Digitizing the Non-Digital: Creating a Global Context for Events, Artifacts, Ideas, and Information

Deborah L. MacPherson

This paper discusses some of the problems associated with search and digital-rights management in the emerging age of interconnectivity. An open-source system called Context Driven Topologies (CDT) is proposed to create one global context of geography, knowledge domains, and Internet addresses, using centralized spatial databases, geometry, and maps. The same concept can be described by different words, the same image can be interpreted a thousand ways by every viewer, but mathematics is a set of rules to ensure that certain relationships or sequences will be precisely regenerated. Therefore, unlike most of today's digital records, CDTs are based on mathematics first, images second, words last. The aim is to permanently link the highest quality events, artifacts, ideas, and information into one record documenting the quickest paths to the most relevant information for specific data, users, and tasks. A model demonstration project using CDT to organize, search, and place information in new contexts while protecting the authors' intent is also introduced.

Statement of the problem

Human history is composed of original events, artifacts, ideas, and information translated into records that are subject to deciphering and interpretation by future generations (figure 1). It's like putting together a puzzle, except that each person assembling bits and pieces of the same information may end up with a different picture.

We are at a turning point in the history of humanity's collective knowledge and expertise. We need more precise ways to structure questions and more interactive ways to interpret the results. Today, there is nearly unlimited access to online knowledge collections, information services, and research or educational networks to preserve and interpret records in more efficient and creative ways.¹ There is no reason digital archiving and dissemination techniques could not also be used to streamline redundancies between collections, build cross-references more methodically.² Content should be presented and techniques utilized according to orderly specifications. This will help to document work more responsibly, making shared records more correct, interesting, and complete.

The open-source system proposed, Context Driven Topologies (CDT), packs and unpacks ideas and information in themes similar to museum exhibitions using specifications created by each author and network. Data layers are formed by registering unique combinations of geography, knowledge domains, and Internet addresses to create multidimensional shapes showing where data originate, where they belong, and how they relate to similar information over time. The topologies can be manipulated to consolidate and compare multiple sources to identify the most reliable source, block out repetitious or irrelevant background information, and broadcast precise combinations of ideas and information to and from particular places. "Places," in this sense, means geographic region and cultural background, knowledge domain and education level, and all of their corresponding online resources.

Modern information must be searchable on multiple and simultaneous levels.³ Today's searches occur for a number of reasons that did not exist when most current collections, repositories, and publications were created. Digital records have the potential to reach far broader audiences than original events, artifacts, and ideas. Therefore, digitized items and the acts of publishing and referencing over networks could theoretically serve a longer-term and more expanded purpose than most individual collections, repositories, or publications are designed to serve.

There is no shortage of interesting work to look at. We live in a complex world that is just recently being digitized, mapped, analyzed, and broadcast over the Internet in fine detail and compelling overall relationships. Many

