SPECIAL SECTION

Health Informatics

18] Introduction
by Prudence Dalrymple, Special Section Guest Editor

20] Informatics and the Future of Nursing: Harnessing the Power of Standardized Nursing Terminology
by Patricia M. Schwirian

25] Public Health Informatics: An Invitation to the Field
by Mary White

30] Empowerment or Anxiety? Research on Deployment of Online Medical E-health Services in Sweden
by Isto Huvila, Gunilla Myreteg and Åsa Cajander

34] Turning Healthcare Challenges into Big Data Opportunities: A Use-Case Review Across the Pharmaceutical Development Lifecycle
by Timothy Schultz

41] Visioning Studies in Emergency Health Care to Support R&D of New Technologies
by Diane H. Sonnenwald

FEATURE

45] Transforming Our Conversation of Information Architecture with Structure
by Nathaniel Davis

48] Using a Taxonomy for Your Database or Website: A Look Behind the Scenes
by Marjorie M.K. Hlava
The special section in this issue is devoted to health informatics. My thanks to Prudence Dalrymple, our guest editor, and her authors for a broad look at developments in this field. Topics include standardized nursing terminologies, public health informatics, e-health, big data in pharmaceutical and other health informatics applications and the use of visioning to explore the potential of future technologies.

Our two feature articles focus on information architecture (IA) and taxonomies, respectively. Nate Davis argues that structure is a defining element of IA as a field, while Marjorie Hlava details for practitioners how a taxonomy can be integrated with a database or website application.

Inside ASIS&T gives the flavor of two recent ASIS&T events: the Research Data Access and Preservation (RDAP) Summit and Information Architecture (IA) Summit. Our annual issues on these topics will be in August/September and October/November. Our RDAP Review columnist for this issue is Rachel Mandell from the University of California at Los Angeles. Her topic is the UCLA data registry.

Finally, on the President’s Page, ASIS&T president Andrew Dillon discusses the future of our Association in the context of declining membership among all scholarly societies.
In Race for Relevance, a critical examination of the future for professional associations (ASAE, 20011), authors Harrison Coerver and Mary Byers argue that the landscape for such associations has altered drastically as a result of new technologies and that few have managed to respond appropriately. Their arguments are somewhat simplified, but their conclusions seem to fit the data. Most professional associations are struggling to sustain, never mind grow, membership, and ASIS&T is no different. Our latest figures continue to show a slow decline in our regular membership, a trend we have witnessed for the last few years. Currently we have 1100 regular members but we can add to this a somewhat healthy 591 student members and a further 100 institutional members. That said, these data points are real and should give us pause as we plan the long-term future of our association. Since there are far more than that number of students pursuing information-related degrees around the world, one imagines we can do a better job of recruiting new members here. And once recruited, a crucial issue is converting these students into regular members once they become employed.

When one considers joining a professional association, costs are measured against benefits, and one wonders if we do all we can to really explain the many benefits one receives
from membership. Typically, we can distinguish two benefit types: tangible and other. The tangibles, such as easy or discounted access to the association’s publications and conferences tend to be taken for granted after awhile, and we must acknowledge that for some, these benefits may not even be valued too highly in the first place. But it is important to note that these benefits involve more than access. *JASIST*, through the publication agreement with Wiley, brings much-needed revenue into the association. Further, annual membership dues alone are completely offset by the reduced rate a member pays in registration at the Annual Meeting. Since our dues are often lower than those for comparable societies, once you add in the other tangibles such as discounted rates for seminars, webinars and summits, access to the jobline, plus free membership to a special interest group or local chapter, there is a suite of benefits to be obtained from joining.

I suspect it is time to revisit the tangibles to align them with member interests, but it is the many intangible benefits of membership that give real purpose to joining an association. Among them are the opportunities for networking or community building and for the professional identity, socialization or credentialing that membership provides.

looking forward, with the anticipated growth in our international coverage, one annual conference, invariably in North America, seems unlikely to provide the real engine for growth. Participation in the life of the association must be increased throughout the year.

We are not yet at the stage of a membership drive, but at the board level we are working on initiatives to strengthen and enhance relationships across the information world. For example, as reported in Inside *ASIS&T* (page 7), we have just completed formation of a management partnership with the Dublin Core Metadata Initiative (DCMI) (www.dublincore.org). This relationship will provide opportunities for synergistic partnerships between *ASIS&T* and those with a commitment to improving the global metadata ecosystem. Beyond management logistics, we are hoping this new partnership can provide increased opportunities for metadata specialists to engage with *ASIS&T* at the Annual Meeting and beyond. As a colleague here reminded me, most scholars in the digital humanities have heard of DCMI, but few of them have heard of *ASIS&T*. Let’s rectify that situation.

We should also recognize that we have tremendous synergies with the broad information architecture community. *ASIS&T* was instrumental in setting up the first IA Summit in 2000, which I attended as a speaker. What was intended as a one off meeting has spawned a very successful annual conference here in the United States and has been joined by a Euro-IA meeting. As an association, these events happen under our auspices, so to speak, and we benefit financially from the healthy attendances they attract. While there is some overlap, these communities have become largely independent. I am keen to explore ways in which we can build more bridges among the IA community (including the IA Institute), DCMI and *ASIS&T*. Staying small has advantages but they are invariably outweighed by the threat of becoming irrelevant. One possible future for *ASIS&T* is for us to become an umbrella professional grouping under which somewhat distinct
practitioner or scholarly communities affiliate. Going forward, I believe we have to be flexible in imagining how ASIS&T will be structured as a professional association if, as some believe, the classic dues-paying membership model will no longer sustain us.

This is all for the future but we are laying the groundwork now. It is all well and good to write self-promotional material, replete with tag lines, claiming to be “the” information society for the information age, but we have to demonstrate this distinction. In an age when the communication network of an association is no longer unique or even the best way for those with similar interests to share ideas, it is time to ask some fundamental questions about what the real benefits of being a member might be and where we want this organization to go in the years ahead. For me, the tangibles and intangibles need to be revisited, and I welcome your input. Previous presidential columns remind you that this is your association. Let us make sure that an increased sense of ownership is the result of more engagement, not a result of severely reduced membership. If you have views on what you’d like to see us do or achieve, let me know.
Summer Balloting to Begin Soon for Board Member Election and Bylaws Amendment

It’s just about time for the ASIS&T membership to exercise its right to help guide the direction of the organization through the election of directors and officers of the Board of Directors. Electronic ballots are being readied for distribution to eligible voters in the coming weeks. Members will be asked to elect a president-elect; two directors and a treasurer.

The following candidates will be on the ballot this year:

**FOR PRESIDENT**
- Sandra Hirsch, professor and director, SLIS, San Jose State University
- Prudence Dalrymple, research and training professor, Drexel University

**FOR DIRECTORS-AT-LARGE (top two vote getters will be elected)**
- June Abbas, associate professor, University of Oklahoma
- Mei-Mei Wu, professor, National Taiwan Normal University, Taipei
- Lauren Harrison, principal information scientist, Hoffman La Roche, Inc.
- Isto Huvila, Abo Akadem i University, Abo, Finland

**FOR TREASURER**
- Vicki Gregory, professor, University of South Florida
- Lisl Zach, assistant professor, Drexel University

**Bylaws Amendment**

Along with the election of new officers and directors of the ASIS&T Board of Directors, this summer’s ballot will include a bylaws amendment for the membership’s vote. The proposed amendment is described at left.

From Board meeting minutes:

Report from Constitution and Bylaws Committee calls for shortening the current voting window by half, to 30-35 days vs. 60-70 days. Discussion of optimum voting window takes place.

**Motion: Dillon Second: Mai**

Move that votes put to the membership must be returned within 35 days from the date of sending.

Approved unanimously

**Bylaw Amendment current wording, with proposed changes noted:**

(strikeout = language to be removed; red and underlined = new language to be inserted):

Ballots. Whenever any action requires a ballot of the membership of the Society, the Executive Director shall prepare the ballots, including instructions for use, and shall notify the Members that they may vote by returning their marked ballots to the Executive Director and that, to be counted, ballots must be received by the Executive Director on or before a specified date. The date so specified shall be no less than 60 nor more than 70 days from the mailing of the ballot. The Executive Director shall supervise the tellers in the counting of all ballots, whether cast by mail/e-mail or at an Annual or Special Business Meeting.

[Other related sections will be changed to be in accord with the above language.]

SUMMER BALLOTING CONTINUED ON PAGE 7
ASIS&T Announces Management Partnership with DCMI

The Dublin Core Metadata Initiative (DCMI), an internationally renowned organization advancing innovation in metadata design and best practices, will become a project of the Association for Information Science and Technology (ASIS&T) upon DCMI’s wrapping up activities at its current location in Singapore.

On June 30, 2013, DCMI Ltd. will cease operations as a company limited by guarantee in Singapore and become a project of ASIS&T. This change for DCMI from independent, non-profit company status in Singapore to a partnership with ASIS&T marks a significant milestone in DCMI’s history. The decision was motivated by the desire of DCMI’s governing Oversight Committee to shape a more flexible and progressive institutional structure, while retaining its mission, goals and objectives and its commitments to an open, consensus-driven community.

Andrew Dillon, president of ASIS&T, stated that the partnering of ASIS&T with DCMI makes excellent sense for both parties. “There is a considerable overlap in participation and subject matter interest in both groups, and information science as a practice and as a discipline will be strengthened by the partnership,” said Dillon. “We look forward to regular cross-fertilization through meeting sessions, workshops, webinars and other forums,” Dillon continued.

The change in DCMI institutional structure was guided by extensive investigations by DCMI’s Oversight Committee into alternative structures and potential partnerships that might better serve DCMI’s global community of metadata researchers and practitioners. Mutual and unanimous decisions were reached by the governing bodies of both DCMI and ASIS&T that DCMI become a project of ASIS&T effective July 1, 2013. Going forward, DCMI will maintain its autonomy as a global community through maintenance of the DCMI “brand,” its governance structure and its programmatic commitments and membership programs while actively engaging with ASIS&T as a partner in mutually beneficial activities.

Thus, DCMI and ASIS&T are committed in the coming months and years to seek out joint activities where both organizations can leverage their strengths to the benefit of both communities. Both DCMI and ASIS&T bring to the new partnership their histories as strong, successful organizations serving research and practice communities.
Plans are well underway for the 76th Annual Meeting of the recently renamed Association for Information Science and Technology. Addressing the dramatic changes taking place in the information world with the advent of cloud computing and ever-greater storage capabilities, the meeting will focus on the continuing development of computing and mobile technology and the ongoing evolution of the web environment. Speakers will address the new ways for accessing, acquiring, retrieving and storing information which constantly defy traditional boundaries.

With the growth of digital content, information objects are blurred, and they challenge information organization. As more people interact and exchange knowledge and information on the web and in the cloud, information environments are transformed, and human information behaviors shift. As information use increases and becomes more complex, the need for meaningful integration and analysis grows. Maintaining its reputation as the main venue for disseminating research on advances in information science, information technology and related topics, this year’s Annual Meeting offers an opportunity to reflect on all the changes that impact human information interaction and their implications for information science and technology.

Jorge Garcia, senior business intelligence (BI) research analyst for Technology Evaluation Centers in Montreal, will deliver the keynote speech at the meeting. With more than 20 years of experience in all phases of application development and database design, Jorge will talk about augmented reality (motion and voice recognition technologies) – an area that has received a great deal of R&D attention and dollars in recent days.

While Jorge prepares for his presentation, the Annual Meeting program committee will continue its review of the hundreds of papers and panel session suggestions it has received from scholars, practitioners, researchers and students vying for spots in the six dozen or so sessions that will appear in the final program for the conference.

And a bit later this summer, following a July 1 submission deadline, the committee’s attention will switch to selecting participants for the interactive showcase, an evening devoted to face-to-face presentations using traditional papers, short videos or demonstrations. Please see the complete Call for Participation at the ASIS&T website for details on requirements for submitting in this category.

Conference Committee
France Bouthillier, McGill University, and Boryung Ju, Louisiana State University, are chairing this year’s ASIS&T Annual Meeting. Serving with them on the program committee are Pascal Calarco, University of Waterloo Library, interactive showcase; Lisa M. Given, Charles Sturt University, posters; and Jens-Erik Mai, Royal School of Library and Information Science, panels.

Also in Montreal
In addition to conducting its 76th Annual Meeting in Montreal, ASIS&T will concurrently host the 9th International Conference on Knowledge Management (ICKM 2013) at the same venue on November 1-2. ICKM will explore the themes of knowledge management metrics, performance measurement, capacity building and certification.

ICKM provides researchers and practitioners from all over the world a forum for discussion and exchange of ideas concerning theoretical and practical aspects of knowledge management. The conference provides a fantastic opportunity for researchers, practitioners, innovators, consultants and academics to exchange ideas and insights and to contribute to the advancement of knowledge management as a discipline and practice.
t was 13 years ago that Lou Rosenfeld, Peter Morville, Gary Marchionini and Victor Rosenberg got together to plan the first ASIS&T summit on information architecture. And from that first effort that possibly exceeded everyone’s expectations, the IA Summit was born. Earlier this year, April 3-7, the 14th IA Summit convened in Baltimore, Maryland, for the latest installment of the premiere, community-curated and volunteer-run gathering on the ever-evolving disciplines of information architecture and experience design. Once again, students, practitioners and anyone interested came together to collaborate and celebrate the great things happening in these communities of practice.

Emboldened by the success of the IA Summit series, in 2010 ASIS&T held its first Research Data Access and Preservation Summit, co-located with that year’s IA Summit in Phoenix. RDAP was designed to bring together leaders in data centers, laboratories and libraries in different organizational and disciplinary settings to share ideas and techniques for managing, preserving and sharing large-scale research data repositories with an eye toward achieving infrastructure-independent access and stewardship. That summit was also the start of something big, and earlier this year, again co-located with the IA Summit, the 4th Annual RDAP Summit featured expert panel presentations, an interactive poster session and lightning talks on the themes of institutional repositories, data citation and altmetrics, data infrastructure, linked data and metadata, and data use and reuse.

