to pieces and rearrange the pieces systematically so as to answer best our object in view” [8, § 16]. The resultant index, so Kaiser averred, will “give an analytic statement of the information, for it has been cut up into pieces, specific facts or opinions, and rearranged in more suitable form” for use by the clientele for whom the index has been constructed [8, § 297]. Kaiser’s systematic index, no less than Otlet’s encyclopedic repertories, was designed to provide immediate access to information.

The Card Index System: Technological Substructure for Otlet’s and Kaiser’s Vision of Information Analysis

As we have seen, Otlet and Kaiser conceptualized documents as recorded expressions of knowledge consisting of constellations of informational units comprising facts, ideas and opinions. That is, both men advocated extracting these informational units from documents and reorganizing them into index files in order to make them more accessible to researchers. Their agreement on the theoretical fundamentals of information analysis for KO was reinforced by a shared commitment to the cutting-edge information technology of their time: the card system.

In today’s digitally inflected culture card indexes housed in cabinets are often considered to be an antiquated and obsolete mode of storing information. In the last decade of the 19th century and the opening decades of the 20th century, however, information systems based on “[c]ards of a uniform size, on which standardized data were transcribed, housed physically in card drawers and related furniture and organized conceptually by classification of schemes of various kinds … epitomized a new ‘modernist’ technology” that was considered to have enormous practical advantages over the previously regnant system of recording information in bound ledgers [9, p. 12; 10]. Finding increasing use not only in libraries (where card files were first employed) but also in business offices and government bureau, card index systems were especially prized for the flexibility in filing that they afforded: not only could cards bearing superannuated information be easily removed and ones bearing new information be added, but files could be easily rearranged if needed.

Otlet and Kaiser were well aware of these advantages and enthusiastically advocated for the use of the card index model [6, p. 384; 11, §§ 71–72].

The underlying reason for the flexibility of the card index system was its segmentation into a number of physically discrete, modular records – cards. Because, in such a system, “[e]ach card is a unit record representing an item of information” [10, p. 405], cards are an ideal tool for registering the results of information analysis. This technology was, indeed, precisely what Otlet and Kaiser envisioned for the KO systems that they designed. Both men held that, in an information index file, each individual card – or, in some documentary contexts, each sheet of loose-leaf paper (Otlet) – should serve as the bearer of a single unit of information extracted from a document. As if to emphasize the one-to-one correspondence between card and information unit, Otlet termed this methodological tenet the “monographic principle” [3, p. 238].

Once the information to be indexed had been entered upon cards in accordance to the monographic principle, it was necessary to organize the index file as a whole. Both Otlet and Kaiser envisioned that each card would be indexed by subject and that different cards containing information about the same subject would be collocated within the index file. Special divisionary, or guide, cards, distinguished from the others by size, shape or color, would mark the place of any given subject entry within the card file and indicate any subdivisions of main entries, as well as give cross-references to related subjects [3, p. 242; 8, §§ 399–416]. Card index files thus had the resources to represent individual pieces of information (by means of subject-indexed unit cards) as well as to indicate the general structure of the card file within which the unit cards were gathered (by means of guide cards) and point out connections between subjects (also by means of guide cards).

In Otlet’s and Kaiser’s eyes, then, the card system was an ideal mechanism for gathering together information units gleaned from many different documentary sources, organizing them according to their intellectual content.
and guiding users to cards containing information on the particular subject of their concern. As such, it was a *sine qua non* for their vision of KO based on information analysis.

**Classified vs. Alphabetical Order: Otlet’s and Kaiser’s Divergent Views of KO Frameworks**

Whereas Otlet and Kaiser were in substantial agreement on both the desirability of information analysis and its technological implementation in the form of the card system, they parted company on the question of how index files were to be organized. Both men favored organizing information units by subject, but differed as to the type of KO framework that should govern file sequence: Otlet favored filing according to the classificatory order of the UDC, whereas Kaiser favored filing according to the alphabetical order of the terms used to denote subjects. It is instructive to examine why each thought the way he did on this important point.