HOSTAETDGO	イタムこじる 1 マル何き	IVNONRTEPHWEMEAMMVEO	
PPTREMARCH	LIRONEEIEDI	SRISGRASOSTECALELDOB	
LARAHAHETI	EGPROSSLEVO	CODMECTCRUMARREGALDE	
	ELOPEXOCNES		
	LOAAAOOISEI		
		NELLEOAPXEBASYOVUOFY	
RNCTAFLWAM	TNLIVEMOROS	SCOITOTLNTVEAREEFLAE	
TANNAMNPIN.	NBOESIRAMSJ	APDTFAEMORNLISEDAMEA	
		INGEATEBAECSPENLLAA	
		INRSZEASSSOYLDCMNLIM	
		ALTTEDADIDDEADTLNTOS	
		L SOHCHASHONKI I SDTOOO	
EMLTAMALLS	DCHNIPHNIAF	RGROFDIAUIEAIAERGVND	
SRNTEDSODD	DACTBUETEOA	APHAANLOIERILOTTOORV	
UAVCOERONN	MRISUDSRUEA	ARNOOTRLFIIBOUDAAEO	
		VBOCFTEHOIGAGROSIPT	
		RONNADNOISUFYLOPECRE	
		ENAVTALAGAEMOIWNNMES	
DOONINEAOT	IEUTESPZAVO	OOCOLARNOVELREASTVBR	
INCRILIGOO	LHMSCHRIGAS	AFPHFTOLECHAPHIAUIIS	
		IMPSOOMETLAONEMCPENT	
		000TSCECRWCSAATIAEAF	
		NAHVOOADGASBERLT000C	
		E DVARIOVUCOE CVIOPMEA	
IVTSSRPPEP:	B 米 Y U C A A A I A S	NNOEOIMRACOPNOREGMAN	
VINECHGEMM	OCNHVEDTEAL	L P A V B O E H T A M N A R D R L R A R	
OGAGCNPEIE	OMTAORAFEEL	PEIEGLPSAWRNRVRMOOD	
	IMEVLSIMMVJ		
		IGEMHPTHSRURSOVVEFIE	
		REPNEEAAILSIEOOEOVDE	
		ODTLICHBTHEDRIVKAPP1	
		SRNVMAEIOPUITROIEERV	
EDVLSAUTOA	BVSLSRTCSIE	COBAENPARDNLXTMDRODA	
LEESFANDET	RAHCIOFOXTS	AANAL ROODECMAOVSCAU	
		LOREOSSNIREALAAPVAT	
I CAMTPOSEV		PAIAAARALOEAGTNENTN	
OF BIVYSOLKL	SEITOOVSNUM	IDENHMUAPMRENRCATIOO	
		BECPUPHVITHLISABATLH	
1.5枚户里河农民民民,	AUOBIYREEES	IRNTURTIOPFLOEAAOMSN	

Figure 1. 50 Word Word-Search-Puzzle (Courtesy of Kevin Lightner)

Deborah L. MacPherson (debmacp@gmail.com) is Projects Director, Accuracy&Aesthetics (www.accuracyandaesthetics.com) in Vienna, Virginia.

of these relationships require mathematics, images, and maps to explain them. We need more than keywords to explore and reference all that has been documented, but we have formed the habit of using keywords and machine-based classification schemes. The entire digital world is in a mire of conflicting priorities, funding opportunities, and intellectual quests toward the future. To advance humanity's collective curiosity and knowledge, and to coordinate similar efforts across disciplines and cultures, we need one form of record keeping. One global context to show:

- 1. Where ideas and information begin;
- 2. If the original is non-digital (e.g., an artifact or real world event), and if so, the location where the artifact resides or the time and place of the event; and
- A marking system to keep track of the ways information has been exchanged, reinterpreted, and reused to create a more comprehensive and simplified guide to humanity's collective knowledge and expertise.

Digitizing the non-digital is a concept to address three issues:

- Tools to assemble the bigger pictures needed to document the best paths to the most relevant information in sets rather than retrieving results item by item;
- Placeholders for information that has not been digitized or was never recorded; and
- Distribution to and from specific places according to the ways it is used, the kind of information it is, and the types of people who are able to understand it.

There is currently little distinction between all data that have been collected or exist, versus the data and techniques selected to draw conclusions. There are no tools to differentiate between information under rigorous discussion by a discipline or culture versus random bits and pieces. There is a need to develop the equivalent of interpretive exhibits to instruct and inspire the general public. There is currently no way to herd information into crowded areas to be consolidated, compressed, and prioritized by its relationship to similar ideas and information. Citation patterns are able to show connections or structure-related information.⁴ However, they currently do not show whether the reference is for or against the other work. There are very few big pictures.⁵ There is no way to trace where an idea has led over time. The global context proposed is not like the ancient Library of Alexandria or large-scale contemporary initiatives. The envisioned process looks beyond the quest to digitize or publish every available event, artifact, and idea. It is not about each item itself. It is being able to make sense of the ways the same information can be viewed in different contexts, and being able to construct a reliable process to search and document the results.