Check out some of the activity at each of the 2013 summits in Baltimore in the following photo montage.
4th RDAP Summit
April 4-5, 2013
Baltimore
The Association for Information Science & Technology (ASIS&T) is seeking a volunteer to serve a 3-year renewable term as ASIS&T Monograph Series editor, acquiring and developing three to six print and digital titles annually on a range of practical and theoretical topics in the broad field of information science and technology. This is an excellent opportunity for an individual who wishes to take a leadership role in promoting and disseminating innovative research, practice and scholarly discourse in the field while making a significant contribution to his or her professional association.

The ideal candidate is an ASIS&T member with strong academic credentials in library and information science and technology; a commitment to keeping current on relevant research and practice in the field; an appreciation for the role of authors, editors and publishers in advancing knowledge and understanding; strong communication, language and networking skills; the ability to identify, motivate and mentor promising and established authors and editors; a commitment to meeting deadlines; and the vision, energy and dedication needed to grow a small but well regarded imprint into a leading program of publications for scholars, students, practitioners and other stakeholders in library and information science and technology.

The ASIS&T monograph series is published in print and ebook formats on behalf of ASIS&T by Information Today, Inc. (ITI). Examples of current print and digital titles can be found at http://books.infotoday.com/asist/ and http://books.infotoday.com/ebooks/.

For more information or to apply contact: John B. Bryans
Editor-in-Chief & Publisher
Book Publishing Division
Information Today, Inc.
609/654-6266, x134
jbryans@infotoday.com

The ASIS&T Task Force on Web Presence is conducting a broad-based survey aimed at identifying desirable characteristics of the organization’s social media. The committee invites all members to take a survey on ASIS&T’s web and social media presence. Of particular interest are users’ experiences with ASIS&T’s website, publications and social media. The survey should take no more than 10-15 minutes to complete. All questions are optional, but the task force welcomes and appreciates all answers and feedback. The responses from participants will be used by the task force to inform recommendations on how to improve ASIS&T’s web presence to better serve its members. You can complete the survey by visiting https://www.surveymonkey.com/s/asistwebpresence

If you have any questions or comments about this survey, e-mail Diane Rasmussen, ASIS&T director-at-large and chair of the ASIS&T Task Force on Web Presence, at diane.m.rasmussen@gmail.com.
The Potomac Valley Chapter of ASIS&T, co-sponsoring with Catholic University of America and its ASIS&T Student Chapter, presented a fascinating program on visual analytics and big data. Yair G. Rajwan, director of Analytics Visualization at Visual Science Informatics, discussed the four core components of visual analytics methodology: data architecture, data analysis, information visualization and evaluation methods. Using concrete examples, he illustrated how information professionals can use visual analytics to provide comprehensive point of views, create actionable intelligence and engage audiences through effective visual communication.

The annual networking dinner of the New England ASIS&T Chapter featured David Weinberger discussing Library as Platform: Opening Up to Support Our Communities in the Production of Knowledge. Weinberger is a senior researcher at Harvard’s Berkman Center for Internet & Society and is co-director of the Harvard Library Innovation Lab. He writes about the effect of the Internet on ideas.

The Northern Ohio ASIS&T Chapter featured Trevor Watkins presenting an interactive overview of robot teams for its May meeting. Co-author of Build Your Own Teams of Robots with LEGO Mindstorms NXT and Bluetooth, Watkins is a network communications and system integrations specialist. He is currently the technology manager at the Wadsworth Public Library and is an adjunct professor at Youngstown State University and Eastern Gateway Community College.

Helen Tibbo, alumni distinguished professor in the School of Information and Library Science at the University of North Carolina at Chapel Hill, received a 2013 Laura Bush 21st Century Librarian Program grant of nearly $500,000 for her project, CRADLE: Curating Research Assets and Data Using Lifecycle Education Data Management Education Tools for Content Creators, Librarians and Archivists. Among other objectives, the project team hopes to produce high-quality massive open online (MOOC) data management instructional courses, provide multimedia self-paced courses through online free university platforms and coordinate face-to-face workshops.

Miles Efron, assistant professor in the Graduate School of Library and Information Science at the University of Illinois, has received a three-year grant of more than $400,000 from the National Science Foundation for his work on new algorithms that build upon the current strengths of search but add a new dimension — time. “We are at a point where soon we won’t have the luxury of ignoring the temporal aspect of data,” said Efron. “In order for search to be successful, time has to make its way into search engines.” In addition to news of his NSF grant, Efron has also learned that he has been awarded tenure at the university; effective with the upcoming academic year, he will be associate professor at GSLIS.

Also from the University of Illinois, Carole Palmer, professor and director of the Center for Informatics Research in Science and Scholarship, delivered the inaugural Ed Mignon Distinguished Lecture in Information Science at the University of Washington iSchool in May. Her talk was titled, “Data Curation and the Reuse Value of Digital Research Data: Meeting the Aims of Multiple Disciplines and Stakeholders.”

Jasy Liew Suet Yan, student at Syracuse University, is the winner of the 2013 NEASIS&T Student Travel Award, presented to the student who submits the most compelling essay addressing the contest’s questions. Award jurors said that Jasy’s essay was well-written and persuasive. In it, she explains the importance of sharing her preliminary research with like-minded individuals who will attend the ASIS&T Annual Meeting. She will receive a cash award to help defray her expenses in attending the meeting in Montreal.

ASIS&T Presents Annual Audit

The report of the ASIS&T auditors on the 2012 financial statements is presented on the following pages.
**ASIS&T Annual Audit**

The report of the ASIS&T auditors on the 2012 financial statements is presented here and on the following pages.

ASSOCIATION FOR INFORMATION SCIENCE & TECHNOLOGY
(FORMERLY AMERICAN SOCIETY FOR INFORMATION SCIENCE & TECHNOLOGY)

FINANCIAL STATEMENTS
AND
INDEPENDENT AUDITORS’ REPORT
SEPTEMBER 30, 2012 AND 2011

INDEPENDENT AUDITORS’ REPORT

Board of Directors
ASIS&T (formerly American Society for Information Science & Technology)
Silver Spring, Maryland

We have audited the accompanying statements of financial position of the Association for Information Science & Technology (formerly American Society for Information Science & Technology) (a nonprofit organization) as of September 30, 2012 and 2011 and the related statements of activities, functional expenses, and cash flows for the years then ended. These financial statements are the responsibility of the Organization’s management. Our responsibility is to express an opinion on these financial statements based on our audit.

We conducted our audit in accordance with auditing standards generally accepted in the United States of America. Those standards require that we plan and perform the audit to obtain reasonable assurance about whether the financial statements are free of material misstatements. An audit includes examining, on a test basis, evidence supporting the amounts and disclosures in the financial statements. An audit also includes assessing accounting principles used and significant estimates made by management, as well as, evaluating the overall financial statement presentation. We believe our audit provides a reasonable basis for our opinion.

In our opinion, the financial statements referred to above present fairly, in all material respects, the financial position of the Association for Information Science & Technology as of September 30, 2012 and 2011, and the changes in its net assets and its cash flows for the years then ended, in conformity with accounting principles generally accepted in the United States of America.

Casadio Waddell, LLC
Columbus, Maryland
January 10, 2013

ASSOCIATION FOR INFORMATION SCIENCE & TECHNOLOGY
(FORMERLY AMERICAN SOCIETY FOR INFORMATION SCIENCE & TECHNOLOGY)
STATEMENTS OF FINANCIAL POSITION
SEPTEMBER 30, 2012 AND 2011

<table>
<thead>
<tr>
<th>ASSETS</th>
<th>2012</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Assets</td>
<td>$1,760,068</td>
<td>$1,006,013</td>
</tr>
<tr>
<td>Cash and Cash Equivalents - Headquarters</td>
<td>127,971</td>
<td>123,212</td>
</tr>
<tr>
<td>Cash and Cash Equivalents - Local Chapters</td>
<td>563,165</td>
<td>-</td>
</tr>
<tr>
<td>Investments</td>
<td>761</td>
<td>9,488</td>
</tr>
<tr>
<td>Accounts Receivable, Net of Allowance</td>
<td>2,565</td>
<td>2,872</td>
</tr>
<tr>
<td>Prepaid Expenses - General</td>
<td>11,477</td>
<td>17,981</td>
</tr>
<tr>
<td>Prepaid Expenses - Conferences</td>
<td>51,829</td>
<td>96,278</td>
</tr>
<tr>
<td>Prepaid Expenses - Taxes</td>
<td>-</td>
<td>60</td>
</tr>
<tr>
<td>Postage Deposits</td>
<td>4,001</td>
<td>5,065</td>
</tr>
<tr>
<td>Total Current Assets</td>
<td>$2,865,092</td>
<td>$2,102,167</td>
</tr>
<tr>
<td>Property and Equipment, net</td>
<td>49,679</td>
<td>54,754</td>
</tr>
<tr>
<td>Other Assets</td>
<td>2,468</td>
<td>2,468</td>
</tr>
<tr>
<td>Security Deposit</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total Assets</td>
<td>$2,952,149</td>
<td>$2,217,729</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>LIABILITIES AND NET ASSETS</th>
<th>2012</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liabilities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts Payable</td>
<td>$27,512</td>
<td>$27,575</td>
</tr>
<tr>
<td>Income Taxes Payable</td>
<td>1,038</td>
<td>485</td>
</tr>
<tr>
<td>Accrued Expenses</td>
<td>68,327</td>
<td>93,192</td>
</tr>
<tr>
<td>Deferred Revenue</td>
<td>441,387</td>
<td>391,881</td>
</tr>
<tr>
<td>Total Liabilities</td>
<td>539,344</td>
<td>443,043</td>
</tr>
<tr>
<td>Net Assets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unrestricted, Undesignated</td>
<td>1,818,392</td>
<td>1,200,024</td>
</tr>
<tr>
<td>Unrestricted, Board Designated</td>
<td>477,443</td>
<td>469,662</td>
</tr>
<tr>
<td>Total Net Assets</td>
<td>2,295,835</td>
<td>1,669,686</td>
</tr>
<tr>
<td>Total Liabilities and Net Assets</td>
<td>$2,952,149</td>
<td>$2,217,729</td>
</tr>
</tbody>
</table>

Note: Independent Auditors’ Report and Notes to Financial Statements.
### ASSOCIATION FOR INFORMATION SCIENCE & TECHNOLOGY (FORMERLY AMERICAN SOCIETY FOR INFORMATION SCIENCE & TECHNOLOGY) STATEMENTS OF CASH FLOWS FOR THE YEARS ENDED SEPTEMBER 30, 2012 AND 2011

<table>
<thead>
<tr>
<th></th>
<th>2012</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash Flows From Operating Activities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in Net Assets</td>
<td>$231,119</td>
<td>$287,540</td>
</tr>
<tr>
<td>Adjustments to Reconcile Change in Net Assets to Net Cash Provided by Operating Activities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depreciation and Amortization</td>
<td>$24,103</td>
<td>$25,844</td>
</tr>
<tr>
<td>Unrealized Gain on Investments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Changes in Operating Assets and Liabilities (Increase/Decrease in Assets)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts Receivable</td>
<td>7,717</td>
<td>(3,306)</td>
</tr>
<tr>
<td>Inventory</td>
<td>3,16</td>
<td>278</td>
</tr>
<tr>
<td>Prepaid Expenses - General</td>
<td>6,44</td>
<td>(1,905)</td>
</tr>
<tr>
<td>Prepaid Expenses - Conferences</td>
<td>4,447</td>
<td>2,250</td>
</tr>
<tr>
<td>Prepaid Expenses - Taxes</td>
<td>60</td>
<td>(900)</td>
</tr>
<tr>
<td>Prepaid Expenses - Postage Deposits</td>
<td>904</td>
<td>1,800</td>
</tr>
<tr>
<td>Increase (Decrease) in Liabilities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accounts Payable</td>
<td>(92)</td>
<td>11,626</td>
</tr>
<tr>
<td>Income Taxes Payable</td>
<td>694</td>
<td>(168)</td>
</tr>
<tr>
<td>Deferred Revenue</td>
<td>2,354</td>
<td>(29,477)</td>
</tr>
<tr>
<td>Net Cash Provided by Operating Activities</td>
<td>445,604</td>
<td>192,324</td>
</tr>
<tr>
<td>Cash Flows From Investing Activities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purchase of Investments</td>
<td>(849,650)</td>
<td></td>
</tr>
<tr>
<td>Purchase of Property and Equipment</td>
<td>(12,228)</td>
<td>(4,940)</td>
</tr>
<tr>
<td>Net Cash Used in Investing Activities</td>
<td>(661,872)</td>
<td>(4,940)</td>
</tr>
<tr>
<td>Net Increase (Decrease) in Cash and Cash Equivalents</td>
<td>(113,760)</td>
<td>181,626</td>
</tr>
<tr>
<td>Cash and Cash Equivalents, Beginning of Year</td>
<td>2,029,826</td>
<td>1,842,147</td>
</tr>
<tr>
<td>Cash and Cash Equivalents, End of Year</td>
<td>$1,915,060</td>
<td>$2,029,826</td>
</tr>
</tbody>
</table>

#### Supplemental Disclosure of Cash Flow Information:
- **Cash Paid During the Year for Income Taxes**: $968, $1,805

---

See Independent Auditors' Report and Notes to Financial Statements.
ASSOCIATION FOR INFORMATION SCIENCE & TECHNOLOGY
(FORMERLY AMERICAN SOCIETY FOR INFORMATION SCIENCE & TECHNOLOGY)
NOTES TO FINANCIAL STATEMENTS
SEPTEMBER 30, 2012 AND 2011

SUMMARY OF SIGNIFICANT ACCOUNTING POLICIES (Continued)

Deferred Revenue: The Association's deferred revenue represents membership dues, conference registration fees and publication subscriptions for services which have not yet been provided.