To understand Kaiser’s preference for alphabetical order, it is necessary briefly to consider the main points of his indexing methodology, which may be summarized as follows. The indexer was to identify key terms within the documents he was processing, extract them and assign each term to one of fundamental categories – concretes (i.e., entities), countries and processes (i.e., the action that a thing does or undergoes) [8, § 73, 299–301]. Once terms had been assigned to their respective categories, they could be joined together in one of three basic combinations, which constituted subject “statements” about the information being indexed [8, § 302]:

- Concrete–Process
- Country–Process
- Concrete–Country–Process
  [var., Country–Concrete–Process]

Within the file, main entry terms would always denote concretes or countries, whereas terms denoting processes would always serve as subdivisions. The main entry terms were to be arranged in alphabetical order, with the same being done for subdivisions under each main entry [8, §§ 389–390]. Although the filing order was alphabetical, Kaiser provided what he called a “logical key” to the index [8, § 389]. He did so by stipulating that the indexer indicate on the guide card for each main entry term a list of the other main entry terms with which it stood in semantic relation: the latter included synonyms, broader terms, narrower terms and related terms [8, § 415, 423]. The “logical key” served as the syndetic structure of the index, indicating a web of conceptual relations otherwise unexpressed by the alphabetical structure of the index file.

Given that Kaiser acknowledged the utility of indicating semantic relations between index terms, why did he prefer alphabetical to classified order for filing? The answer lies in his view of language. Kaiser considered words in natural language to be imprecise expressions of the concepts that they are intended to convey. In addition, he held that there is little agreement among users of a language as to the precise definition of individual terms [8, § 60–61, 112]. Such semantic indeterminism, in his opinion, makes it difficult to determine precisely what the authors of documents mean by the words they are using, and any attempt by an indexer to substitute preferred terms for the author’s own words runs the risk of misinterpreting the meaning of the words in the original document. It was for this reason that he preferred using index terms extracted from the document itself [8, § 114].

Kaiser’s perception of the semantic lability of language also led him to distrust classified order because it presupposed universal agreement about the precise definition of words – something that he held to be impossible. The alphabetical approach circumvents this problem by organizing information files by the formal characteristics, rather than the meaning, of subject terms: it thus reduces the possibility of misinterpretation on the part of the indexer [8, §§ 114, 178]. Furthermore, it makes use of an organizing principle that forms part of the common knowledge of indexer and user alike, whereas a classed organization based on the meaning of its index terms would require special knowledge of the principles underlying that classification [8, § 131–132]. For Kaiser, then, alphabetical order represented the interpretatively safest and most user-friendly way of organizing information units by their subject terms.

Otlet, by contrast, was strongly opposed to organizing information units by the alphabetical order of their index terms. In his view, such a mode of organization “scatters the [subject] matter under rubrics that have been classed arbitrarily in the order of letters and not at all in the order of ideas” and so obscures the conceptual relationships between them [6, p. 380]. By
the same token, it removes individual terms from the intellectual contexts that help to determine their meaning, and this isolation makes it difficult for users of an alphabetical index file to “handle the complex expressions that one finds in the modern terminology of discipline[s] such as medicine, technology, the social sciences” [6, p. 381]. A further difficulty, claimed Otlet, is the tendency of individual alphabetical indexes to have their own particular sets of index terms: this “arbitrariness in the choice of words” gave them “a personal character” that hindered standardization and, hence, bibliographical cooperation across different information centers [6, p. 381]. Finally, because the subject terms in an alphabetically arranged system must be expressed in a certain language, such indexes would be, of necessity monolingual and so usable only to members of a single language community. This restriction, in Otlet’s view, would be an impediment to international cooperation in the organization of information [6, pp. 380, 381].

Otlet’s preferred framework for organizing information units was the classified approach, as embodied in his UDC. One reason lay in the UDC’s capacity for providing intellectual context for individual subject entries: unlike alphabetical indexes, it assured that “all related subjects are grouped together” and so was likely, in Otlet’s opinion, to be “preferred by scientific men” [12, p. 99]. A more fundamental reason, in Otlet’s view, was that the UDC was comprehensive and systematic in its subject coverage and eminently suitable for use in an international setting. Not only did the decimal classification’s general outline “embrace the universe of knowledge” but its use of a numerically-based decimal notation to designate classes and their subdivisions expressed in a transparent way “the location of each subject, no matter how specific it may be, in the whole corpus of knowledge” [13, pp. 31, 34, 36]. No less important, the numerical notation served to “translate ideas” into “universally understood signs,” namely numbers [13, p. 34]. This feature would permit users from different linguistic communities to use the classification despite the language barriers that might otherwise separate them. For Otlet, then, the UDC provided a comprehensive and systematized codification of knowledge whose notation allowed it to be used on an international scale, and such a system would be necessary if the results of information analysis were to be made available to the broadest range of users.