Having bigger pictures will allow researchers, curators, and others to see what is missing or decide which archival works should be converted into digital form. We do not have the time, resources, or reasons to digitize every item in every collection. The aim is to gradually identify what the most telling examples are in different areas so someone new to an event, artifact, idea, or information can see it in various contexts and automatically be shown the most compelling or instructive sequences first (figure 2).

A coordinated effort to overlap and see all archives and publications by ranking accuracy and appeal to the public in relationship to all knowledge will make it possible for entirely new lines of inquiry to be established. It will help researchers coordinate work across disciplines. An example of this principle today is the International Virtual Observatory Alliance (IVOA).⁶ IVOA is a coordinated effort by astronomers worldwide to document our universe more efficiently by systematizing their records; showing where they originate; indicating how they were collected; meeting their rigorous mathematical



Figure 2. Photomosaic® Thousands of miniature images of the civil war combine to make one large portrait. (Courtesy of Robert Silvers)

standards; and deciding themselves how and where their records belong in relationship to each other, and which ones are most important. Only astronomers are qualified to do this. The same is true in any area of humanity's specialized knowledge and expertise. The most difficult aspect of creating a global context is accommodating and expressing each area in its unique way as created from within, while still being able to get the most descriptive examples from all areas to fit together in a sensible and appealing overview.

Until digital archives and publications can be deeply searched on a global level using simpler tools and predetermined pathways accessible by anyone, two researchers in different geographic or academic areas may be investigating the same topic from different points of view and will not know it. There is no way to be led to the best Internet resources. Today, as so much information surrounds us, it is hard to believe that common lines of inquiry could be discovered by accident. Context of the place, time, idea, or education level should be able to drive Internet topologies to the most appropriate online resources.

Constructing a reliable and beautiful digital history of all events-both natural and man-made-artifacts, ideas, and information means contributing to and combining a wide range of knowledge, expertise, networks, archives, and tools. Mapping digital knowledge to historical knowledge means arguing about and perfecting an entirely new set of checks and balances. Historical and digital knowledge are different. Historical knowledge is fluid, continuous, and held by traditionally separated cultures and disciplines. Digital knowledge goes everywhere that can be marked and traced by the times and places it was created, captured, and distributed. Trying to visualize what is happening and relating it to working practices and the types of information that came before it is not like tracing the history of the human race back to Adam and Eve or the universe back to the Big Bang, where substantial guesswork beyond our memory or experience is involved. The entire conversion into the networked age is happening before our eyes in less than one generation without the benefit of reflection, careful review, and storytelling. We're collecting everything indiscriminately over and over again while all datasets are rapidly expanding. We need to step back, slow down, and acknowledge that many current digitization and publication methods do not consistently generate reflective or reviewed results that are able to tell a story.

We do not currently have one shared map, context, mathematical record, language, or set of symbols to interpret from different points of view for a variety of purposes over time. We do not currently mark the original versus subsequent interpretations of the same information as an integral component of most digital records. There is no financial support for one single shared storage space to preserve only the highest-resolution, most agreed-upon versions because we may never be able to agree on what they are. Therefore, there is also not one system that can be fine-tuned to discover research and results that may be accidentally overlapping. Instead, unusual approaches get watered down by constrained words designed to fit metadata requirements developed by archivists and engineers rather than the original authors. Links get broken, Web sites are no longer maintained, trends change. There are currently very few feasible ways to pick up on a line of inquiry previously initiated by others without sorting through and regenerating the same information again.⁷ A simplified version of the work needs to be preserved on the network, able to be referenced by others even if they are far away, live in a different time, or are more or less advanced in their ways of thinking.

If digital information is reliable, someone in a remote place or in the future should not need to collect the same information again or unintentionally retrieve out-of-date or duplicate results.