Income Taxes: The Association is exempt from federal and state income tax (except taxes on unrelated business income) under Section 501(c)(3) of the Internal Revenue Code. Although the Association is exempt from income tax, it has certain activities considered unrelated to its exempt status that are subject to income tax. The Association has accrued taxes of $1,105 and $484 for unrelated business income for the years ended September 30, 2012 and 2011, respectively.

The Association follows the provisions of Accounting for Uncertainty in Income Taxes under the Income Taxes Topic of the Codification. The Codification requires the evaluation of tax positions, which include maintaining its tax-exempt status and the taxability of any unrelated business income, and does not allow recognition of tax positions which do not meet a "more-likely-than-not" threshold of being sustained by the applicable tax authority. Management does not believe it has taken any tax positions that would not meet this threshold. The Association's income tax returns are subject to possible federal examination, generally three years after they are filed.

Fair Value Measurement: Generally accepted accounting principles define fair value, establish a framework for measuring fair value, and establish a fair value hierarchy that prioritizes the inputs to valuation techniques. Fair value is the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction between market participants at the measurement date. A fair value measurement assumes that the transaction to sell the asset or transfer the liability occurs in the principal market for the asset or liability or, in the absence of a principal market, the most advantageous market. The Association's financial assets, recorded at fair value on a recurring basis, primarily relate to investments in available-for-sale securities.

The fair value hierarchy prioritizes the inputs to valuation techniques used to measure fair value into three broad levels:

Level 1: Unadjusted quoted market prices in active markets for identical assets or liabilities.

Level 2: Unadjusted quoted prices in active markets for similar assets or liabilities, unadjusted quoted prices for identical or similar assets or liabilities in markets that are not active, or inputs other than quoted prices that are observable for the asset or liability.

Level 3: Unobservable inputs for the assets or liabilities that rely on management's own assumptions.

The following table summarizes the valuation of the Association's assets and liabilities measured at fair value on a recurring basis as of September 30, 2012:

<table>
<thead>
<tr>
<th>Asset Type</th>
<th>Fair Value (Level 1)</th>
<th>(Level 2)</th>
<th>(Level 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corporate Bonds</td>
<td>$200,043</td>
<td>-$</td>
<td>-$</td>
</tr>
<tr>
<td>Mutual Funds</td>
<td>147,621</td>
<td>147,621</td>
<td>-</td>
</tr>
<tr>
<td>U.S. Treasury and Agency</td>
<td>124,374</td>
<td>124,374</td>
<td>-</td>
</tr>
<tr>
<td>Accrued Interest Receivable</td>
<td>2,131</td>
<td>2,131</td>
<td>-</td>
</tr>
</tbody>
</table>

See Independent Auditors' Report.
NOTE 4. PENSION PLAN

The Association sponsors a defined contribution retirement plan that operates under section 403(b) of the Internal Revenue Code. The plan covers all full-time employees and part-time employees with more than 1,000 hours of service. Participation in the Plan begins after completion of twelve months of service. Employers may contribute to the Plan and the Association contributes 5% of the portion of an employee's salary within the Social Security wage base. The Association's contributions to the Plan totaled $21,314 and $21,320 for the years ended September 30, 2012 and 2011, respectively.

NOTE 5. BOARD DESIGNATED NET ASSETS

The Board of Directors of the Association has designated net asset balances into the following funds as of September 30:

<table>
<thead>
<tr>
<th>Fund</th>
<th>2012</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserve Fund</td>
<td>$233,932</td>
<td>$233,932</td>
</tr>
<tr>
<td>Chapter Development Fund</td>
<td>17,815</td>
<td>20,000</td>
</tr>
<tr>
<td>Local Chapter Funds</td>
<td>127,871</td>
<td>123,212</td>
</tr>
<tr>
<td>Special Interest Group Funds</td>
<td>5,020</td>
<td>5,600</td>
</tr>
<tr>
<td>Special Interest Group Funds</td>
<td>48,819</td>
<td>41,605</td>
</tr>
<tr>
<td>Chapter Project Funds</td>
<td>3,300</td>
<td>3,300</td>
</tr>
<tr>
<td>SIG Digital Scholars Fund</td>
<td>11,773</td>
<td>7,700</td>
</tr>
<tr>
<td>History for Information Science Fund</td>
<td>18,415</td>
<td>16,415</td>
</tr>
<tr>
<td>Scholarship Fund</td>
<td>4,518</td>
<td>4,618</td>
</tr>
<tr>
<td>New Initiatives Fund</td>
<td>10,029</td>
<td>10,029</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>477,443</strong></td>
<td><strong>469,662</strong></td>
</tr>
</tbody>
</table>

NOTE 6. COMMITMENTS

Leases: The Association is obligated under non-cancelable leases agreement for office space which will expire in the year 2015. The amounts due under the office leases are subject to increases based on the greater of a fixed 4% annual escalation or on a percentage of the change in the Consumer Price Index.

The minimum future rental commitments through the remaining term of the leases are as follows:

<table>
<thead>
<tr>
<th>Year</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>$42,183</td>
</tr>
<tr>
<td>2014</td>
<td>43,870</td>
</tr>
<tr>
<td>2015</td>
<td>45,624</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>131,677</strong></td>
</tr>
</tbody>
</table>

Rent expense paid for all operating leases was $40,569 and $39,030 for the years ended September 30, 2012 and 2011, respectively.

See Independent Auditors' Report.
Health Informatics: Introduction
by Prudence Dalrymple, Special Section Guest Editor

EDITOR’S SUMMARY
This special section explores informatics in the health arena through a collection of illustrative examples. The evolution of controlled vocabularies customized for nursing is described, with appreciation of the importance of nurses’ notes in patient records for later retrieval and analysis. The specialized field of public health informatics is presented, offering valuable perspectives and applications geared to serve entire populations rather than individual patients. DOME, an e-health project in Sweden, explores the effects of patients’ access to their own medical records, looking at the pros, cons and implications of online availability of personal health data. Approaching pharmaceutical development as a big data challenge complemented by electronic visualization is shown to facilitate analysis of the complex interplay of numerous variables. Last, a demonstration of information technology applied to emergency healthcare simulation shows how a proposed new technology can be tried out without expensive and time-consuming prototypes.

KEYWORDS
informatics nurses
medical science demographics
biomedical information public policy
medical records simulation
controlled vocabularies pharmacology

This issue of the Bulletin contains a suite of articles grounded in the emerging field of health informatics. While an earlier issue of the Bulletin contained a definition and introduction to the field (www.asis.org/Bulletin/Jun-11/JunJul11_Dalrymple.pdf), this issue’s articles illustrate specific aspects of this large domain having to do with human health and disease, its etiology and treatment. Clearly, such an undertaking is challenging in scope and complexity and is well beyond the resources of a single issue of the Bulletin. Thus, this collection of articles is indicative, rather than exhaustive. Its intent is to pique the interest of the Bulletin’s readers, to provide illustrative and informative examples of the diversity of approaches within health informatics, to describe a few of the challenges in the field and to indicate where appropriate relationships exist among information sciences, information technology and informatics.

The issue opens with a discussion by Patricia Schwirian of the historical evolution of controlled vocabularies in nursing. Nursing care is frequently invisible, included as part of organizational overhead. In the past, nursing notes have frequently been stripped from records before they are archived for storage. Yet it is estimated that well over three-quarters of a nurse’s time is spent in documenting care. Nurses are taught: “If it isn’t documented, it isn’t done.” As more and more healthcare organizations adopt electronic health records, the opportunity to represent and document the care delivered by nurses moves from potential to reality. Controlled vocabularies in nursing (referred to as standardized nursing terminologies or SNTs) have been included in the UMLS (Unified Medical Language System) and are now often included in commercial EHR systems, thus enabling the extraction and analysis of nursing activities. Schwirian, one of nursing’s informatics pioneers discusses what this change means for the development of nursing informatics and the nursing profession.
Mary White’s introduction to public health informatics defines and describes the application of informatics at the population level. She distinguishes it from the application of informatics techniques in managing the care of individual patients, noting that in public health the focus is on the health of populations, including the social problems associated with the burdens of chronic illness. Awareness of public health issues is relevant to all members of society, whether or not they are health professionals, particularly in an increasingly global and interconnected world. She offers suggestions on how readers may learn about this intriguing field.

The role of the patient in health informatics has increased rapidly in the past few years. The term *e-health* is often used to denote consumer and patient perspective, whether it refers to the personal health record or to access to the medical record created by the health team during the course of care. Once almost impossible to obtain without a lengthy wait and without paying for copying charges, one’s own medical record in electronic form is potentially only a few clicks away. Isto Huvila and his colleagues describe the early phases of DOME, a Swedish initiative that is studying a current program to provide patients access to their electronic health records. He provides a balanced view of the benefits and challenges of this program, noting that the findings of the research team currently studying this initiative will have far-reaching implications.

The issue concludes with articles that illustrate two different research approaches. In the first, Timothy Schultz presents examples of the application of big data technology across the pharmaceutical life cycle – from genomics and drug development through clinical monitoring, and finally pharmacovigilance, the detection, assessment and prevention of the adverse effects of drugs. In his explanation, he also cautions that the simple brute force application of computational power is insufficient to solve the complex problems in health care, and he argues for greater attention to data quality on the input side and data visualization on the output side. While Schultz’s article provides insight into big data applications in the pharmaceutical domain, clearly data analytics of all types are central to advancing our understanding of many aspects of health, particularly in an era when full adoption of electronic health records is becoming a reality.

In the second, Diane Sonnenwald describes a novel technique called *visioning studies*, which facilitate early-stage understanding of the potential impact of future technology in complex contexts, such as emergency medicine. A visioning study consists of two complementary techniques. One investigates task performance using an experimental design involving task simulation, observation, questionnaires and interviews. The other explores domain implications using a qualitative design, including video depicting the technology vision and semi-structured interviews. She describes how this technique can inform the design of new technology, enhance its adoption and reduce unintended negative consequences. It can also uncover potential conflicts with current social structures, facilitating the identification of enhancements to social structures and/or practices needed to derive additional benefits from the technology. This article discusses a visioning study focused on 3D telepresence technology in emergency health care.

In an era in which the acquisition, processing, retrieval and analysis of data, information and knowledge are as essential to optimal treatments as pharmaceuticals and X-rays, I hope that you will find these articles not only informative, but a springboard toward your own observation of health informatics applications in action.
Informatics and the Future of Nursing: Harnessing the Power of Standardized Nursing Terminology
by Patricia M. Schwirian

EDITOR’S SUMMARY
Rolled into general health services, an invisible part of a hospital room rate, nursing is overdue for its own professional identity. One hallmark of a true profession is universal use of a specialized vocabulary by members. Standardized nursing terminologies provide a consistent basis for communicating the unique contributions of nursing to direct patient care and to research. Nursing has the dubious advantage of having seven recognized term sets serving numerous purposes but undermining the goal of consistency, key to widespread recognition and acceptance. Adoption of a single vocabulary appears doubtful, though two stand out for active support, currency and free availability. Widespread use of a standardized nursing terminology will come with time, consistent use, incorporation into nursing education and reference in professional healthcare and nursing journals. Cross mappings to important medical terminologies are necessary for interoperability and use in health and nursing informatics research.

KEYWORDS
nurses
controlled vocabularies
interdisciplinarity
interoperability
medical informatics

M ost people agree that nursing is a profession. Others, however, would argue that it is still evolving because it lacks some of the hallmarks of a “true” profession. For example, a “true” profession entails the right to practice independently. Nursing isn’t there yet. While states vary considerably in the extent to which a nurse’s practice must be supervised by a physician, supervision in some form is still the case. Another defining characteristic of a “true” profession is a specialized vocabulary that is used and understood universally by its members. Professions such as medicine and law have specialized vocabularies of long standing.

A significant challenge that nursing has always faced is how to differentiate nursing’s contributions to patient care from those of medicine – particularly in the hospital environment. With the increasing importance of the electronic health record (EHR) and the Nationwide Health Information Network of which the EHR is the foundation, the challenge becomes even more important. Rutherford [1] has stated it well: “Practicing nurses need to know why it is important to document care using standardized nursing languages, especially as more and more organizations are moving to electronic documentation (ED) and the use of electronic health records. In fact, it is impossible for medicine, nursing or any health care-related discipline to implement the use of ED without having a standardized language or vocabulary to describe the key components of the care process.” Another colleague and I recently wrote, “We need to identify nursing’s contribution to patient care and determine the best way to incorporate elements of nursing care into the EHR and other documentation that describes and evaluates the quality of patient care. Fortunately, the answer is already at hand; it lies in the use of standardized terminologies that reflect the uniqueness of nursing care. A key element in differentiation between the purposes and practices of nursing and medicine lies in the professional languages that are used and understood within and between these disciplines” [2].
Using standardized nursing terminologies (SNTs) has several important—and necessary—benefits. They include better communication among nurses and other healthcare providers; increased visibility of nursing interventions and resultant patient outcomes; improved patient care; greater adherence to standards of care; and furthering the nursing research agenda by generating data about patient care in a consistent manner [2].

It was exactly this last element—furthering the nursing research agenda—that prompted Harriet H. Werley, a nursing leader of incredible intellect, dedication and foresight, to take the lead in what was to become nursing informatics in the very early 1980s. As she pointed out, “Nurses must learn to take advantage of the advances made in the information and computer age and recognize that the opportunities and technology are there for them to conduct research and develop their computerized nursing information systems. Only then can the nursing care documentation and related nursing administrative data be collected, retrieved and manipulated in systematic ways to advance nursing practice…as well as the profession.”[3, p. xix]

Standardization of the language that nurses use to document care and outcomes had to be an integral element of this effort.