**Universalism vs. Localism: Otlet’s and Kaiser’s Differing Views on Framework Scope**

For Otlet, one of the major advantages of the UDC over alphabetical systems was that “every alphabetical filing scheme has, through the arbitrariness of the choice of words, a personal character, whereas the [UDC] has an impersonal and universal character” [6, p. 380]. His preference for impersonality and universality in KO must be viewed in light of the larger project of which the UDC was a key element. Otlet’s initial purpose in creating the UDC was to establish a KO framework for a bibliographic card catalog of universal scope, which he dubbed the Universal Bibliographic Repertory (RBU) and developed under the auspices of one of the institutions that he founded, the International Institute of Bibliography (IIB) [3, p. 238]. As he elaborated his idea of information analysis, he developed the idea of a “universal encyclopedia,” which was a series of interrelated, discipline-based card (or loose-leaf sheet) indexes synthesizing the totality of information relating to the different branches of knowledge [5, p. 83-84; 6, p. 409]. These comprehensive collections of facts and ideas were likewise to be indexed and organized in accordance with the UDC and their development was to be overseen by the Office of International Bibliography (OIB), the branch of the IIB responsible for administering the dossiers of information that Otlet was creating. Otlet’s intent, then, was that the UDC would be a classification of universal scope that would serve as an international standard for organizing information.

In stark opposition to Otlet’s insistence that an ideal KOS be impersonal and universal, Kaiser firmly held to the view that, ideally, KOSs should be constructed to meet the needs of the particular organizations for which they are being created. For example, with regard to the use of card indexes in business enterprises, he asserted that “[e]ach business, each office has its individual character and individual requirements, and its individual organization. Its system must do justice to this individual character [11, § 76]. Because different businesses have different information requirements, “each office must devise its system in accordance with its own requirements, and it should itself be the best judge of what these requirements are” [11, § 76]. The inevitable corollary to this was that “it is impossible to devise a system that could be applied universally” [11, § 76].
What applied to the business office applied to the library as well. Kaiser deprecated the use of universal classifications in special libraries and information bureau, claiming that prefabricated general schemas for organizing documents or information could not take into account the particular information needs of a given organization [8, §§ 246–247]. Furthermore, he rejected the idea that an index should aspire to comprehensive coverage of a universe of knowledge. Rather, he maintained, it should include index terms only for those subjects that actually pertain to the information needs of the organization for which it was designed [8, §§ 309, 311]. In the same vein he recommended that the relationships between main entry terms established in the synectic structure of an index should reflect the particular point of view of the organization using it [8, § 425]. Kaiser’s insistence on designing KOSs in accordance with the particular needs of a given organization stands in stark antithesis to Otlet’s belief that such a system should be universal in scope, standardized and sufficiently impersonal to be useful to a large number of users from a host of different backgrounds; this antithesis continues to be a topic of discussion about KO within IS to this day [14].

Conclusion: Worldview and Professional Culture as Determinants of Frameworks for Facts

We have seen that Otlet and Kaiser believed that documents could be disarticulated into information units (namely, statements of facts and ideas), that these units could be recorded on unit records in the form of cards and that they could be filed in card indexes serving as repositories of information directly accessible to users. However, they held almost diametrically opposed views about the appropriate KO framework for organizing information units: Otlet advocated the “grand narrative” of a universal classification system applicable across different contexts and encompassing the whole universe of human knowledge, while Kaiser supported the creation of numerous, local “micro-narratives” in the form of alphabetical indexes customized to the specific needs of individual institutions. It remains to consider why, despite their agreement on the fundamentals of information analysis, their ideas of what constitutes the ideal type of KOS were so divergent.

Two factors explain Otlet’s and Kaiser’s different visions of KO: their personal worldviews and the professional cultures that they inhabited. Strongly influenced by the Comtean and Spencerian versions of positivism that he had imbibed in his youth, Otlet was concerned with the progress of the sciences (in the widest sense of the term), believing that a holistic integration of human knowledge into a single, well-articulated system of sciences could form the intellectual basis for the universal amelioration of human life [15, pp. 20, 26–29, 354]. This integrative, universalist conception of human knowledge was of a piece with Otlet’s passionate devotion to the internationalist cause of fostering social unity on a planetary scale and his lifelong commitment to the building of institutions for international cooperation [15]. For him, information analysis provided a mechanism for identifying and isolating those facts and ideas that belonged to the general fund of human knowledge and organizing them in a universally codified fashion that would allow researchers to leverage this information more easily for the benefit of humankind as a whole: in this way, he hoped, the better organization of information produced by documentation would contribute to the progress of humanity.