Searches in the public domain should not be boring. They should be as easy to click through as TV channels, with more directions to go and better content. All searchers should not have to start at the top like everyone else on the first page of Google, CiteSeer, or ArXiv with a blank white space and a box to enter key words. Investigators should be able to outline the facts they know, dial in measurements, specify relationships, and generally be able to use their own knowledge and expertise to isolate and extract entire ideas over broad spectrums or select only relevant portions of archives and publications to reintegrate into larger bodies of work for further discussion.

Digital objects are able to depict more than the unaided eye can see. An example is the evaluation of the center of mass of Michelangelo's David performed for David's restoration by the Visual Computing Lab based on a 3D model of the statue built by Stanford University (figure 3).⁸ The digital David does not have mass. The original David is a beautiful object sculpted of a known and predictable material. The model makes it possible to test restoration techniques without permanent damage in ways no one would dare attempt on the irreplaceable original without first knowing more. The documentation process is an enhanced original that should be permanently bound to the digital history of the original sculpture. The evaluation method could be applied to other objects, but this model belongs with this object and this type of research. A global context built upon a solid, mathematically linked foundation would mean this conscientious work would not be lost or need to be repeated.

Digital records are not being used nearly to their full potential. So many influences on humanity's intellectual evolution could be examined as history takes shape over time. Concurrent and conflicting interpretations can take on more meaning than the original by itself. For example, how could the Internet and legal citations be used to map the subsequent interpretations of the U.S. Constitution from the time, place, and reasons where it was written to every Supreme Court case and related citation since the original context? What would this map look like (figure 4)?

The impact that these four pages of ink on paper have had to the United States and the entire world cannot currently be examined in one volume to see where the most contentious and useful passages are. Similar dynamics in Wikipedia are shown in History Flow by Martin Wattenberg at IBM Research.⁹ What if techniques developed in one field could be applied to content from another area? For example, what if computer models created to track storms and hurricanes could be used to arrange and watch the evolution and real world impact of all the documents and actions associated with a war?

Being able to see how originals evolve in their interpretation and impact on society over time is practical because not all records are worth keeping. Even worse,

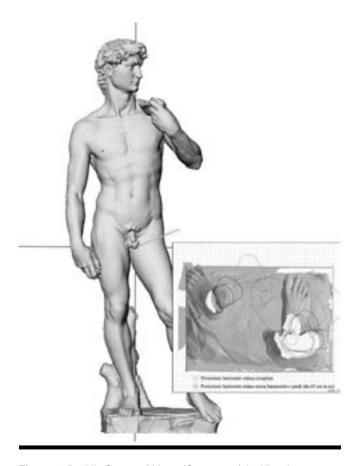


Figure 3. David's Center of Mass (Courtesy of the Visual Computing Lab and Stanford University)

mundane or meaningless events, artifacts, ideas, or information may seem more important than they actually were if they are not translated into digital form or distributed in the right way.¹⁰

The task today is to make the most advanced ways of thinking and working more approachable and appealing to someone new, which is everyone outside a particular discipline or culture, while traversing a map of humanity's collective knowledge and expertise. Because shared memories of this magnitude would be so far-reaching and complex, the record itself needs to be able to show every user how to use it. Every unique purpose for looking around, publishing, or referencing work, and adding to or taking away from a collaborative global context should be geared toward improvement and simplification. While millions and millions of people are accessing enormous numbers of files and collections, some paths are better than others. In order to sort and choose the best parts of vast collections, documenting everyone going in and out of various semantic places can ultimately identify the best paths to information everyone understands. What if someone who does not care at all about paintings makes an inquiry—which ten should they be shown to get them interested? There is also the issue of gearing the Internet to provide more efficient pathways to widely accessed preapproved and curated information.

Every mouse click could accumulate to document the most reliable pathways in and out of shared information spaces to generate an assortment of scenarios for looking at the same information in different ways (figure 5).¹¹

We think there is far too much information to consolidate into one big picture, that our ideas and methods are too incompatible to coexist comfortably in one space, but perhaps this is not really the case. Perhaps we can understand what is happening more clearly by working backwards.

