THE DISCOVERY REVOLUTION

George L. Paul and Bruce H. Nearon

E-Discovery Amendments to the Federal Rules of Civil Procedure

ABA
Defending Liberty
Pursuing Justice

THE DISCOVERY REVOLUTION

George L. Paul and Bruce H. Nearon

E-Discovery Amendments to the Federal Rules of Civil Procedure

Cover design by ABA Publishing

08 07 5 4

Library of Congress Cataloging-in-Publication Data
Paul, George (George L.)
The discovery revolution : e-discovery amendments to the Federal rules of civil procedure / George Paul, Bruce Nearon.
p. cm.
Includes index.
ISBN 1-59031-605-3
1. Electronic discovery (Law)—United States. 2. United States. Supreme Court. Federal rules of civil procedure. I. Nearon, Bruce. II. Title.

KF9650.P38 2005 347.73'50285—dc22
2005033641

Discounts are available for books ordered in bulk. Special consideration is given to state bars, CLE programs, and other bar-related organizations. Inquire at Book Publishing, ABA Publishing, American Bar Association, 321 North Clark Street, Chicago, Illinois 60610.

www.ababooks.org

PREFACE

The "E-Discovery amendments" to the Federal Rules of Civil Procedure will have a significant impact on the practice of law. The amendments address the new "information infrastructure" which appeared suddenly in our economy —just a few years ago in fact—and which quickly outdated assumptions about how information behaves. Indeed lawyers, society's specialists in gathering and testing information, will no longer do their job effectively unless they understand the new rules and the reasons they came about.

As one might expect, the amendments were the subject of robust public comment. Seventy-four witnesses testified before the Civil Rules Advisory Committee, and over 254 comments on the new rules are currently found on the "U.S. Courts.gov" website maintained by the Administrative Office of the U.S. Courts. There, one can access both the comments and the testimony about the amendments that occurred in San Francisco, Dallas, and Washington, D.C.

Docket numbers for each comment, and for the testimony of individuals and organizations, are found on the website and cited in the form of "04-CV-001" for the first comment received; "04-CV-002" for the second and so on. In addition, complete transcripts of the hearings can be downloaded and searched.

Many have suggested that this voluminous commentary is a treasure trove. Litigators, corporate counsel and vendors with long, practical experience were compelled to succinctly state the key issues for rule-makers. Various groups testified, from defense counsel and corporations to public interest and plaintiffs' firms. So did bar associations, including the ABA's Litigation Section, the ABA's Section of Science & Technology Law, the American Trial Lawyers Association, the American College of Trial Lawyers, and the Association of Corporate Counsel, among others.

Surveys were conducted by think tanks. Law students chimed in, as did their professors. Forensic specialists were heard. Even U.S. District Court Judges got into the act, including the Federal Magistrate Judges Association. Insightful readers will often be able to predict the stands taken by the various

groups—making the positions of the district judges and the magistrate judges all the more intriguing, not to mention the solutions of the rule-makers.

For these reasons, this book cites liberally to the public commentary. The commentary represents not only a sort of "legislative history," but also contains valuable practice points. Whether specific comments were incorporated into the rules or not, they are instructive and eminently "practical." Sometimes a comment is fascinating precisely because it was *not* incorporated, or, because a rule was changed to avoid a thorny issue implicated in the testimony.

But it would be unfair to write this book exclusively from a legal perspective. One of the authors is a litigator with experience in electronic discovery and trials. The other is a CPA who specializes in auditing information systems, and who knows such systems, including their imperfect human element, intimately. For this reason we have attempted to present a "real world," non-idealized picture of the complexity facing lawyers today. Several experts in their field also contributed to sections on evolving case law, statutory background, and the substantial amount of information processing that now attends to most electronic discovery.

The E-Discovery amendments to the Federal Rules of Civil Procedure were approved by the Standing Committee on Rules of Practice and Procedure at its meeting in Boston on June 15 and 16, 2005, and then by the Judicial Conference of the United States, on September 20, 2005. In December 2006, the new rules will become effective unless disapproved by either the U.S. Supreme Court or by Congress.