In contradistinction to Otlet’s universalist perspective, Kaiser’s worldview was rooted in an ethos of individualist particularism. Holding to the conviction that “[o]ur individuality is our greatest asset” and that “[o]ne cannot standardize the human intellect,” he took the goal of information analysis to be to create indexes of information units that would provide persons working within a particular organization with access to just those facts and ideas that would be useful to them in their work for that organization [8, §§ 23, 57]. Whatever the intellectual sources of Kaiser’s fervent individualism may have been, his worldview was perfectly consonant with the world of commercial and technical libraries within which he developed and refined his method of systematic indexing, a world that placed a premium in the customization of classification and indexing to the specific needs of the enterprise for which the librarian was working [16, pp. 184].

Ultimately, the seemingly paradoxical differences between Otlet’s and Kaiser’s views of how KOSs based on information analysis should be elaborated must be explained by cultural differences, both personal and professional. Both men lived and worked during an epoch when cultural modernism was at its height [17]. It is unsurprising, then, that the approach
to information analysis that they developed is consonant with the modernist impulse towards the systematization, organization and rationalization of cultural processes. Yet if Otlet and Kaiser were both products of a modernist milieu, they also inhabited different worlds within that milieu. Motivated by a universalist vision of benefiting humankind as a whole, Otlet elaborated his KO theory within the world of international organizations and universal bibliography, while Kaiser developed his within the world of commercial and industrial libraries, where pragmatic concerns strictly tied to particular organizational needs were of primary importance. The role of culture in shaping Otlet’s and Kaiser’s differing views on KO frameworks for organizing facts derived from information analysis is well worth keeping in mind for proponents of information-analytic approaches to IS today.

AUTHOR’S NOTE – I thank the members of the Research Writing Group at the Graduate School of Library and Information Science, University of Illinois, Urbana-Champaign, for their helpful critique of an earlier draft of this article.

Resources Cited in the Article

Events as a Structuring Device in Biographical Mark-up and Metadata
by Michael Buckland and Michele Renee Ramos

There is little structure or best practice in the concise biographical texts found in biographical dictionaries and Who’s Who volumes. This paper is a progress report on an investigation of using events as a structuring device for mark-up and metadata structures in biographical texts as part of a project entitled Bringing Lives to Light: Biography in Context [1]. The idea is that anyone’s life can be usefully decomposed into events at any desired level of granularity and that each event could be described as a 4-tuple of the four facets what, when, where and who.

Purpose

The difference between seeing and understanding lies in knowing the context, and it should be emphasized that we approached this problem area from a particular perspective: helping readers to understand. This paper reports on one part of a series of studies of how learning can be facilitated by making it easier to find relatively trustworthy explanatory resources, suitable both for a text being read and for the reader. During 2004-2006 a project entitled “Support for the Learner: What, Where, When and Who” explored this area in general terms [2]. A four-facet “4W” approach was adopted – what, where, when and who – because each has distinctive characteristics leading to different genres of search aid and different display requirements.

Where involves a duality of place (a cultural construct) and space (a physical construct) and, for this, place name gazetteers and map displays are well-developed genres [3][4]. When similarly involves a duality of events and calendar time. Historical events are calibrated by calendar time, and calendar time is calibrated by events such as solar years and cesium radiation cycles. In practice, people tend to mark time by mentioning personal and historical events (for example, “after I graduated,” “during [the] Vietnam [war]”) more than by calendar dates, so an approach similar to place name gazetteers using named time period directories and timelines can be adopted [5]. What tends to be a residual category when other specialized concepts have been removed. Here thesauri, subject indexes, library classifications and other tools are used, and ways to express and display relationships and cross-references are well developed.

Who, however, emerged as a relatively underdeveloped area. The disambiguation of personal names – associating multiple names for the same person and distinguishing different persons with the same name – is a well-understood area with its own standards. Also, genealogists understand family relationships and how to represent them in family trees, but other kinds of interpersonal relationships have not received the same careful attention, although a variety of limited examples can be found. Further, there seemed to be a distinct lack of accepted standards or best practices for structuring the very concentrated, stylized biographical texts found in biographical dictionaries. Our impression of the situation was validated and detailed in a report issued by the Text Encoding Initiative [6]. Fortunately we have been able to explore these issues as part of a project entitled “Bringing Lives to Light: Biography in Context” [1].