Proposed solution and design for a running prototype

Even though many networks are in place and countless computers have been manufactured, technology advances rapidly. There are very few reasons to repair obsolete equipment or maintain outdated web resources. Therefore, why not go back to the drawing board on all of it? We may have completely new computers and networks within ten years, anyway.

A record-keeping and referencing system this ambitious needs to incorporate every type of record, classification scheme, symbol, style, and quirk. When visiting a new place outside your comfort zone, it needs to be obvious what the best local techniques are to filter and understand the results. People new to an area need to have the option of using tools they can invent or already know.

The Visualization of CDT's model demonstration project will bring together research scientists, artists, integrators, and institutions to develop a running prototype. The purpose is to establish and record a series of planned and spontaneous situations in different parts of the world across a range of disciplines and existing networks so that these situations can be mapped. The project will be a

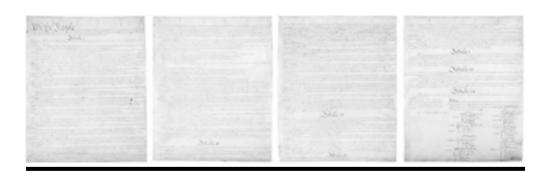


Figure 4. The Constitution of the United States (Courtesy of the U.S. National Archives and Records Administration [NARA])

group of people thinking together to confront the roadblocks in assembling incompatible ideas and information into one context. The group will collaborate in larger and smaller groups in roughly three-month intervals as participants continue with their existing work. The development of this system has to be dynamic, changing piece by piece both from the bottom up and the top down while everyone's regular work continues. Therefore, the system will be geared toward sample sets of active work products, rather than the record-keeping system by itself.

The current objective is to establish a network of ten art museums, ten scientific research institutes, and ten new media/new technology efforts in ten cities that speak different natural languages (for example: English, German, French, Italian, Hindi, Mandarin, Ga [belonging to the cluster of KWA languages in Ghana], Uzbek, Spanish, and Arabic). The overall intent is to use mathematics, art, and individual ways of knowing to develop a series of professional sketches to serve as shortcuts between languages and key words in the search process.

The first step is to map the background of each of the project participants' previous work by time, location, and discipline. The database will include scientific visualizations, art objects, performances, algorithms, mathematical formulae, musical recordings, and many other forms of creative and scholarly expression. The next steps will be to hold a series of interactive workshops. At the first workshop, the research scientists will explain the mathematics and images they use in their work. Two sets of artists will isolate the aesthetics to render their own map through the scientists' ideas. Two traveling exhibits will be created, one to be experienced in person, the other to be presented through a new media and online exhibit. Both will be tracked physically and conceptually using CDT. The results will be generated and interpreted using GIS, MATLAB, Photoshop, and flow visualization software. For more information, please contact the author.

A survey of individual and institutional requirements

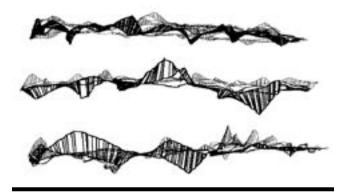


Figure 5. Thick and Thin (Courtesy of the artist John Simon)

will be undertaken to define practical ways to move and organize ideas and information into a unified sample map of previously unrelated content and techniques. For example, at one institute, perhaps only two participants and four local professors will understand what that part of the map is showing. Another part may only have meaning to one artist. A unified map for everyone, with built-in copyright protection for the participating artists, scientists, and institutions, will be presented to nonspecialist general publics around the world for feedback and further change within specified limits. The participating publics will be people interested in contemporary art, cutting-edge scientific research, new media, and events where all three communities can interact.