A factor that Werley did not address, but one that has been a sore point for nurses for many years is “being part of the room rent.” When one goes to a hospital for diagnosis, treatment and care of a disease, disorder or injury, one expects to pay for the services received from physicians. Typically, however, nurses perform most of the direct care that a patient receives. One never expects to get a bill for those services. Yet, as Moss and Saba [4] have pointed out, “The cost of nursing care constitutes a high proportion of healthcare expenses. . . . The current practice of billing for inpatient healthcare services is based on the outdated and often impractical accounting system from the 1930s wherein hospitals bill patients at a fixed daily room rate and the costs of nursing services are included within this per-diem change. . . . Thus, nursing costs are included in overall department and organizational summaries and are not itemized, devaluing the critically important role of the nurse in the delivery of direct and indirect patient care” [4, p. TC149].

In addition to Harriet Werley, another major player in the evolution of standardized nursing terminologies was Norma M. Lang. She was a key figure in the conceptualization and development of nursing diagnosis—a concept that arose in the mid 1970s. Lang built on the work of Marjory Gordon who, in 1976, offered conceptual and structural definitions of nursing diagnosis. According to Gordon [5], a nursing diagnosis has three essential components: the health problem (P), the etiology (E) and the defining characteristics or signs and symptoms (S). The development of a nursing diagnosis taxonomy and its use in nursing practice, many argued, was a means of defining nursing practice both inside and outside the profession and of contributing to the growth of nursing autonomy. Led by Norma Lang, this line of work ultimately produced a set of diagnostic statements that were accepted by the North American Nursing Diagnosis Association (NANDA). NANDA diagnoses were developed using a Delphi-like system in which experts provided input to the development and refinement of each statement. There was—and still is—considerable criticism directed toward the NANDA taxonomy, but it still holds a significant place in the world of SNT’s.

A Standardized Nursing Terminology or the Standardized Nursing Terminologies?

The development of SNTs proceeded rapidly through the late 1980s and into the 1990s. Nursing is a very broad field. Most nurses practice in hospitals, but many engage in practice—home health, office nursing, community nursing, occupational nursing. Moreover, there are subspecialties under each of these areas, so it is difficult to find one terminology that will meet the needs of all these groups of nurses. Accordingly, different terminologies have been developed that reflect different kinds of practice and incorporate varying conceptual foundations. Many argue that progress would be made more rapidly if nurses would simply agree on one best or most-inclusive SNT, thereby allowing nurses to present a more unified front in their efforts to define nursing practice in a unified, understandable manner. That agreement has not come to pass.

Instead, nursing’s most powerful voice and standard-setter—the American Nurses’ Association (ANA) has defined a set of ANA-Recognized Terminologies. Carol Bickford, a senior policy fellow in ANA’s Department of Practice and Policy, has stated, “Nursing is both blessed and challenged
by the wealth of terminologies available for describing nursing practice and nurse’s contributions to healthcare. This diversity offers practitioners choices in how to best describe their patient population and practice.” [6, p. 182-183]. The Committee for Nursing Practice Information Infrastructure (CNPII) provides oversight for the recognition program. The recognized terminologies include seven interface (or point-of-care) terminologies and three multidisciplinary terminologies.

1. **Clinical Care Classification (CCC).** The CCC is a research-based, empirically developed terminology that started with a community focus, but has evolved under the direction of Virginia Saba at Georgetown University into a comprehensive terminology designed to be applicable to all settings. It includes nursing diagnoses, interventions and outcomes. The CCC was designed to be computerized. Some of the others are not.

2. **Omaha System.** The Omaha System was also designed from the beginning to provide a structure that was computer compatible. Like the CCC it is also research-based, and under the leadership of Karen Martin at the Omaha Visiting Nurses Association “…was developed, tested and revised through the efforts of public health and home care nurses in Omaha, Nebraska and community health agencies in seven other geographical areas.” [7, p. 297]. The Omaha System includes an assessment component, an intervention component and an outcomes component. While originating in community and home care, it has evolved over time and is currently applicable in a wide variety of nursing care settings, including acute care.

3. **Nursing Intervention Classification (NIC).** The NIC is a “…comprehensive, research-based standardized classification of interventions that nurses perform. It is useful for clinical documentation, communication of care across settings, integration of data across systems and settings, effectiveness research, productivity measurement, competency evaluation, reimbursement and curricular design. The classification includes the interventions that nurses do on behalf of patients, both independent and collaborative interventions, both direct and indirect care.” [8]. Developed by nurse researchers at the University of Iowa, the NIC interventions include both the physiological and the psychosocial. The 554 interventions are grouped into 30 classes and seven domains – physiological: basic; physiological: complex; behavioral; safety; family; health system; and community. Each intervention has a unique number (code). The NIC is continually updated and is mapped into SNOMED (Systemized Nomenclature of Medicine).

4. **Nursing Outcomes Classification (NOC).** NOC was also developed at the University of Iowa. A nursing outcome is defined as “…a measureable individual, family, or community state, behavior or perception that is measured along a continuum and is responsive to nursing interventions. The outcomes are developed for use in all settings and all patient populations. Clinical sites used to test the NOC included tertiary care hospitals, community hospitals, community agencies, nursing centers and a nursing home. . . . The 490 outcomes are grouped into 32 classes and seven domains: functional health, physiologic health, psychosocial health, health knowledge and behavior, perceived health, family health and community health. Each outcome has a unique code number that facilitates its use in computerized clinical information systems [9]. Like its sister program, NIC, it has been mapped into SNOMED.

5. **NANDA International (Nursing Diagnoses, Definitions and Classification).** The historical importance of NANDA as a foundational element of the SNTs has already been mentioned – as has the centrality of its developer and champion, Norma Lang. It is a nursing diagnosis classification developed to describe judgments made by nurses in providing care. These diagnoses are the basis for selecting appropriate nursing interventions and establishing desirable patient outcomes. The terminology is updated every two years.

6. **International Classification for Nursing Practice (ICNP).** This classification is the newest of the SNT development efforts and the only one of the ANA-recognized terminologies that is international in scope. ICNP is the outcome of a proposal to the International Council of National Nurses to develop a standardized vocabulary and classification of nursing phenomena, interventions and outcomes that would be useful in both paper and electronic records. Moreover, it was to be part of a worldwide infrastructure for the purpose of improving healthcare policy and patient care globally. It was to involve a common language to
facilitate communications describing nursing care and stimulate nursing research. Current efforts focus on mapping the ICNP to the ADA-recognized terminologies and SNOMED CT [7].

7. Perioperative Nursing Data Set (PNDS). Members of the American Operating Room Nurses association started the PNDS development in 1993. It describes perioperative nursing diagnoses, interventions and patient outcomes that are specific to the perioperative environment from preadmission until discharge using standardized elements.

Is the Adoption of a Single SNT Likely to Occur?

The simplest answer is “probably not.” It is my opinion that the two single SNTs that are the most likely to have the most widespread use are the CCC and the Omaha system for three primary reasons. First, each has a vigorous champion. Virginia Saba is an outspoken advocate for the CCC, which she developed. She is a very active contributor to the online discussions of the American Nursing Informatics Association (ANIA). ANIA is an organization composed primarily of the nurses who are active in the field of informatics—primarily in hospitals. When questions arise pertaining to technical/structural problems an informatics nurse may have, Dr. Saba is quick to explain how the CCC handles that problem. She is also active in research that evaluates the utility of the CCC in nursing practice. For example, in a 2012 issue of CIN: Computers, Informatics, Nursing (considered by most to be the premier journal in the N I field), she and Jacqueline Moss, a colleague from the University of Alabama at Birmingham, reported on a pilot study that demonstrated the use of the CCC System Costing Method to tease out the actual costs of nursing care [4]. Could this be a ticket to “getting out of the room rent”?

The Omaha System also has an equally dedicated champion in Karen Martin, who directed the development of the Omaha System when she was employed by the Visiting Nurses Association of Omaha, Nebraska. Like Dr. Saba, she too is actively involved in research that demonstrates the value of her system. In 2011 CIN published a paper in which she and colleagues from the University of Minnesota and the University of Pennsylvania discussed how implementation of the Omaha System can contribute to achieving the goals of meaningful use—a critical element in the American Recovery and

Reinvestment Act of 2009. Achieving meaningful use goals has significant financial implications for eligible professionals and hospitals that use certified electronic health record technology [9]. Dr. Martin is also a tireless presenter of Omaha System workshops and programs for users at all levels of expertise.

Beyond the dedicated efforts of Saba and Martin, the CCC and Omaha systems have another advantage: they are both in the public domain and they are free! Finally, both systems are religiously kept up to date—a “must” if a system is to survive in the highly competitive healthcare marketplace.

Are We “Harnessing the Power”?

In my opinion, the best answer is “We are getting there.” Progress is being made, but in a fragmented way. For example, the NANDA typology has been a regular part of baccalaureate nursing programs for at least the past 10 years, so students are familiar with the notion of a standardized terminology and structured language. The younger generation of nurses therefore has that experience as a piece of their education. However, most older nurses have no familiarity with NANDA or any of the more recent structured languages [2] [11]. How are practicing nurses going to become familiar with SNTs? Linda Thede and I raised that question in our second national survey of nurses. While the workshops and orientations provided by vendors of EHRs as part of the installation package are often their only source, many are learning about SNTs from the professional literature and continuing education. However, the lack of follow-up education available in their everyday practice makes it difficult for most nurses to make SNTs a part of their everyday practice [12]. Provision of ongoing, follow-up education and support is a must.

Another indicator of progress is the increasing frequency with which articles about the value of SNTs and research reports describing their effectiveness in achieving patient care goals are appearing in the respected professional healthcare journals focused on informatics such as CIN: Computers, Informatics, Nursing and JAMIA (Journal of the American Association of Medical Informatics). The nursing specialty journals are also picking up on SNT articles applicable to their own areas of practice.

In order for the informatics goal of full interoperability across nursing information systems to be achieved, two steps must take place. The first is
that nurses must be supported consistently in their use of SNTs; the second is that cross-mappings among the various terminologies must be completed. Although all the major nursing terminologies have been included in the Unified Medical Language System (UMLS) developed and distributed by the National Library of Medicine (NLM), research is needed to map these terminologies fully and accurately. Fortunately, work in this area is underway, suggesting that the importance of capturing and manipulating data regarding nursing effectiveness is now much more widely recognized [13] [14].

Finally, in my role as an educator (I teach an online graduate nursing informatics course and almost all of my students are also employed full time in hospitals) I’m seeing progress. One of the discussion questions I present has to do with the extent to which they have encountered SNTs in their professional practice and what they think about the impact of SNTs on nursing practice and patient care. Three or four years ago their responses were mostly a blank slate. That situation has changed dramatically for the better in the past year as vendors are incorporating SNTs in varied forms and combinations into the nursing-care related elements of the EHR. Sometimes they tell me that until they read about SNTs in my class they didn’t know that they were working with them, but now they do. That might be a backdoor approach to learning and appreciating SNTs, but I am encouraged.

Now, about getting out of that room rent!
Public Health Informatics: An Invitation to the Field
by Mary White

From the advent of Web 2.0 to recent policy recommendations in the Affordable Care Act, change is creating a rapidly growing demand for public health informatics specialists as well as public health professionals with proficiency in informatics. Information science professionals may find in public health a new and exciting arena to apply their analytical and systems skills. Public health professionals may find increased informatics knowledge and skills helpful in their work. This article gives an overview of public health informatics (PHI) and describes a variety of resources to help those interested in finding out more.

Brief Introduction to the Field

While many are familiar with the more clinically driven specialties such as medical or nursing informatics, PHI is a relatively new field. But what exactly is PHI, and how does it distinguish itself in the health informatics world?

Broadly stated, PHI can be defined as “the systematic application of information and computer science and technology to public health practice, research and learning.” [1, p. 68] Regardless of the health domain, all informatics subspecialties apply the informatics pyramid, the relationship and transformation of data, information and knowledge, to making decisions and solving problems [2]. In contrast to a health IT professional working specifically on troubleshooting technology or infrastructure issues, a PHI specialist leverages information and computer science to support public health goals and decision-making and defines the “how and why” science behind the technological tool or approach [3].

As a profession and a discipline, in general, public health focuses on “population and society’s role in monitoring and achieving good health and
quality of life.” [4] That is, in contrast to the more clinically oriented health (and informatics) disciplines, public health focuses on the health of populations versus that of individuals and on prevention in lieu of treatment. It addresses vulnerable points in the causal chain of health problems and considers governmental/policy contexts [1].

As a field, public health is inherently interdisciplinary and collaborative. PHI is no different, applying a range of disciplines, including information science, engineering, law and the social sciences, to public health issues and processes [3]. Therefore, public health issues and responses draw a diverse group of skill sets to the table, and information scientists are especially well suited to this call.

Information scientists and data analysts may find this branch of informatics especially appealing because it naturally deals with gathering, managing and analyzing large datasets. Practitioners in this field also have much in common with information scientists who are interested in the relationships among information, communication and behavior change. Public health professionals may find a way to enhance their work through cutting-edge applications as well as traditional systems methodology.

While initially focused on disease surveillance and epidemiology, opportunities for applying informatics skills to public health have expanded with the development of information and communication technologies, changes in policy and creative approaches to information and communication needs as well as public health interventions. Examples of the variety of these opportunities and roles were evident in the recent American Medical Informatics Association (AMIA) fall 2012 conference pre-symposium event entitled “Current Issues in Population Health Informatics for Healthcare and Public Health.” Late-breaking topics ranged from information exchange (among clinical care, public health and health departments), disease registries, Web 2.0 technologies and public health, privacy and ethical implications and the impact of government policy on public health and informatics. Presenters were from various educational backgrounds and institutional affiliations and were drawn from the federal government, private healthcare industry, local and state health departments, non-profits and academic institutions.