Accordingly, practitioners can get a head start by reading about the new rules now. But more, the rules go to the heart of electronic discovery as it is practiced today. The issues addressed in the new rules and discussed in this book are obviously already present in litigation. These include such things as the approach to pre-discovery meetings; the "form" of production of electronic evidence; preservation of evidence when routine processes may destroy it automatically; who pays for relatively "inaccessible evidence" and determining what is reasonably accessible and what is not; sanctions for spoliation; and privilege waiver issues, including "claw back" of materials inadvertently produced. Although not discussed in the rules, the book introduces the related issue of the authenticity of evidence. This is a meaty list of topics that are all immediately relevant as a result of the rapidly evolving way in which Society handles information.

Indeed, with a fundamentally new information infrastructure, we seem to have reached a turning point in the legal profession. Before, lawyers dealt with artifacts of information that were well-organized by file-keeping procedures that made access an easily mastered aspect of the lawyer's trade. Things

are much different now. Whether lawyers can absorb the changes brought on by the complexity of information systems, without becoming overly dependent on technical specialists or perhaps being crushed by costs, is probably *the question* for the future of the profession. One of the themes of this book is that the new complexity demands a new collaboration. The profession must undergo an evolution in its human relationships and attitudes.

Unless lawyers are willing to seek and share information about systems, they will act counterproductively on behalf of their clients. They will be condemning their clients to overly broad preservation orders, or worse, draconian spoliation sanctions. They will simply drown in information, or in the expense of unnecessary privilege reviews. Or, they will be caught off guard by the shifting of costs for the production of information.

The authors wish to acknowledge that this book has been a team effort. A group of lawyers and information technologists, the American Bar Association's "Digital Evidence Project," has worked on this project for some time, as have experts in the fields of electronic discovery, civil procedure, and auditing of information systems. The authors wish to acknowledge in particular the contributions of Gib Sorebo, Ben Wilson, Hoyt Kesterson II, Joel Kazin and Steven Wu of the Digital Evidence Project; Kim Taylor of Encore Lex Solutio; and Robert G. Schaffer and James Belanger of the law firm Lewis and Roca, LLP.

The authors also wish to thank members of the Civil Rules Advisory Committee for their tolerance and for sharing working papers and memos, particularly Special Reporter Professor Richard Marcus for his answers to questions. Judge Ronald Hedges of New Jersey and attorney/forensicist Craig Ball of Houston were kind enough to read the transcript and offer suggestions and improvements. Ken Withers of the Federal Judicial Center was, as usual, supportive. And thanks go to the wonderful people of the Administrative Office of the U.S. Courts, including Judy Krivit.

George Paul and
Bruce Nearon

George L. Paul, J.D., is a partner in the Business Litigation Section of Lewis and Roca LLP.

Bruce H. Nearon, CPA, is the director of the IT Security Auditing practice of J.H. Cohn LLP.

CONTENTS

CHAPTER 1

The Revolution in Discovery

Society stores information in a profoundly different way than it did in 1990. The change has affected commerce, everyday life and culture, and therefore the world of law, which constantly must access and test societal information in its discovery processes.

Indeed, physical information storage reigned supreme for almost the entirety of civilization. As a result, our cultural systems have long incorporated certain assumptions about information: We have assumed that there were original records, with molecules that had been altered to record information. There were also copies of such originals. And there were altered records, which were not to be considered authentic. The basic unit of information storage was usually referred to as a "document," particularly if connoting a more formal informational record. The word dates back to the Latin documentum, meaning "proof."

The modern form of business document can be said to have appeared in the 1870s with the invention and diffusion of the typewriter and its use in conjunction with carbon paper. To deal with the resulting proliferation of paper, new technologies for storing documents emerged in the 1890s. Notable among these were the vertical file cabinet and the Rand card index system, the combination of which allowed large numbers of documents to be stored and accessed efficiently. Records management techniques evolved whereby documents could be systematically organized and retrieved. One could request "Please bring the file" and a set of organized papers would promptly arrive, often facilitated by a clerk whose job was to keep track of things.