Each part of the prototype will be able to be examined in groups to compare and contrast different elements against different backgrounds. Some arrangements will be assisted by the computer and network. The project will map everything with which each event, idea, and artifact has ever been associated in scale, proportion, and relative placement in the record overall. For example, if the records in question are paintings, any group could be gathered together into the same reference window without copying the images. The assembly window has a built-in scale for the items it is showing, so they will be displayed in the correct proportion to each other. The system binds images of physical objects with their dimensions and the times and places they were created while this information is known—so a user does not ever have to guess later when looking back at any part of the record. Any group of paintings can be automatically arranged chronologically, by size, culture, or any number of comparisons and curatorial issues. A sample sequence is:

- 1. A zoomed-in map showing a group of paintings in an exhibit. Each painting links to its history.
- 2. Within the map of all paintings shown in an intricate collage.
- 3. Inside the map of all human endeavor shown as an appealing landscape.

Higher levels can then be used to reorganize a theme, for example, "only Germany 2005 to 2007," and drilling back down to generate other exhibitions. This would lead to other paintings and other curators' conclusions, which would provide a more complete representation of each painting, exhibition, museum, curator, culture, and era. When the records in question are scientific visualizations, problems of presenting unrelated files together are more complex. The records may not share a common scale or system of reference. It may only be possible to place mathematical constructs in contexts based on where they originate geographically and by knowledge domain. An important part of the work will be determining the best contexts by which to introduce ideas or information to untrained viewers and devising methods to start deeper in the records using mathematical, cultural, or other prior knowledge and preferences.

The same concept can be described by different words, the same image can be interpreted a thousand ways by every viewer, but mathematics is a set of rules to ensure that certain relationships or sequences will be precisely regenerated. Therefore, unlike most of today's digital records, CDTs are based on mathematics first, images second, words last.

Ideas and information will be encoded to persist over specified periods of time. Better examples will find higher placement by connecting to more background information and showing stronger relationships to larger numbers of open questions. Cycles will be implemented to return to the same idea later and remove information that is never referenced or has not changed the course of the record's flow. Out-of-date, irrelevant, or rarely used information has to either be compressed or be thrown away,

A new type of identity and a process to assemble and eliminate information will be created in thirty prototype forms showing the intertwined history of the events, artifacts, ideas, and information generated by the project and all it branches out to when connecting back to the publications, exhibits, ideas, artifacts, and other information generated by the participating individuals and institutions. The CDT model will relate and join tables to display all the different forms together in one map. Each piece of information and the patterned space around it will be documented a special way to generate drawings leading back to originals reliably structured to transfer to other computers and networks. They will transfer without ambiguity because the transactions and paths to the Internet addresses are based on mathematical relationships that can be checked.

Each contributor has the first opportunity to place his or her ideas in context and define the limits of how their originals can be referenced, changed, and presented. At the end of the project, the set will be closed so that it can be cleaned of information that was only temporary, placeholders can be examined, and the entire model can be manipulated as one whole. For more information, please see www.contextdriventopologies.org

The more specified a single piece or set of information, the easier it will be to define its history and place it in context. Each unique placement and priority assigned by each individual or institution may not agree with the priorities and placements envisioned by others, but sooner or later, there will begin to be correspondence and everyone will be looking very generally at the same emerging map.

Conclusions

There will be innumerable contexts to create, discover, and remark upon in the future by creating a shared pace of curiosity and knowledge acquisition. A global context could be used to extrapolate new knowledge from trends that occur over longer periods of time in more places than we currently share or document. As the envisioned system is fine-tuned, it will become an ideal place to test an idea that is only partially complete to see where the idea fits or to determine if it has already been done. The results could be immediately applied to improve education. In today's frantic information overload, we should not forget that digital information—and even cold, hard, raw data—is more than ones and zeros. They represent peoples' work, their fingerprints; people are attached to their data.

One wishes networks of computers could understand one's ideas and work, but we only show them the boring parts. The proposed system will capture beauty so computers can help to find where it is hidden inside all the repositories, publications, and collections through which no person has the time to sort. The system will allow users to specify how they think their information relates to the rest of the world so their intended context can be traced in the future. One hopes that using networks and computers to compare ideas and works on larger levels will restore craftsmanship and attention spans to make users want to spend more time with better information.