If order is to be brought to disorder, appropriate organizing principles are needed. Time is one obvious principle for biographical arrangement, but, by itself, it provides little beyond a useful sequencing. What additional approaches would be more useful for organizing the details of people’s lives? Events shape our lives, and we engage in actions and activities, so combining actions and activities with other kinds of events seemed a promising analytical device for organizing the untidiness of peoples’ lives.
About Events

We use event in an everyday sense. The most basic characteristic of an event is that something happens. There is activity, some change; otherwise we would say it was a “non-event.” Change happens through time. An event may be very brief, seeming instantaneous, or it might be very prolonged, but all events occur in time and require at least some time to happen. Because an event is presumed to involve some change, events cannot be preserved, only described, represented or re-enacted. For this reason and also because people see and define events differently, we treat events as narrative constructs.

Events happen in some place. It could be everywhere (like the Big Bang), the location might be unknown or it could be an imaginary place, but an event that happens nowhere would be a non-event.

When considering events, one quickly recognizes a duality between state and action. For example, one could say that one spent the year 2008 studying (an activity), or one could say that one was a student during 2008, a status. These are alternative, comparable autobiographical statements that appear to be different but are equivalent for most purposes. (Strictly speaking, of course, being officially a student does not in itself guarantee that much studying is being done and, vice versa, one can study without officially being a student, although in a broader sense anyone who studies is a student.)

There is a corresponding duality with respect to defining what the event is. It could be the prolonged activity of studying, or it could be the transition to (or from) student status. That these alternative perspectives are different does not matter much because they are complementary, mirror images of the same life experience, and one can choose either approach to suit one’s need. We have preferred to adopt the activity as the event, rather than the change of state, believing it to be more useful unit for description and analysis for our purpose.

A 4W Structure for Representing Events

We have identified activity, time and location as salient characteristics of events. Agents, human or other, are also commonly involved. In biographical narratives the biographee is ordinarily the primary agent, or at least a participant, but others may also be involved, starting with the mother at birth. It is reasonable to expect there to be values for each of what, where, when and who when describing or representing any biographical event even though our knowledge of the details of each might sometimes be lacking. What emerges is a case for using the 4W facets – what, where, when and who.

The what facet initially proved to be problematic. There is a tendency to equate “what” with the grammatical object of any narrative sentence, but so long as we are concerned with activities, what should be more associated with verbs, and the solution is to reserve what for what kind of activity, or happening or change. If from 1940 through 1944 he grew potatoes, potatoes were what was produced, an outcome, but the activity was a kind of agriculture.

The Granularity of Events

If we adopt “event” as the unit of analysis and representation for biographical narratives, how detailed an approach is desirable? This question takes two forms. First, should one be concerned with only large or also small events? Our answer is that so long as an event-based approach proves effective, it should be effective for any granularity – any level of detail – so the choice of how far to extend the treatment to smaller events should depend on the purpose of the analysis. Second, as with any form of mark-up, there is a tendency to want to be complete in the sense of including more and more detail. But for our purposes, this temptation should be to strongly resist because the more detailed the mark-up and the more encyclopedic the description, the greater the cost and the less the likelihood that the work would ever be completed. The mark-up itself (and its structure) can be treated as a convenient abstraction of the narrative text (and its structure), but that quality should not imply that it is a replacement. For detailed questions, reference should be made to the text itself.

Collective Biography and Prosopography

A biography is ordinarily a narrative of a single person’s life, although other people will be mentioned. An account of the lives of several persons is usually called a collective biography and a volume of biographical descriptions is a biographical dictionary. The word prosopography originally meant the description of the features of a single individual, but its meaning has been extended to the description of a defined group of people sharing some
significant cultural characteristic and, especially, the relationships among them and the attributes they have in common. Nowadays, a prosopography is ordinarily a database describing the same attributes of each of a set of people in relatively standardized terms. In this way, prosopographies are in the mainstream of humanities computing: the creation of a corpus and of methods to look into it to analyze in various ways the content of the corpus.

Our own interests are significantly different. We are primarily interested in making texts, and especially biographical dictionaries, easier to understand. Since the difference between seeing and understanding lies in knowing the background and context of whatever it is we see, our concern is in facilitating convenient access to relatively trustworthy explanatory resources. So instead of the inward-looking perspective characteristic of prosopographical studies, our interest is in looking outwards from any fragment of any description of any person to any accessible trustworthy explanatory resource outside of the text.