The various health informatics subspecialty areas are distinguished from one another by the content and principles of their domains (that is, whether “medical” or “bio” or “public health”), as well as through the “differing natures and challenges of their informatics applications”. [1, p. 69] These points are useful in framing differences between PHI and other informatics discipline approaches. The recent policy implications of the HITECH/Affordable Care Act (ACA) offer one example. It motivates the adoption of the electronic health record (EHR) through incentives that are based on attesting to the meaningful use of data collected by the EHRs, including for public health purposes, such as immunization or disease registries or surveillance of data to detect disease outbreaks [5]. This program offers opportunities and challenges for both the clinical and PHI approach. Whereas clinical informatics specialties may focus on in-hospital EHR adoption and necessary specifications important to patient care, public health entities may consider how to access and meaningfully analyze and apply such data to population-level health questions and problems. But both involve an application of health informatics principles – including to such issues as system design, system selection and the ability of the systems to connect and share data through standards and interoperability.

The HITECH/ACA is but one specific, timely example of the potential of PHI. The true range of opportunities offered through PHI is as varied as the topics studied and addressed by the field of public health and as varied as the methods and approaches offered by the field of informatics. Regardless of career role or industry type and location, there are many opportunities to put PHI principles to work: an epidemiologist at the state or local level can learn how to connect to disease outbreak information via a health information exchange (HIE); a health communication specialist at the federal level can design smartphone apps to address chronic disease; an IT administrator at a private health institution may consider the privacy implications of putting health data in the cloud; an academic librarian can develop a web portal to deliver timely information to disaster responders. In short, whether you might wish to focus solely on a career as a PHI specialist, apply PHI principles to your work or apply your skills to public health problems, there is a place for you.
Training and Opportunities

A variety of educational experiences and opportunities in PHI are available for professionals and students, with a range of time and monetary commitment, from professional associations, webinars/virtual events or short courses to traditional academic programs and training fellowships. Several organizations provide formal and informal training and opportunities to learn more about PHI, including the American Medical Informatics Association (AMIA), the Centers for Disease Control and Prevention (CDC), the National Association of County and City Health Officials (NACCHO), the National Institutes of Health (NIH)/National Library of Medicine (NLM), the Public Health Informatics Institute (PHII) and the Association of Schools of Public Health (ASPH). The following list is not meant to be exhaustive, but selected and timely opportunities are described below.

Online Resources/Webinars/Virtual Events. Many governmental and nongovernmental organizations and professional associations provide webinars and online resources for those who want to “get their feet wet” in PHI without the time or financial constraints of physically attending courses or conferences. They offer an opportunity to learn about the variety of intersections between applications of public health and information science and informatics approaches. AMIA, NACCHO [7] and the CDC are just a few offering PHI webinars as well as archived access to some of those events.

During the spring of 2013, the CDC National Prevention Information Network (NPIN) hosted a series of free, live webinars under the theme “In the Know: Social Media for Public Health.” The webinars covered topics including the technology, applications and evaluation of a variety of social media channels, ranging from Twitter to YouTube. Archived versions of these webinars and slides are available at the CDC NPIN website (www.cdcnpin.org/scripts/features/feature_itk.asp).

One exciting, upcoming opportunity, which requires no geographic or financial commitment from the participant, is a national PHI web conference. The CDC, the National Association of County & City Health Officials (NACCHO) and Association of State Health Officials (ASTHO) are sponsoring the free Public Health Informatics Virtual Event, July 16-18, 2013. This year’s theme is “Strengthening Public Health – Health Care Collaboration” and topics will include policy and practice, research and innovation and evidence-based practice. Conference information is available at www.cdc.gov/virtual/PHIVirtualEvent2013.html.

Short Courses. The Public Health Informatics Institute (PHII), a program of the Task Force for Global Health, offers training and an Informatics Academy to advance their mission of improving “health outcomes worldwide by transforming health practitioners’ ability to apply information effectively.” During the summer of 2013, the PHII Informatics Academy is offering an online, 8-week short course in conjunction with the University of North Carolina at Chapel Hill. The course, “Designing and Managing Public Health Information Systems: 8 Steps to Success,” will cover the information technology (IT) lifecycle. Find out more about the short course or other opportunities from PHII at www.phii.org/.

AMIA also offers a variety of informatics virtual courses through the 10x10 program, whose goal is to train 10,000 health professionals in 10 years. These courses are conducted in collaboration with partner academic institutions across the country and are offered with a variety of start dates across the calendar year. While the courses are offered online, participants have the opportunity for intensive in-person sessions, typically aligned with AMIA or other professional meetings. Interested participants may sign up for individual courses, typically lasting three to four months. Course dates and topics may vary throughout the calendar year. Find out more at www.amia.org/education/10x10-courses.

Although not public health specific, through the Woods Hole Medical Informatics course, the National Library of Medicine (NLM) fully sponsors health or information professionals in a week-long intensive survey course on the application of computer technologies and information science in the health sciences. The National Library of Medicine also offers grant opportunities for a variety of informatics training, with a range of possibilities including conference grants, early career development,
research-specific training support and small business grants. Find out more about the Woods Hole course and these additional informatics funding opportunities at [www.nlm.nih.gov/grants.html](http://www.nlm.nih.gov/grants.html).

**Professional Associations.** Professional associations provide an opportunity for professional networking and continuing education through association meetings, and their websites are often good places to start to get to know more about current hot topics. The American Public Health Association (APHA) and the American Medical Informatics Association (AMIA) are at the forefront of the public health and informatics fields. Both associations have specialty groups bringing together those with a shared interest in public health, information technology and informatics. After exploring these larger organizations, you may find additional conference or networking opportunities as you gravitate toward specific concentrations or topics, for example at an mHealth summit ([www.mhealthsummit.org](http://www.mhealthsummit.org)).

The American Public Health Association (APHA) offers members access to the “oldest, largest and most diverse organization of public health professionals in the world.” Through the Health Informatics Information Technology (HIIT) section, the vision is to “promote, enhance public awareness and formulate policies on best application and/or methods of information technology and informatics for use in public health” ([www.apha.org/membergroups/sections/aphasections/hiiit](http://www.apha.org/membergroups/sections/aphasections/hiiit)).

The American Medical Informatics Association (AMIA) is composed of professionals from a variety of informatics subspecialties and provides a forum for those interested in informatics to develop working relationships with each other. By joining AMIA’s 500-member Public Health Informatics Working Group (PHI-WG) ([www.amia.org/programs/working-groups/public-health-informatics](http://www.amia.org/programs/working-groups/public-health-informatics)), you have access to listserv discussions, as well as a network of colleagues at the forefront of this field. The AMIA PHI-WG also hosts educational opportunities, such as webinars on a variety of PHI topics, as well as the “Current Issues” pre-symposium event mentioned above. Those interested in PHI, as well as presentations from the pre-symposium, are encouraged to review the proceedings available through the PHI-WG website.

**Traditional Academic Programs and Training Fellowships.** Finally, traditional academic programs and training fellowships offer formal ways for those interested to gain in-depth training, skills and experience in PHI. Academic institutions offer a variety of levels of training and focus areas in health informatics, including individual courses, graduate certificates, masters and doctoral training, both online and face-to-face. Post-graduation, several fellowships are available to augment the training and experience of informatics and public health professionals.

Some academic informatics programs may take a general approach to health informatics or specialize in specific subspecialties such as biomedical/clinical, bioinformatics, translational and PHI. While there is currently no accreditation process, selected institutions offer a public health informatics-specific focus. Conduct an Internet search or visit the AMIA website for a list of many health informatics and PHI training programs ([www.amia.org/education/programs-and-courses](http://www.amia.org/education/programs-and-courses)).

The National Library of Medicine (NLM) supports pre-doctoral, post-doctoral and short-term (three months) funded informatics research training at selected academic institutions across the country. Fourteen “University-based Biomedical Informatics Research Training Programs” cover up to four areas of emphasis: health care, translational, clinical research and PHI. To find out more about the program, and the institutions focusing on PHI, visit [www.nlm.nih.gov/ep/GrantTrainInstitute.html](http://www.nlm.nih.gov/ep/GrantTrainInstitute.html).

Through the Applied Public Health Informatics Fellowship (APHIF), the CDC, in collaboration with the Council of State and Territorial Epidemiologists (CSTE), the Association of State Health Officials (ASTHO), the Public Health Informatics Institute (PHII) and the Association of Schools of Public Health (ASPH), offers fellows on-the-ground experience at the state or local level. This one-year fellowship is an accelerated training program for recent graduates of informatics (or related) masters or doctoral programs, with an interest in the practice of public health. More information is available at [www.aphif.org](http://www.aphif.org/).

In the two-year CDC Public Health Informatics Fellowship Program (PHIFP), fellows experience an intensive orientation to the CDC, public health and PHI. This orientation is followed by short-term assignments at the CDC.
and with public health partners (domestically or internationally) geared towards solving PHI problems. Graduates will be prepared to assume leadership roles in creating and evaluating PHI projects. Interested applicants should have masters or doctoral level training, informatics, public health and research experience. More information is available at www.cdc.gov/PHIFP.

Conclusion

PHI holds opportunities for those who wish to become informatics specialists or those who would like to more fully apply informatics skills in their own public health work. Many public health organizations, including the Association of Schools of Public Health (ASPH), the Public Health Foundation (PHF) and the Association of State Health Officials (ASTHO) [4] [7] [8], among others, have begun defining and highlighting the importance of informatics competencies for public health professionals. Opportunities exist at many levels for interested students and professionals to gain more exposure to and proficiency in PHI. I welcome your interest, involvement and leadership in this exciting and growing field and invite your suggestions on other training and professional opportunities in PHI.

Acknowledgements

I would like to give special thanks to Rachel Graber for her support with this manuscript.

Resources Mentioned in the Article

Empowerment or Anxiety? Research on Deployment of Online Medical E-health Services in Sweden

by Isto Huvila, Gunilla Myreteg and Åsa Cajander

As early e-health efforts progress, it is critical to step back to study their effectiveness and inform expanded implementation. The DOME project is a Swedish research initiative to examine the methods and implications of providing patients access to their own medical records and other e-health services. The focus is on SUSTAINS, a patient access system established in 2012 in 11 European countries. Sweden’s Uppsala county opened access to residents to view their electronic health records with provider notes, lab results, diagnoses and treatments in late 2012. Pre- and post-implementation studies are being carried out through DOME and mixing with the public debate on benefits and drawbacks of easy online access to personal health information. Supporters and critics share concern for quality of care and data security. But advocates focus on better decisions by empowered patients, while detractors, including many healthcare providers, see patient anxiety resulting from insufficient information and lack of consultation. The DOME project is providing the opportunity to analyze all sides of expanded access to electronic health records.

KEYWORDS
health information personal information psychological aspects medical records information access Sweden

Isto Huvila is an associate professor in the department of archival, library and museum studies, Uppsala University, and in the School of Business and Economics, Åbo Akademi University; he is also a research scientist, DOME. His primary research area is information and knowledge management. He can be reached at Isto.Huvila<at>abm.uu.se.

Gunilla Myreteg is a member of the faculty in the department of business studies, Uppsala University, and is a research scientist, DOME. Her specialty is management accounting. She can be reached at Gunilla.Myreteg<at>fek.uu.se.

Åsa Cajander is a member of the department of information technology, Uppsala University; project manager, DOME. Her principal area is HCI and software development methods. She can be reached at Asa.Cajander<at>it.uu.se.

Development and deployment of e-health services and patient access to medical information is currently a priority all around the world. This attention is due to the increase in healthcare expenditures that is being experienced in the industrialized countries. In Sweden the first attempts to implement online access to medical records for patients were made in the 1990s [1]. However, the process has only recently accelerated due to necessary changes in legislation and also through the launch of extensive regional and national initiatives. In 2012 a major Swedish-lead, EU-funded project, SUSTAINS (Support USers To Access INformation and Services), was launched to implement a patient access system in 11 European countries. Earlier experience with such initiatives has shown that an effective deployment of relevant and successful services needs to be based on a thorough understanding of the technical, societal and organizational contexts of these implementations [2] [3]. In spite of the large number of studies in the field, the dynamics of the different factors is not fully understood. In this respect, SUSTAINS represents a unique national and international context of study for researchers to develop urgently needed theoretical and directly applicable knowledge, as well as guidelines for informing future implementations.

This paper will briefly present the DOME project, a major Swedish research initiative, that was founded to exploit the research potential of SUSTAINS and simultaneously to address the broad research needs in the field of e-health. The aim of DOME is to gather systematic evidence on the national and European-wide implementation of online patient access to electronic health records and e-health services. The article will also highlight some early findings from the project and offer some observations on the public discussion that was started in Sweden by the implementation of the online service.
The DOME Project

DOME is a collaborative research project among Lund University, the University of Skövde and Uppsala University (all in Sweden), consisting of researchers representing different fields of scholarship including, for example, management accounting, information technology, information and knowledge management and medicine. The purpose of the project is to create and disseminate knowledge about implementation and use of e-health services based on studies of on-going implementation projects. During the initial phase of DOME, the research focus has been on SUSTAINS and the deployment of e-health services in the Swedish counties of Uppland and Norrland and comparative research in the regions of Skåne and Västgötaland. Due to the complexity of the research context and the large number of stakeholders across society in the field of e-health, the practical relevance of the research findings is considered to be critical for the success of the project.

In order to ensure that the research is firmly anchored in practice, the DOME project has invested much effort in developing an effective network of reference groups. Currently, 25 national and regional organizations participate at different levels in the consultative activities. One goal of the reference groups is to help the DOME researchers identify, understand and highlight the practical problems and possibilities with e-health services from different perspectives. Another goal is to feed back the results of the research project to the national development of e-health services. A specific practical goal of the DOME project is to develop a stable and empirically informed base for the next phase: a national implementation of e-health services in the County Council of Uppsala (LUL). This project is to establish a stable and empirically informed base for the next phase: a national implementation of e-health services.

Online Access to Medical Records

In the autumn of 2012 LUL gave all patients in the county access to their personal electronic medical records, together with a number of other e-health services, online. LUL is the first county in Sweden to offer this e-health service. The online access to medical records enables patients in LUL to access and read their electronic health records (EHR) containing information such as appointment bookings, medical notes, drug prescriptions, medical lab results, diagnoses, referrals and log lists with names of medical staff who have accessed the records. For testing purposes, a minor launch limited to employees at the county was made in August 2012. This test phase was followed by a complete launch covering all patients in the county of Uppsala. In April 2013 over 18,000 patients out of the population of 200,000 inhabitants had read their EHR in a total of 67,102 sessions.