A shared visual language driven by mathematical relationships that can be checked will allow future historians to see where records simply will not harmonize. Users will be able to analyze why different ways of looking can shape and divide knowledge and history as it changes. Visiting online archives and publications will change. Developing processes to pre-organize searches and results for public viewing can change now by creating a system for curators and others to develop sets of information, rather than publishing individual items on their Web sites. Library facilities can change, and research rooms can become multimedia centers. Networks can broadcast content and techniques in one package.

There is not one clearly defined reason why being able to see these kinds of overviews or make these types of comparisons can be useful. The Internet is a worldwide invention being constructed for a variety of purposes. A perfectly legitimate reason to capture the history of transactions across it in a simple form is just to see what might happen with the objective of increasing our understanding and respect for each other. The most important reason for establishing a global context is to allow users to transfer and update complex histories, thoughts, images, studies, visualizations, drawings, flow diagrams, sequences, transformations, cultural objects, stories, expressions, and purely mathematical or dynamic relationships without depending on constrained keywords or illegible codes that do not describe this information as well as the information can describe itself. All cultures and disciplines would be able to construct their parts of the record precisely the way they prefer. We would finally be able to use computers to show why and how we think information is related—a huge leap forward in the world of digital record keeping.

References and notes

1. CiteSeer, 2005, http://citeseer.ist.psu.edu (accessed Apr. 6, 2006); Internet2, 2005, www.internet2.edu (accessed Apr. 6, 2006); Jane's Information Group, 2005, www.janes.com (accessed Apr. 6, 2006); Machine Learning Network Online Information Service (MLnetOiS), 2005, www.mlnet.org (accessed Apr. 6, 2006); National Technical Information Service, 2005, www.ntis .gov (accessed Apr. 6, 2006); Smithsonian Institution Libraries, Galaxy of Knowledge, 2005, www.sil.si.edu/digitalcollections (accessed Apr. 6, 2006); Thompson Scientific, ISI Web of Knowledge, 2005, www.thomsonisi.com (accessed Apr. 6, 2006); Visual Collections, David Rumsey Collections, 2005, www

.davidrumsey.com/collections (accessed Apr. 6, 2006); World Health Organization, Statistical Information System, 2005, www3.who.int/whosis/menu.cfm (accessed Apr. 6, 2006).

2. G. Ammons et al., "Debugging Temporal Specifications with Concept Analysis," in *Proceedings of the ACM SIGPLAN 2003 Conference on Programming Language Design and Implementation* (New York: Association for Computing Machinery, June 2003).

3. W. Huyer and A. Neumaier, "Global Optimization by Multilevel Coordinate Search," in *Global Optimization* 14 (1999): 331–55

4. A. Bagga and B. Baldwin, (Workshop paper), in COLING-ACL '98: 36th Annual Meeting of the Association for Computational Linguisitics and 17th International Conference on Computational Linguisitics, Aug. 10-14, 1998, Montréal, Quebec, Canada: Proceedings of the Conference (New Brunswick. N.J.: ACL; San Francisco: Morgan Kaufmann, 1998); S. Deerwester et al., "Indexing by Latent Semantic Analysis," Journal of the American Society for Information Science 41, no. 6 (1990): 391–07; A. McCallum and B. Wellner, "Toward Conditional Models of Identity Uncertainty with Application to Proper Noun Coreference," in Proceedings of the IJCAI Workshop on Information Integration on the Web (Mountain View, Calif: Research Institute for Advanced Computer Science, 2003), 79-84; T. Nisonger, "Citation Autobiography: An Investigation of ISI Database Coverage in Determining Author Citedness," College and Research Libraries 65, no. 2 (Mar. 2004): 152-63; K. van Deemter and R. Kibble, "On Coreferring: Coreference in MUC and Related Annotation Schemes," Computational Linguistics 26, no. 4 (Dec. 2000); K. Boyack, "Mapping All of Science and Technology at the Paper Level," presented at the session Mapping Humanity's Knowledge and Expertise in the Digital Domain as part of the 101st Annual Meeting of the Association of American Geographers (Denver: Association of American Geographers, 2005): 54; Metacarta, 2005, www.metacarta.com.