Terminological Issues

Any description needs vocabulary. The nature of our interests has two consequences for choices of vocabulary. First, of course, the terminology has to be suitably descriptive of the 4W facets. As already noted, the naming of persons, places and time has received attention, so our interest has mainly been on what, what kind of event.

Our interest in linking outwards to external explanatory resources suggests that we should draw on a widely used external resource such as the Library of Congress Subject Headings (LCSH) [7] or at least a vocabulary that will map to it. Our rather limited initial examination of biographical records suggests that LCSH could be used. The attraction, of course, is that LCSH, used in library catalogs and some other bibliographies, takes one immediately to the available scholarly literature on whatever the activity is. Also LCSH is more likely to have (or acquire) interoperable mappings to other widely used vocabularies than a locally created one. Another possibility is to use article names from Wikipedia [8], which are organized in a thesaurus-like structure, are more comprehensive and up-to-date than the LCSH and are also being mapped to a variety of other vocabularies.

Locally developed vocabulary may be necessary for specialized texts where fine distinctions and technical terms matter. Specialists use and need specialized terminology. With specialized vocabularies and, indeed, with more general ones, the need arises to map the terms used locally to the corresponding terms used in catalogs and other scholarly resources. This problem resembles the need to map from the geographical description codes (aka feature types like castle, lake, inhabited place, airport) used in place name gazetteers to LCSH in order to connect objects on the ground to literature concerning that kind of object. In this case we have found that comparison of the National Geo-Intelligence Agency’s Geographical Description Codes [9] with Library of Congress Subject Headings reveals differences in style and in emphasis, as well as scope and scale, with some 600 NGA GDC codes to over 150,000 Library of Congress subject headings. Sometimes LCSH has greater detail, especially for kinds of historic sites; sometimes NGA has more detail (for example, in submarine geomorphology), but, in general, they match quite well [4, pp 380-381]. There is a need here for the adoption of search-term recommender services when moving to or between vocabularies.

Whatever vocabulary is used for the mark-up of biographical texts, if, as in our case, the intention is to connect fragments of the text to external explanatory resources, it is increasingly important that the vocabulary chosen lends itself to easy interoperability with the proliferation of naming services associated with the Semantic Web.

Conclusion and Unresolved Issues

This paper is a progress report. So far we have experimented primarily with event-based mark-up of concise biographical texts created by archivists and by editors of scholarly texts. The former have been <bioghist> fields found in Electronic Archival Description-compliant descriptions in the UK Archives Hub, a repository of records for British archival collections [10]. The latter have been brief biographical records created as part of The Emma Goldman Papers project at the University of California, Berkeley [11].

We assume that RDF (Resource Description Framework) and OWL (Web Ontology Language) specifications should be used in order to achieve maximum interoperability in the emerging Semantic Web environment.
Interoperability is especially important since we are trying to relate words, names and phrases in biographical texts to external explanatory resources. A report on our preliminary vocabulary development will be published separately.

A number of questions and unresolved issues remain:

1. Institutions play a large role in archival and biographical texts, so the “biography” of institutions needs comparable attention.

2. We have found a need to distinguish between personal biographic events and contextual events. The Great Depression and the Second World War were major events that affected the environment of personal lives, but they were not major personal events in the same way as, for example, getting married is. These differences regarding life events and cultural-historic events are generally addressed in terms of mereological (part-to-whole) relationships in most event ontologies. Although in some instances biographic events can have a part-to-whole relation to larger contextual events, in many cases they may not. For instance, a person can get married during the Second World War, but it does not follow that the marriage was a part of the war. We see a need for a vocabulary of biographic event types that is distinguishable from the event vocabularies specified in other ontologies.

3. Events are often related to each other. To make these links explicit would imply explicit enumeration of events within or external to the mark-up.

4. The terminology used in the mark-up of biographical texts is likely to vary from the wording in the text being marked-up. Retaining or associating fragments of the original text along with the mark-up terminology could provide a training set for programs to provide computer-aided mark-up.

More experience is needed, especially in mapping between different vocabularies and in the varieties of uses of event-based mark-up, but the provisional conclusion is that an event-based 4W form of mark-up will work as intended.

ACKNOWLEDGMENTS – We are grateful for support from the Institute of Museum and Library Services for Award LG-06-06-0037-06 “Bringing Lives to Light: Biography in Context” and from the Advancing Knowledge grant PK-50027-07 “Context and Relationships: Ireland and Irish Studies” jointly funded by the National Endowment for the Humanities and the Institute of Museum and Library Services. We benefited from initial analyses by Ms. Aurélie Bénard of Paris.