Empowerment or Anxiety?

The early research studies of the DOME project have so far focused on implementation of online access to medical records and development of related new e-health services in the County Council of Uppsala (LUL) in Sweden. This implementation is a central part of the SUSTAINS project. The pre- and post-implementation studies in the county of Uppsala have been complemented with comparative baseline research in other parts of the country. In the next phase the aim is to follow the implementations in other county councils in Sweden and, in collaboration with international research partners, in other European countries.

Not entirely unexpectedly, considering the findings published in earlier literature [5], online access to medical records has led to a debate in media. Proponents of the system regard it as a means to change the role of the patient in the healthcare system. The proponents emphasize potentially positive outcomes of patient empowerment. Whereas such interested parties as politicians and managers in healthcare organizations tend to be primarily positive toward the idea of online access to the EHRs, there is a notable
divide among the healthcare professionals. Some professionals have been very positive toward the initiative, but at the same time many have expressed concerns regarding implementation of the EHR. Even if most critical voices in the public debate have belonged to professionals, findings of the DOME project show some patients also share the anxiety. At the moment it seems that in general, the opinions of the patients seem to vary the most. Some patients are indeed very positive about the possibilities offered by the EHR and the opportunities to have a say. At the same time, other patients are not sure whether they would benefit at all from consulting their medical records, and some are concerned with the security of the service.

An interesting aspect of the controversy in the media is that the proponents as well as the critics of the online access to the EHRs refer to largely similar arguments. They both express a concern for quality of care and security, for example. Those in favor of allowing citizens to access their medical records online regard the consequent empowerment of patients and their possibilities to take charge of their health as a central means to improve the quality of care. Informed patients are assumed to be better equipped to make decisions about their health care, treatments and lifestyle. Empowerment has also been suggested to lead to a deeper engagement in personal health care and to a more meaningful implementation of informed consent. The patient’s access to the record is also expected to yield more accurate information in the EHR, since the patient can spot mistakes and make them known to the doctor. In the long run, the changing role of the patients is expected to lead to healthier lives of citizens as well as a more effective use of healthcare resources.

The critics have expressed their concerns regarding, for instance, how some patients might become anxious after reading about preliminary diagnoses without having someone to talk to or how some individuals might misunderstand the medical record and take actions that even might compromise their health and safety. Other critical arguments relate to the possibility of an unauthorized or ineligible access to, or misuse of, data. All of these issues are said to have a potentially negative effect on the quality of health care. At the same time, many professionals have expressed their concerns about the current implementation process and how the implications of the new system for the occupational safety and work environment of healthcare professionals were neglected. The anticipated problems in working conditions would also, according to them, lead to a poorer quality of care. In order to reduce the risk of patient anxiety it has been suggested, for instance, that doctors should have sufficient time to consult a colleague and complete notes before new entries in records become accessible to patients. There are also patient and staff safety concerns related to providing access to the medical records of some psychiatric patients, which is an issue that was treated also in the earlier literature [6].

At present, the initial hypothesis of the DOME project finds strong support; there is great need for multi- and interdisciplinary research on the deployment of e-health services. In order to develop and understand the current and future processes in e-health, it is necessary to apply a simultaneous consideration of how technical, professional, political and patient-related concerns are linked. The DOME research has already shown this linkage to be an important factor for the success of the initiatives. In the next stage of the DOME research the aim is to track the implementations in other county councils in Sweden as well as in other European countries. The DOME researchers, through the European-wide SUSTAINS project and its national deployment projects in Sweden, have been given a unique opportunity to analyze the contingencies, opportunities, threats and implications related to the development and implementation of online access to EHRs and other e-health services in projects that are both large in scale and embedded in national and international contexts.

**DOME project**

More information about the DOME project can be found at [www.it.uu.se/research/hci/dome/index.php?lang=1](http://www.it.uu.se/research/hci/dome/index.php?lang=1).
Resources Mentioned in the Article


The quest for efficient data acquisition, processing and consumption methodologies has been a topic of critical interest for decades across enterprise and academic institutions alike. In order for organizations to remain agile, novel insights must be derived and decisions made in a relatively short amount of time – often in light of limited observations and sparse data. With the relatively inexpensive cost of computational power and data storage and with a society more interconnected through powerful mobile devices, increased access to invaluable data is beginning to be realized. However, an unintended consequence to these developments is managing and processing massive amounts of data securely and efficiently. These challenges are computationally and storage intensive in nature and are further complicated by an increasing emphasis on fault tolerance, redundancy and scalability. As a popular example, Facebook must continuously deal with the demand of hosting a massive number of social interactions per day, resulting in 500 TB of data. In order for Facebook to suggest relevant content to its projected 1.11 billion members in a timely manner based on past activity, computational resources must be seamlessly incorporated into its infrastructure to address increasing usage demands and a growing user base.

Big Data Overview

The term big data has become a buzzword in the field of information technology. It epitomizes the challenges just described in the broadest of terms. In the past, solutions to such challenges have focused on massive dedicated mainframe computers, distributed computational grids and, more recently, so-called cloud services. The aim of these solutions has been to distribute a computationally intensive workload across a series of dedicated...
resources, such that the overall execution time is reduced as a function of available compute nodes. Popularized by third-party services such as Amazon’s Elastic Compute Cloud (EC2) and Rackspace Cloud, the cloud-based offerings allow organizations to scale technology infrastructure in an agile manner by completely outsourcing and abstracting away hardware layers. As a result, these solutions in principle allow users to add computing resources to their infrastructure in an ad hoc manner. However, more specifically, big data has also become synonymous with Apache’s software stack built on Hadoop, including MapReduce, HDFS, HBase, Hive and Cassandra. Inspired by a series of research papers published by Google to solve distributed, contemporary challenges facing search engines, Hadoop was the first to gain widespread notoriety. This attention is due to its distributed nature and built-in, fault-tolerance features. The software stack allows users to divide complex processes across a number of computational devices whose results are then merged by a central server. Services such as the ones provided by Cloudera allow organizations to easily administer the many working components of a Hadoop cluster through user-friendly, web-based administration tools. While Hadoop is not the only stack that offers distributed computational services, it has gained a great deal of attention for its approach to big data problems.

**Current Applications of Big Data in Health Care**

The most prevalent examples of big data currently reside with social media because user-contributed content provides a robust source of data for analysis. However, the issues discussed above are certainly applicable to the healthcare industry as well [1]. Biological and health-based data are naturally much more complicated and difficult to collect than social-media data. Modeling biological phenomena is typically very complex and has always been understood to be a computationally intensive process. In addition, the Affordable Care Act will undoubtedly continue to transform the healthcare environment in the United States. The ability of pharmaceutical companies to continue bringing new life-saving/life-enhancing medicines to patients in a timely, yet cost-effective, manner will be dependent on their ability to skillfully manage big data generated during all phases of pharmaceutical development. Experience shows that the complexity of such a challenge will be directly related to the complexity of the targeted disease state.

From early research to clinical application and ultimately to pharmacovigilance, the process of delivering therapeutic interventions to patients is an inherently complex, interdisciplinary and interconnected workflow. The pharmaceutical development process consists of many functional research areas. Each contributes a specialized set of insights and knowledge to the underlying mechanisms of a disease or disorder. Not surprisingly, each area also has its own unique set of data challenges. These challenges are further complicated by the many challenges in sharing and translating insights across functions. The big data healthcare problem is based on identifying novel ways to acquire, process and disseminate biomedical data more efficiently. The concept of achieving computational efficiency is certainly not new. It has been the focus of intense research for decades. Nevertheless, the recent focus on challenges associated with big data beyond the scope of healthcare research and development has also...
generated opportunities for closer partnerships between researchers and supporting business functions. Brief examples of big data use-cases sampled across the pharmaceutical development lifecycle include those found in genomics, clinical monitoring and pharmacovigilance.

**Genomics.** One particular area of interest is identifying the relationships between a disease and its genetic, environmental and/or health-based risk factors. Each not only gives a unique view into the underlying mechanisms of diseases and disorders, but also reveals the interplay between different types of risk factors. Identifying risk-based genes is a means to uncover biological pathways for direct therapeutic interventions, while personal risk factors establish corrective interventions, which patients can implement to reduce their risk of developing particular diseases. An example is discussed in [2], where researchers use Bayesian network (BN) learning to untangle the complex relationship between the intrinsic and extrinsic factors that drive bladder cancer. In brief, a Bayesian network is a probabilistic graph model used to expose the conditional dependencies between random variables. To construct the BN structure in this example, variables themselves are operationalized as a set of binary values. Categorical variables can be expressed simply as binary True/False, while continuous variables (such as the level of a specific biomarker) can be similarly expressed by whether the value has exceeded some operational threshold (biomarker positive/negative). Each variable (gene, environmental condition, health condition, existence of biomarker and so forth) is expressed as a node, where directed edges connect those variables that exhibit a directional dependency. Each node is associated with a function that captures the probability of all occurrence combinations with its direct parents. If Node A is connected to Node B, we say that B depends on A or A is a parent of B. Constructing a BN is complex, and adding additional variables to the analysis exponentially increases the quantity of possible comparisons for consideration.

To give an example, researchers applied BN analysis to a set of 1447 single nucleotide polymorphisms (SNP) belonging to genes known to be associated with cancer, smoking history, environmental risk factors such as arsenic exposure and demographic information. Three different approaches were used to learn the structure of the bladder cancer BN, and their results were compared. Overall, it was determined that the three approaches differed slightly in their results, but revealed similar underlying patterns. Interplay between three SNPs in linkage disequilibrium was identified; however, the directionality of these relationships did not always agree between the learning methods. No genetic risk factors were identified to have any relationship with other health factors, but, as the researchers note, this finding is not very surprising since only cancer-risk genes were selected for the analysis. In theory, genetic polymorphisms could show relationships with one’s affinity to smoke, should addiction risk-based genes be selected.

While many conventional BN analyses can be executed on a single central processing unit (CPU), as was done in this case, BN can require scalable processing in the conventional big data sense depending on the size of the data and the size of the directed graph. Bayesian networks, however, have been ported to run in a distributed fashion over Hadoop to accommodate significantly larger graphs [3] than the one constructed for this analysis.

**Clinical Monitoring.** The uses of biosensors, or devices that can detect and measure some physiological activity on or within the body, are quickly becoming an area of intense research. New technology start-up companies...
such as Ginger.io are using the popularity of feature-rich smartphones to monitor the well-being of patients by assessing various metrics of biofeedback, including blood glucose, heart rate and accelerometer monitoring. Understanding how to detect reproducible patterns in signals acquired from biosensors suggests a non-invasive way of learning underlying physiological processes. Ginger.io allows for real-time capture of biometrics to track health-based outcomes for patients by securely monitoring biometrics through a smartphone. In recent years, many types of research-grade instrumentation which were once available to organizations with large budgets, such as EEG headsets and stress sensors, are now more commercialized and obtainable by research enthusiasts. Open source APIs allow researchers to directly access the devices on both desktops and mobile devices, such that novel applications can be devised. As the form factors of these sensors continue to shrink and become less cumbersome when placed on the body, as well as when integrated directly into clothing fabrics, access to large amount of continuous streaming biometric data for analysis increases.

The most attractive area of biosensor analysis can be found in pattern detection, by linking the behavior of biosignals to known phenomenon that occur within the body. For example, researchers use data acquired from accelerometers to capture the movement of subjects along a three-dimensional axis. By finding patterns in the accelerometer data, researchers can predict what activities the subjects are currently partaking in, such as running, walking or typing, and develop metrics for daily energy expenditure [4], although the accuracy of such predictions greatly diminishes outside a controlled laboratory. The possibility of such prediction suggests using accelerometer data from a mobile device to detect early and subtle signs of Parkinson’s in at-risk patients. As another example, researchers have used EEG headsets to monitor the brain activity of autistic children when carrying out certain tasks, such as facial and emotion recognition [5]. These examples require a very large amount of both healthy control data and targeted patient data to establish a “ground truth.” From these pools of subjects, some must be allocated for training for a semi-supervised approach, which is explicitly annotated with the actual actions the person was taking at that time. The uniqueness of each patient, coupled with noise from the sensor itself, can mean that the amount of training data required can be quite substantial.

The capture, indexing and processing of continuously streaming (and possibly annotated) fine-grained temporal data is another big data challenge of increasing interest. The processing of biosignals can include a feature selection step that is executed across metrics calculated within windows of varying and overlapping lengths of time, with resolution ranging from a millisecond to seconds. Understanding the variations within population samples and between comparative populations can be a computationally intensive process, considering the number of subjects required and time-slices under scrutiny. Patient monitoring, assessing health outcomes and understanding the physiological impact a disease has on a patient are just a few examples in which big data analysis has a potential for major impact in this area. Accurate and non-invasive means of inquiring about the inner workings of diseases and disorders can allow researchers to develop innovative ways of applying these insights as future clinical trial endpoints or of remotely alerting caretakers to an impending medical issue. The implications are seemingly limitless.

Pharmacovigilance. The World Health Organization defines pharmacovigilance (PV) as “the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problem.” When drugs are evaluated for efficacy and safety in large Phase 3 clinical trials, studies are conducted in controlled experiment states in which inclusion and exclusion criteria are enforced. However, once drugs are approved and available on the market, a great deal of monitoring by the pharmacovigilance function must be done to ensure that drugs are performing as expected outside the clinic and in the commercial environment. All reports of adverse events (AE) experienced by patients are taken very seriously. They are reported to the FDA and captured in their Adverse Event Reporting System (FAERS) database. The FDA publicly releases reports of these incidents on a quarterly basis with identifying personal information removed. Challenges associated with the understanding of AEs in the commercial setting are often complicated by the fact that patients may take drugs in combinations beyond those included
in well-controlled clinical trials. Since clinical trials may exclude patients who are taking specific medications concomitantly, insights into drug-drug interactions may not be entirely understood at time of commercialization. One solution to this challenge is to analyze the millions of adverse event reports each year to determine whether a particular set of drugs, taken together in a regimen, may be responsible. Expanding drug regimens into all possible two-way, three-way and four-way combinations suggests a dataset of potential interaction combinations in the order of trillions of possibilities. The goal is to identify drug combinations of interest that result in adverse events that occur more than would be expected.