5. J. Burke, "KnowledgeWeb Project, 2005." www.k-web .org (accessed Apr. 6, 2006); Visual Browsing in Web and non-Web Databases, Iowa State University, www.public.iastate .edu/~CYBERSTACKS/BigPic.htm (accessed Apr. 6, 2006).

6. International Virtual Observatory Alliance, 2005, www.ivoa.net (accessed Apr. 6, 2006).

7. S. Bradshaw, "Charting Excursions through the Literature to Manage Knowledge in the Biological Sciences," presented at the session *Mapping Humanity's Knowledge and Expertise in the Digital Domain*, as part of the 101st Annual Meeting of the Association of American Geographers (Denver: Association of American Geographers, 2005): 56, project paper available from http://dollar .biz.uiowa.edu/~sbradsha/Beedance/publications.html (accessed Apr. 6, 2006).

8. M. Callieri et al., "Visualization and 3D Data Processing in the David Restoration," *IEEE Computer Graphics and Applications* 24, no. 2 (Mar./Apr., 2004): 16–21.

9. M. Wattenberg, "History Flow," 2005, http://research web.watson.ibm.com/history (accessed Apr. 6, 2006).

10. K. Börner, "Semantic Association Networks: Using Semantic Web technology to Improve Scholarly Knowledge and Expertise Management," in *Visualizing the Semantic Web, 2nd ed.* Vladmire Geroimenko and Chaomei Chen, eds., (London: Springer Verlag, 2006) 99–115.

11. G. Sidler, A. Scott, and H. Wolf, "Collaborative Browsing in the World Wide Web," in *Proceedings of 8th Joint European Networking Conference, Edinburgh, Scotland* (New York: Elsevier, 1997); J. Thomas, "Meaning and Metadata: Managing Information in a Visual Resource Reference Collection," in *Proceedings of Association for Computers and the Humanities and the Association for Literary and Linguistic Computing Meeting* (Charlottesville, Va.: University of Virginia, 1999); H. Yu and A. Vahdat, "Design and Evaluation of a Conit-based Continuous Consistency Model for Replicated Services," in ACM Transactions on Computer Systems 20, no. 3 (Aug. 2002): 239–82.

12. Visualization of Context Driven Topologies/CDT Model Demonstration Project, 2005, www.contextdriventopologies.org (accessed Apr. 6, 2006).

Image acknowledgments:

50-Word Word-Search Puzzle

www.synthfool.com/puzzle.gif

Permission: Kevin Lightner, Synthesizer Enthusiast. Wrightwood, California

Abraham Lincoln

www.photomosaic.com/samples/large/AbrahamLincoln.jpg

Permission: from the artist Robert Silver.

David's Center of Mass

http://vcg.isti.cnr.it/projects/davidrestoration/restaurodavid.htm

http://graphics.stanford.edu/projects/mich/book/book.html

Permission: Roberto Scopigno, Visual Computing Lab, ISTI-CNR, Via G. Moruzzi, 1, 56124 Pisa Italy and Marc Levoy, Stanford Computer Graphic Lab, Gates Computer Science Bldg. Stanford, CA 94305

U.S. Constitution

www.archives.gov/

Repository: National Archives Building, Washington, D.C.

Permission: NARA government records are in the public domain.

Thick and Thin

www.numeral.com/drawings/plotter/thickandthin.html

1995 11" × 15" Ink on Paper.

Permission: from the artist John Simon, New York City. Specializing in algorithms and conceptual art.