Resources Cited in the Article

[10] Archives Hub: www.archiveshub.ac.uk/
Am honored and, in fact, thrilled to receive the ASIS&T Award of Merit. ASIS&T is my professional home and has always been very special to me. My first meeting was the 1976 meeting in San Francisco, and I have only missed one or two Annual Meetings since then. Submitting a paper or panel presentation to ASIS&T is a top priority for me each year, as is submitting my best work to JASIST and ARIST. All of my mentors have been active in ASIS&T (in fact, many served as ASIS&T president). On a very personal note, my husband Jerry Lundeen and I met at the 1979 Mid-Year Meeting in Banff. Our son Andy attended his first ASIS&T meeting in 1985 when he was 3 months old.

These memories are not merely a self-indulgent preface to my remarks; many of them exemplify aspects of my main message today – that is, first, the importance of professional societies; second, preserving research for future generations; and third, building strong mentoring relationships. Research and scientific discovery are facilitated and enhanced through strong professional societies, strong interpersonal mentoring relationships and ensuring that the building blocks and products of our scholarship including articles, reports and data are preserved in ways that make them findable, accessible and usable into the future.

First: Importance of Strong Professional Societies

Strong professional societies bring together present and future colleagues, help identify potential collaborators, expand knowledge, challenge assumptions and stretch thinking. They provide an intellectual home as well as local, national and international venues for publication and presentation. These interactions can be face-to-face as we are here, virtual or, better yet, both.

Membership in many small professional societies has decreased over the last decade. Some people blame the economy – saying dues are too expensive or employers no longer pay for dues. Others blame changing cultures – saying the younger generation doesn’t value professional memberships or see the benefits. Still others blame technology – saying why belong to a professional society when you can easily and instantly interact with anyone virtually? Perhaps all of these may be contributing factors.

For all of us who know the value of professional societies, it is important to work together to make them flexible and relevant into the future by providing multiple ways to participate and interact. For those of us who work with students or young professionals, it is important that we encourage participation and provide opportunities for leadership and new ideas. And, although I am preaching to the choir by saying this to those who are in attendance here, attending and participating in face-to-face conferences when possible is important to building and sustaining a career and professional identity and hearing new ideas.

At the start of my career, my employers did not always pay for me to attend ASIS&T meetings. After rent, utilities and food, ASIS&T conferences were next on my priority list for several years. It was an investment in my future that I will never regret.

Second: Preserving Research for Future Generations

High-quality publication outlets are another important contribution from professional societies. No matter your age, it is time to start thinking about the preservation of your professional legacy by participating in the web of scholarship. Publishing your best work in high-quality peer-reviewed journals such as JASIST is one way, placing your reports, lecture materials and data in institutional or subject repositories is another. Multiple modes of publication and access will help preserve our intellectual heritage into the future and enable new discoveries. There is not just one solution, nor one type of information that we should focus on. ASIS&T members have the obligation...
in our role as authors, but also the expertise and interest to take a leadership role in preservation and sustained access to other people’s information. ASIS&T members are leading participants in the first two National Science Foundation (NSF) DataNET projects, described in a session yesterday afternoon. DataNet (or Sustainable Digital Data Preservation and Access Network Partners) from the NSF Office of Cyberinfrastructure seeks to build exemplar organizations and interdisciplinary solutions to digital data capture, preservation, use and reuse. NSF describes it this way:

- The new types of organizations envisioned…will integrate library and archival sciences, cyberinfrastructure, computer and information sciences, and domain science expertise to
  - provide reliable digital preservation, access, integration and analysis capabilities for science and/or engineering data over a decades-long timeline;
  - continuously anticipate and adapt to changes in technologies and in user needs and expectations;
  - engage at the frontiers of computer and information science and cyberinfrastructure with research and development to drive the leading edge forward; and
  - serve as component elements of an interoperable data preservation and access network. [1]

Information science in general and ASIS&T members in particular are natural leaders in this important work.

Although we are slower than some others, the University of Tennessee libraries have launched an institutional repository focusing on materials that otherwise might be lost, rather than duplicating the existing peer-reviewed-journal system. I have volunteered to be one of the first faculty members to put my speeches, reports and other materials into TRACE (Tennessee Research and Creative Exchange). A personal as well as a professional commitment to such initiatives is important to preserving our digital heritage.