One example depicting a solution to this problem is explained in [6], where an analysis similar to that described above of FDA data is processed over Hadoop. Provided source code contains data manipulation procedures to first merge plain-text CSV data together on a per-subject basis, expand drug-reaction possibilities and express results in the form of a contingency table. When this logic is executed on a Hadoop cluster, the computation of the final contingency table can be distributed across computational nodes using MapReduce and reassembled on a central server. Overall, researchers demonstrated how results could be expressed and visualized as a graph, where nodes represented a particular drug, and a weighted, undirected connection between them represented the strength of a possible interaction. Clusters formed in the resulting graph showed tightly interconnected groups of similar drugs such as those used to combat HIV and various types of cancers. However, they also note the need to ultimately include qualified clinicians in the analysis process to help interpret results.

Lastly, while the monitoring of dangerous drug-drug interactions in clinical practice is an important activity, another interesting area takes the opposite approach in identifying drugs with synergistic efficacy. In other words, identifying comorbid conditions common in a particular disease can suggest other therapeutic interventions that further alleviate secondary symptoms. Disentangling symptoms directly attributed to the mechanism of a disease state from those caused as a consequence is an incredible challenge. A recent investigatory trial has since suggested that depression medication not be used in dementia patients [7].

**Barriers to Entry and Areas of Opportunity**

The daunting task of managing big data will frequently require pharmaceutical information scientists to develop innovative approaches to address complex data computation processes. However a number of barriers currently impeding implementation of such strategies must be overcome before these data can be effectively and routinely managed. The time-consuming process of data collection commensurate with the scope of the task (for example, Phase 3 clinical trial management) and its secure internal and external dissemination must first be addressed. The quantity of longitudinal and/or cross-sectional observations needed to be deemed “big data worthy” often results in complications from both a temporal and logistical perspective. Comprehensive, multi-site research collaborations continue to expand in the era of translational medicine. One such collaboration is the Alzheimer’s Disease Neuroimaging Initiative (ADNI), which has yielded many great insights into the etiology of Alzheimer’s disease over the past eight years. Yet during this time very few big data applications have been generated. But this paucity may not be due to a lack of effort. Relying on a conventional brute force big data approach to understanding an insidious disease such as Alzheimer’s by following patients in the clinic over the course of the disease in its entirety may simply not be feasible. However this challenge is certainly not unique to Alzheimer’s disease. Researchers exploring cures for other poorly understood medical conditions such as autism face a similar dilemma. When, due to the nature of a particular disease or condition, larger quantities of data are required to improve the statistical power of a clinical study in a highly regulated environment, researchers must continue to develop novel ways of challenging old data analysis and modeling paradigms in light of these limitations [8]. In this way, current challenges may be turned into future opportunities.

While some of the most full-featured electronic medical record (EMR) systems may have some big data applications in the form of detailed plain-text clinician notes and medication information, these setups are in the minority [9]. Along those lines, obtaining insights from disparate data sources and incorporating them into the same analysis workflow is vital.
However such integration is still an issue when lack of data-standard adherence is prevalent. In addition, legacy clinical data trapped within flat files must be seamlessly integrated into research workflows. Data curation and interoperability are two areas that will assist in the timely discovery and dissemination of data for analysis. By enabling researchers to locate useful data, quickly assess applicable endpoints and understand how studies were constructed, more robust and integrated datasets can be created. Using domain-specific ontologies such as the Neuroscience Information Framework (NIF) [10] and a general descriptive-study ontology as discussed in [11] to annotate independent variables can facilitate this process, making big data use-cases more viable from a data standpoint.

Lastly, while the content thus far has focused on the quantitative aspects of big data, one must also note qualitative considerations. The area of visual analytics presents an area of synergistic research with big data by conceptualizing the output of complex processes through intuitive graphical means. Metrics dashboarding, real-time interactive visualization and giga-node graph exploration are some examples that would serve as appropriate visualization solutions to the big data examples discussed above. By enabling researchers to scrutinize visual representations of solutions, latent patterns in the data can be identified through quantitative means. Since the application of big data solutions proposes comprehensive workflows that convert unstructured data into analysis-ready datasets, consideration of the structure of the end-data models is vital for the visualization process. An example would be a large weighted adjacency matrix to express the structure of a graph network. Generalizable data models simplify the visualization workflow process as well. By allowing data to be seamlessly piped across specialized visualization applications, researchers can gain insights of varying dimensions into the data.

**Conclusion**

The domain of healthcare research poses an incredibly challenging set of problems that require the synergistic insights of clinical research and information technology. By understanding the complex nature of acquiring, processing and maintaining health-based data along the entirety of the pharmaceutical development lifecycle, novel technologies and statistical applications can be devised. In light of current legislation such as the Affordable Care Act and mounting medical challenges such as the Alzheimer’s epidemic, the solutions discussed above are more pertinent now than ever before. However, while big data solutions are an attractive area for researchers, many data integration and interoperability issues still impede their widespread implementation. Regardless, the concepts of big data, as has been shown here, have very practical applications for aiding in the development of future life-saving therapeutic interventions.

---

**Resources Mentioned in the Article**


### Resources Mentioned in the Article, continued

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
</table>
Visioning Studies in Emergency Health Care to Support R&D of New Technologies
by Diane H. Sonnenwald

EDITOR’S SUMMARY
Given the risk and cost of proposed but unproven technologies, a method to project likely outcomes offers great value. Visioning studies deliver that opportunity by enabling potential users and other stakeholders the chance to explore new technologies. Goals include validating benefits, exposing unanticipated consequences, discovering hurdles to deployment and revealing considerations for funding. A visioning study was used to investigate 3D telepresence technology in emergency health care, enabling paramedics to work on a simulated trauma victim along with a remote physician. In another phase of the study, three groups of stakeholders viewed a brief video on the proposed technology and were interviewed to understand their thoughts on advantages, obstacles and potential fit. Through the visioning study, researchers were able to explore implementation of the technology and experience its potential benefits and challenges in a realistic test mode. Knowledge gained through visioning can lead to better-informed decisions and more effective use of limited resources.

KEYWORDS
medical science
medical informatics
research and development
simulation
evaluation

Significant resources are spent each year globally to support research and development efforts, and an ultimate goal of this research is to create new technologies that will improve lives and/or reduce costs. Funding agencies, foundations, citizen groups and businesses are often looking for evidence that the proposed technology will indeed accomplish one or both of those goals. Providing such evidence is especially important when the proposed new technology will require large amounts of intellectual and financial resources to be developed and deployed and/or is targeted for deployment in life and death situations, such as those that emerge in emergency health care.

But how can we know if a new technology will fulfill its promise without spending resources to create a working prototype and then to evaluate the prototype in meaningful ways? In some instances creating a prototype will require multiple years of effort and a fairly large amount of resources, especially if the prototype requires advances in multiple technological domains such as telecommunications, computer vision and computer graphics.

Furthermore, we know that investigating and understanding the perspectives of users and other stakeholders can lead to the discovery of new features that will enhance the adoption and use of a technology. The earlier these features are discovered and incorporated into the R&D process, the lower the costs of development overall because the cost to introduce and/or modify features usually increases as the R&D cycle progresses.

Thus the challenge is to investigate the potential of a new technology as early in its research and development cycle as possible to validate its proposed benefits; discover unintended consequences and unintended benefits; identify deployment challenges and solutions; provide insights regarding technology features and help inform funding decisions. To address this challenge we have been developing a new research approach called visioning studies.
A visioning study is conducted in collaboration with computer science and engineering researchers and stakeholders in relevant domains. It can consist of two complementary components. One component investigates task performance, comparing current task performance with future task performance as envisioned with the proposed technology. To do this task component an experiment incorporating simulation, observation, questionnaires and interviews is conducted.

The other component explores implications of the proposed technology in the larger domain context. It employs qualitative methods and includes showing each participant a video depicting the technology vision followed by open-ended interview questions seeking participants’ perspectives on issues that influence adoption and use of technology such as relative advantages, compatibility and complexity of the technology, the impact of social influence on adoption of the technology and best practices with respect to deployment processes [1] [2].

We conducted a visioning study to investigate the potential of 3D telepresence technology in emergency health care [3] [4] [5] [6]. An experiment investigating task performance and a qualitative study exploring implications of the technology for the emergency health care were conducted.

Trauma, i.e., serious physical injury, is a significant health problem, frequently referred to as the “hidden epidemic of modern society” because it is responsible for more productive years lost than heart disease, cancer and stroke combined [7] [8]. Today paramedics in the field collaborate with remote physicians in complex emergency situations via cellphone or radio. Results of medical tests, such as electrocardiograms (ECGs), may also be transmitted from the paramedic’s equipment to a remote physician. Future technology, such as 3D telepresence technology, could potentially be utilized for collaboration between paramedics in the field and physicians located elsewhere. Ideally, 3D telepresence could improve patient care through earlier diagnosis and treatment in complex trauma cases [9]. Research on 3D telepresence technology could lead to a system that presents physicians a seamless 3D view of a remote trauma scene, with the view dynamically changing as the physician walks around or changes position with respect to the 3D view. There would be no need for navigation tools that require user input or manipulation such as zooming or changing the angles of the cameras. A digital laser pointer could be developed that would be controlled by the physician and projected at the remote scene, enabling the physician to virtually point at objects in the remote scene. To understand the potential of 3D telepresence technology in emergency health care we conducted a visioning study.

We first conducted a post-test-between-subjects experiment to investigate task performance [5] [6]. In the experiment we compared current, near-term and future task performance. This was achieved simulating a complex emergency medical situation in which study participants, i.e., practicing paramedics, had to diagnose and treat a trauma victim. The trauma victim was in fact a METI human patient simulator, that is, a sophisticated computerized mannequin. Each study participant was asked to diagnose and treat the victim under one of three conditions: working alone (as currently typically occurs), collaborating with a remote physician using state-of-the-art videoconferencing (a near-term vision for paramedic-physician collaboration) and collaborating with a physician using a 3D proxy or surrogate (the long-term vision).

The 3D proxy was simple in its design, yet true to the vision of 3D telepresence technology. In the 3D proxy condition the physician was physically present in the same room as the mannequin and paramedic. The physician was allowed to freely move around in the room. However, the physician could not touch anything in the room and could only point to medical equipment and the victim using a laser pointer. This mirrors the 3D telepresence technology vision.

Results indicated that paramedics who participated in the 3D proxy condition performed fewer harmful interventions, reported higher levels of self-efficacy, found the information provided by the remote physician more useful and of higher quality and showed less variation in task performance times irrespective of years of professional experience in the 3D proxy condition compared to paramedics who collaborated with the remote physician using state-of-the-art 2D teleconferencing. These results are promising; however, technology adoption and use is significantly influenced by the interplay of multiple sociotechnical factors in addition to task performance outcomes [1] [2].
Therefore we conducted a second component of the visioning study that explored factors that may facilitate and/or impede adoption and use of 3D telepresence technology in emergency health care. We first created a 5-minute video illustrating the concept of 3D telepresence technology and its possible use in emergency healthcare situations. In the video we also invited the viewers to share their perspectives on the vision with us. We showed the video in interview sessions conducted with three types of individuals: individuals who might use the technology to provide emergency health care, including emergency room (ER) physicians, nurses, interns at large and small medical centers and paramedics; individuals who make decisions regarding technology purchases and adoption and who manage potential users of the technology, including ER and emergency services administrators; and individuals who would potentially support the technology, such as IT operations personnel. We asked these individuals to discuss what they perceived as the relative advantages (if any) of the technology, its potential fit with their current ways of working, features the technology should or should not have, best practices with respect to deploying the technology and so forth.

Results of this component are presented in two recent articles [3] [4]. Overall the study participants reported the technology could lead to improved patient outcomes and major changes in how paramedics perform their work and how they interact with physicians. The technology would make paramedics’ work visible in ways not previously possible. This could lead to increased prestige for the paramedic profession, increased learning, higher levels of trust and additional permissions to perform procedures in emergency situations. However study participants also reported the technology would require additional training, changes to existing financial models used in emergency health care and increased access to physicians. Legal issues and privacy concerns were also raised, as well as ideas regarding technical features.

The visioning study illuminated potential benefits from 3D telepresence technology in emergency health care, as well as technical features and organizational and social changes that appear to be necessary to facilitate the adoption and use of 3D telepresence technology in emergency health care. It uncovered potential synergies and conflicts with current social structures, facilitating the identification of enhancements to social structures and/or practices to derive additional benefits from the technology. This knowledge has influenced technical decisions and should, ideally, increase the technology’s rate of adoption and reduce its unintended negative consequences.

Acknowledgments
My sincere thanks to the study participants; Greg Welch and Henry Fuchs, who lead the research and development of the 3D telepresence technology, for supporting this research; Jim Mahaney for technical support; Greg Welch and Andrew Dally for their work on the visioning video; James E. Manning and Bruce Cairns for providing expertise in emergency health care and participating in this research, including providing assistance in gaining access to study participants; and to Hanna Söderholm for her work in the project. This research was supported by the National Library of Medicine Contract N01-LM-3-3514, 3D: Telepresence for Medical Consultation: Extending Medical Expertise Throughout, Between and Beyond Hospitals.
Resources Mentioned in the Article, continued


Transforming Our Conversation of Information Architecture with Structure
by Nathaniel Davis

EDITOR’S SUMMARY
Since the origin of the concept, information architecture has been viewed as an art and a science, rooted in library science but borrowing from multiple disciplines. Though there are recognized elements, some say it lacks a foundation of consistent internal theory. The central concept of information architecture is structure. Though invisible and often taken for granted, effective structure is the quality that makes websites functional. It rests on navigation, information organization and information relationships and can extend to user experience and spatial representation. Information architecture reflects elements from a number of disciplines but, by aggregating them, is greater than the parts. Advancing information architecture from art to science depends on shared strategies and solutions for website structure.

KEYWORDS
information architecture
information science
analytic models
web sites
user experience

Information architecture has been characterized as both an art and a science. Because there’s more evidence of the former than the latter, the academic and research community is justified in hesitating to give the practice of information architecture more attention.

If you probe the history of information architecture for the web, its foundation appears to be rooted in library science. But you’ll also find a pattern of borrowing methods and models from many other disciplines like architecture and urban planning, linguistics and ethnography, cognition and psychology, to name a few. This history leads many to wonder if the practice of information architecture is anything other than an art of induction for solving problems of architecture and design for the web.

Haverty [1] and Dillon and Turnbull [2] were early proponents of this idea. In particular, Haverty observes how the field’s lack of internal theory requires practitioners to create by way of constructive induction. She summarizes constructive induction as

“…a process for generating a design solution using two intertwined searches. The first search involves identifying the most adequate representational framework for the problem; the second search involves locating the best design solution within the framework and translating it to the problem at hand.”

For instance, in the absence of an internal IA theory, information architecture practitioners often refer to the study of wayfinding as a compatible representational framework for physical places. By equating digital information environments to spatial constructs, practitioners often translate solutions from the practice of wayfinding to improve how users navigate complex websites.
In 2002, despite the speculation on information architecture’s intellectual foundations, Peter Morville and Lou Rosenfeld seeded the field with the notion that navigation, labeling, content organization and search are the field’s central concerns [3]. This view of information architecture is arguably the most widely recognized one that equally frames the scope of the field’s problem space and serves as the basis of most training on the subject. While these concerns are indispensably relevant to the practice, there are more, which makes the framing of IA practice difficult for practitioners and a challenge for academics who want to perform research and plan more adequate courseware.

Certainly, there is one concept that has persisted under the radar for many years with limited exploration. It is littered throughout countless articles, books and papers and is present in the most cited IA practice definitions. It may be the single concept that truly bridges practitioner and academic interests around a central and worthwhile topic. That concept is **structure**.

**Structure: An Idea We Should Be Talking About**

While *information architecture* may be an intriguing phrase, I’ve come to realize that it’s just not exciting to most people. For me, that’s pretty hard to admit. Even in the prime of its popularity (roughly between 1998-2002), the term never had the appeal of terms like *mobile*, *social* or *user experience*. There is a reason. These other concepts are more tangible to grasp and quantify or more easily packaged and sold. *Information architecture* is not an equally self-explicit term and in all practicality falls much lower in our conscious interaction with computing systems.

Take, for instance, the device you are using to consume this article. What do you see? What do you touch? There are more components inside your device than those you interact with on the surface – for a reason. If you had to complete every circuit and process for every logical function yourself, in order to fulfill the desired task, your device would be impossible to use. In context, while mostly hidden from their users, these electronic components, like the display panel, central processing unit and hard drive, provide the structure that supports the general interactive experience of your device. While these components are part of the whole that contributes to making your device appealing, their primary function is to make what is appealing possible. The same is true for information architecture. While there are aspects of information architecture with which users engage, it’s mostly operating under the “hood” of our information interactions, providing the structure that enables a much broader communicative experience in a web environment (Figure 1).

![Structure makes use possible.](image)

*In use, since site structure is not typically a physically tangible expression like visual and interaction design or some consumable content, it’s easy to take the structure afforded by information architecture for granted – and trust me, it is. However, I’ve come to accept this fact.*

*I’ve become comfortable with knowing information architecture is important even if it is not the marketing darling or business buzzword of the day. In fact, without the practice of information architecture – albeit one*
performed by experts and non-experts – and the type of structure that it enables, the World Wide Web as we know it would not be possible.

**Without Structure, Where Would We Be?**

As we all know, the web was founded on the hypertext transfer protocol, invented by Tim Berners-Lee. At the time of its inception, the Internet was a network for archiving documents and computational resources for the academic and science communities. Hypertext and HTML leveraged the broad reach of the Internet and catalyzed distributed knowledge and research by enabling interconnected “webs” of documents. Through the hypertext node and its inherent link within HTML documents, Berners-Lee created the most basic, but by no means least important, form of structure that is essential to any web environment.

We’ve come a long way from simply connecting disparate web pages as part of creating website structure. In the classic sense of information architecture that I practice and independently research, website structure is the whole that consists of the physical and abstract constructs for

- **navigation** – the methods for enabling wayfinding in an information space;
- **information organization** – the expert and non-expert classification of content and objects of all types;
- **information relationship** – the association of content with lower levels of information, concepts, meaning and behaviors [4].

In a more expansive viewpoint held by contemporary practitioners, the field of information architecture is also interested in applying structure that integrates the edges of digital and physical spaces and places as part of a ubiquitous ecology [5] as well as the structure of user experiences [6]. Regardless, even contemporary practitioners must be grounded in the concepts of web structure or their grander efforts of site visioning and pervasive architecture will not hold.

While the interests of IA practice can be wide and possess a disparate set of vocabularies, the concept of website structure rests at its center.

**Parity Between Academics and Practitioners**

Academics have argued that there is no published IA theory or discipline upon which to explore beyond a web-centered process of inductive reasoning, and they are right. At times, one’s practice of information architecture may be informed by a collection of existing disciplines, but only to create a whole that is greater than the sum of its disciplinary parts. That whole is website structure. While, at the moment, IA practice may be known mostly as an art, its potential science and future internal theory lie in how we understand, strategize and find solutions for site structure.

I hope this discussion promotes common interests in research and greater collaboration between practitioners and their academic counterparts in the future.

To follow my research on this and other topics related to information architecture, visit the DSIA Portal of Information Architecture. ■

**Resources Mentioned in the Article**


The advantages of using a controlled vocabulary (such as a taxonomy or thesaurus) in a database or website project can seem mysterious. What does it do? How does it work? And why should I use one? Let’s take a look behind the scenes to find out why utilizing controlled vocabularies is so valuable. Figure 1 shows basic software components of a controlled vocabulary.

First of all, the taxonomy or thesaurus must be in a digital format. It can be kept either as a separate document file (a spreadsheet, for example) or as it exists in a specialized software application (such as a taxonomy

FIGURE 1. Basic software components of a controlled vocabulary

A quick look behind the scenes

Database Management System

- Search thesaurus
- Validate term entry
- Block invalid terms
- Record candidates

Thesaurus tool

- Establish rules for term use
- Suggest indexing terms

Indexing tool

- Validate terms
- Add terms and rules
- Change terms and rules
- Delete terms and rules
management tool). The screenshot in Figure 2 is from the editorial user interface of a thesaurus software application.

The left panel shows the taxonomy in the hierarchical view. This hierarchy organizes terms and concepts in branches. The broadest subjects are located at the top of this hierarchy. Depending on the size of the thesaurus, each of these broad subjects often contains thousands of terms.

Each term has its own intricate set of relationships, which are found in the term record. The right panel shows a term record containing the broader term, narrower terms, status, related terms and other fields such as synonyms, history, scope notes and so forth.

This amounts to quite a bit of information stored as an object. In this example, the taxonomy term-object is *Heating, cooling, and ventilation*. Treating terms as objects is a useful and easy-to-use way to access your taxonomy, as the object is the term along with all its pertinent and related data.

The thesaurus terms are all organized into a hierarchy or other preferred format. But how are these thesaurus terms connected with, for example, a website? Most often, it will be through controlling the terms used as metadata for the objects and pages on the site.

As shown in Figures 3 and 4, metadata can be found by going to any website, choosing View and selecting the *Source view* or *Page Source View* (depending on which Internet browser is used).
When working with a relational database management system (RDBMS), the taxonomy terms are placed in a table somewhere (Figure 5). This table of terms is then related to the primary key or main records; this table will subsequently be linked to the records directly.

Whether using an object system or relational database management system it is vital to have a place to put those terms. Whoever is building or maintaining the database must find a place for them.

In object-oriented code, a very similar kind of model applies (Figure 6). Again, it is extremely important that the data transfers over from the thesaurus to the primary records.

The terms and their connections must be defined in the relational database. In the various relational database models, there are a lot of options for how to carry this out. See, for example, Figure 7.

In the case of an XML-based database system, new text can be input and the system will have a way to suggest the terms automatically and add them to the system (Figure 8).
When looking at the Mediasleuth site, below, for example, the hierarchical list shown on the website is directly connected to the site from the hierarchical list of the original taxonomy (Figure 9).

Oftentimes the narrower terms in a term record become the narrower terms in the search interface (Figure 10); the related terms from the term record may also be posted in the search interface. All of this integration illustrates that there is a fairly direct connection between the original taxonomy and the website, the user interface and the search experience.

Integrating the taxonomy with the content and user interface enhances the findability of the terms on the website or database (Figure 11).

The terms are used as labels in search as well as for tagging the records behind the scenes. Rather than merely having simple terms connected to a webpage, all of the intertwining relationships that define each concept are linked directly to the search.

When the taxonomy terms are attached to the record and loaded into the search system, while using a variation of that same taxonomy on top of the search system, the taxonomy is being used at the same time to search and to tag. Then when the search is being used, the results are vastly improved.

It doesn’t matter whether a relational database management system, MySQL, Lucene, Autonomy or Google is being used as the search software if the taxonomy term is attached to the term record, and the taxonomy terms are placed in the inverted file for search. When choosing a taxonomy term on the user interface, it will go to that inverted index and pull back the appropriate records regardless of the search software.

The Figure 12 workflow diagram shows that it may be necessary to have a lot of raw data placed into a data repository. The taxonomy terms will be added to the records in that repository. That repository could then be stored as an SQL file for e-commerce, in an XIS repository or in a search system. The system may or may not use a presentation layer for performing search. So, from the original repository where the terms have been added to the records, they can also be spun out to all of these different places for storing the records. Use of this feature is not required, but is certainly available, and often times valuable.

Finally, the same set of taxonomy terms and relationships can be inserted in many places on a website. Taxonomies are easily accessible, easily edited, easily stored and easily utilized.
Following the sudden flood of interest in research data, information professionals have played an important role in communicating to the greater research community the advantages, as well as concerns, associated with sharing and reusing research data. The new task of academic librarians is to develop the best practices and methods for how to make data and other valuable forms of research output available to a wider audience. In response to this challenge, UCLA is developing a data registry, which may be one step toward dealing with the data deluge – one that other academic libraries can also explore.

Many different tools are aimed at making digital data and research output discoverable by others. Two significant efforts are institutional repositories and data repositories. While originally designed to store and maintain access to electronic documents, such as article pre-prints or theses and dissertations, many institutional repositories have begun accepting research data as well. In addition, certain disciplines have domain-specific data repositories dedicated to caring for the datasets produced in those fields. Data repositories are a premier tool for discovering data; however, they are usually geared toward the data produced by a single discipline, limiting their scope. Many fields, especially in the humanities, lack the support necessary for managing the digital content they produce.

In light of this need, additional efforts are underway to assist those interested in managing their data. The data registry, for example, has recently gained international recognition as a possible solution. A data registry functions much as a library catalog: it maintains descriptive records of datasets, allowing the end user to discover the data via a surrogate record, rather than containing data themselves. A data registry can also include a link to the location of the data, allowing interested parties to contact the original researcher directly to discuss a possible data exchange. For libraries that are unable to host a data repository, a data registry may be a cheaper and more viable solution to maintaining access to the data produced by campus researchers.

The UCLA library first became involved in such an enterprise in July 2011, when Christine Borgman, professor of information studies, and Todd Grappone of the library received grant funding from the Institute for
Digital Research and Education (IDRE). The goal of the UCLA Data Registry project was to make a data registry that was both general enough to accommodate multidisciplinary data and specific enough to contain the details necessary to ensure data location and usability. The UCLA library saw this project as a first step toward developing more integrated data management tools.

Project personnel interviewed 20 researchers from disparate fields to determine a) if they would be interested in registering data with the UCLA library and b) what kind of data they would be willing to register. Based on interviews conducted between January 2012 and March 2012, project personnel determined that certain kinds of research projects would benefit from the UCLA Data Registry. For scholars with data stored on UCLA servers, a registry page would provide additional exposure to their data. Some researchers welcomed such supplementary methods for making their data available, because publications only reflect a limited amount of the data collected for a project. Interviewees also acknowledged incentives to registering their data, such as fulfilling funding requirements. In addition, they highlighted motivations that extend beyond data management services, such as the opportunity to establish new collaborations.

Humanists, who typically lack the infrastructure to support the discovery of their non-published work, also noted that they would benefit from registering their projects. However, one of the most interesting interview findings was that, while humanists were among the most interested in registering their work, conversations with them almost always broke down when participants were asked to discuss their “data.” For many of these researchers, curating or sharing the “data” is meaningless without the accompanying software and tools that are created by these projects. This contextual need poses a new challenge for the data registry, as well as data management tools in general.

As of the summer of 2012, the UCLA Data Registry was in its early prototyping and user-testing stage of development. The UCLA library is building this tool with hopes that it will link to current or future tools that build on and add value to its services and functions. In many cases, the most effective tools are not stand-alone services, but instead work in conjunction with other tools to function as a suite of services that collectively form an information infrastructure or a “value chain of scholarship” [1]. Links between journal publications and the datasets on which the articles are based allow an interested user to discover either the dataset or the publication first and then easily access the other components in the chain.

In the long run, the UCLA Data Registry may not be the tool researchers need to help make their research discoverable. The speed at which scholarship and scholarly practices change makes it difficult for academic librarians to predict what kinds of tools will truly benefit research. For this reason, it will be important for UCLA and for other academic libraries to continue collaborating with researchers to implement useful services that do not impose standards or force services where they are not helpful.

 resource mentioned in the article