**Third: Building Strong Mentoring Relationships**

And, finally, I would like to stress the importance of mentoring. Just as we build on past scholarship through citations, record present scholarship through publication and encourage future scholarship by depositing and providing access to data, we pass on our knowledge and experience to new generations of scholars through mentoring.

I have been fortunate to have strong mentors throughout my career. At the start of my worklife, Pamela Cibbarelli, president of her own information consulting firm, gave me guidance in the work world and the freedom to make decisions. As a doctoral student at the University of Illinois, past ASIS&T president Martha E. Williams and incoming ASIS&T president Linda Smith shared their attention to detail, research expertise and knowledge of the information industry. For the last decade and a half Donald W. King, another former ASIS&T president and also an Award of Merit winner, has shared his extensive research experience, dedication to writing and even his data. Don is in Japan this week at a conference that is featuring the results of our surveys of reading at 20 Japanese universities. We decided it was more important for me to be here.

One never outgrows the need for a mentor, even when turning to mentoring others. Passing on our ideas, insights and even our data to mentees is an important way that scientific discovery advances. I am proud to count ASIS&T members Suzie Allard, Lorraine Normore and Peiling Wang as past and current mentees.

But I am straying into personal reminiscing again, so let me conclude by reiterating the importance for every scholar to bridge generations and build scholarship by serving as both a mentee and mentor; to preserve and pass on our best work, data and reports in scholarly journals and repositories; and to support and improve strong professional societies that provide formal and informal opportunities for collaboration and learning.

ASIS&T has been the center for these activities for my entire career, so I thank you once again for this very special honor.

---

**Resource Cited in the Article**

I’m sitting with my elderly parents at the lunch table. My mother folds the newspaper, hands it to me and says, “You’ll find sad news in here.” “Oh, no, an obituary?” I ask and then read of the sudden passing of an 87-year-old woman I had known for seven months. Actually, I had not really known her for seven months. I merely met her seven months ago when she shared a rehabilitation room with my mother following their respective surgeries. Yet Wanda’s passing was a surprise. I would often think of her as I drove near her house, and she had come up in conversation the day before. Despite her age we didn’t think of her as someone who would pass away following a “sudden illness.”

I don’t take well to sad news. So I raked leaves. Drank bourbon. Rubbed my eyes a lot. And tried to figure out why I felt moved by Wanda’s passing. What was it about Wanda? I only spent “visiting hours” with her for one week. Yet I felt a big loss.

After some reflection, I suspect I appreciated Wanda’s big spirit, unorthodox nature and robust energy. I also realized that, in a short time, she made a profound impact on me.

And that’s why I’m mentioning this story in my new *Bulletin* column. I always wanted work that made a profound impact on others. I gravitated to the field of information architecture because I was attracted to the profound impact available to practitioners in this field.

Years ago, when I was beginning work as an information architect, I was sitting at a local café. I listened to a young woman, a recent college graduate (or so I assumed as someone dipping into her conversation), who sat nearby describing her work-world frustrations to her father: “Don’t you see, I don’t want to be a marketer,” she whined. “I want to be in a field where I can help people.”

“Information architects help people,” I might have blurted. Surprisingly, I didn’t.

As a teacher I get to evangelize about the possibilities inherent in our work as information architects. I show a photo of Thom, grumpy in his mauve cubicle, with the text: “Early work days. Often built deliverables nobody wanted to use.” The slides show a problem with the process – how the contracts define and the attention is paid to the *deliverable*, not to the *use* of the product or site.

As information architects we tend to build a lot of deliverables – products we provide to communicate design (preferred) or clarify the end of a period of work (costly). Deliverables can be fascinating and useful. But the profound impact of our work takes place at the moment someone is able to use a product or content we have helped develop.

I see we do our best work when we think about the helping humans. Although I have attended eight IA summits,

Thom Haller, the *Bulletin*’s new 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.
I’ve only attended one ASIS&T Annual Meeting. But I do recall the enthusiasm that practitioners brought to their case studies on helping people retrieve information. I wonder if our enthusiasm for helping others differentiates us from other professionals.

How profound is our impact? I’m still trying to learn that, but I think it’s a question worth addressing. I’m hoping this column provides me a space to investigate and articulate that question. And perhaps, synthesize and share some answers.

Help me out:
- Do you see your work having a profound impact? If so, how? If not, why not?
- Do you find inherent possibilities in your work? Do you share these with others? How?
- Is helping people part of the reason you pursued your schooling or professional choices?
- How do you perceive information architecture – especially as someone who didn’t enter the field from a design or user experience perspective?