The Rise and Consequent Miniaturization of the Computer

But an onslaught of technological change increasingly affected society's information infrastructure after World War II. Change was at first gradual and did not affect the technologies of information storage utilized by most businesses. For this reason, for a long time there was minimal impact on the law.

The first computers were simply not that relevant to the way everyday records were kept by businesses or, critically, for the interpersonal communication of business workers. These were gigantic, expensive machines used only by large institutions such as government agencies, big companies, and universities.

It is important to note for our purposes that such machines did not process in "real time." Instead, workers brought programs and data, which others input into the machines. The processing occurred in batches, and it was not uncommon to receive the results of a computation hours or sometimes even a day after the request to the data processor. This simply would not work for many of today's business applications, as, for example, at a supermarket point-of-sale checkout counter or a Web-based airline reservations system.

During the 1970s and 1980s, however, a multi-tiered revolution turned the initially fanciful idea of a small "personal" computer into reality. Microprocessors, hardware, and software were developed that led to ever-increasing computing power and storage capacity, and to a decrease in size and cost. Suddenly, small personal computers spread into businesses and homes during the 1980s. The technical innovation made it easier to create and store electronic information, and millions were doing it.

The Connectivity Revolution

But there was another aspect to the rising infrastructure, involving the concept of "connectivity." In the early to mid-1980s, distribution of electronic documents was slow and cumbersome. People printed documents on paper for distribution, or physically exchanged "floppy disks." Then, the 1990s saw the rise of corporate intranets, wide-area networks, and the Internet, which collectively served to connect many previously isolated computers. This "connectivity revolution" was made possible by the widespread adoption of com-

munications protocols like TCP/IP,[1] and by the rise of the client-server architecture as the most common design for computer networks.

Today, networks have become extremely complex, often connecting hundreds of servers and tens of thousands of clients. The composition of each network is constantly changing as the hardware and software are adapted to evolving business needs. In addition, computer networks grow "organically." They tend to evolve of their own accord, not by any predetermined design but in bits and pieces effectuated by different individuals, who are ignorant of the others' actions. In addition, over the course of a decade or less, software and hardware products age until they become obsolete in the sense that their manufacturers reduce or stop providing support or updates. In short, components can become "out of sync" with the rest of the evolving organism. As these components age, the data they support becomes increasingly difficult to access.[2]

The New "Digital Realm" of Information

These intertwined technological advances, including "real-time" computing, the personal computer, networks, widespread communication protocols, e-mail, and, of course, the Internet, have combined to create a new information infrastructure for the industrialized world. As a result, there is a new "digital realm" of information storage. And the digital realm differs radically from civilization's long-standing physical information storage systems.

In the physical realm, information is static. It is stored on artifacts of matter, altered in ways to record information. Physical informational records last decades, centuries, even millennia. Consider that the Egyptian pyramids carry information that has endured several thousands of years.

---

1. TCP/IP means Transmission Control Protocol/Internet Protocol, the suite of communications protocols used to connect hosts on the Internet. TCP/IP uses several protocols, the two main ones being TCP and IP. TCP/IP is built into the UNIX operating system and is used by the Internet, making it the de facto standard for transmitting data over networks. Even network operating systems that have their own protocols support TCP/IP. *See* http://www.webopedia.com/TERM/T/TCP_IP.html (last visited June 7, 2005).
2. Some of the details recounting the evolution of computer technology are from Microsoft's comment to the Advisory Committee (04-CV-001).

But in the digital realm, information is not stored by altering matter. Information is stored in a more abstract way by means of machines utilizing *flows* of information, with records stored and processed in multi-layered systems with interacting components. As a result, information is much more dynamic in the digital than in the physical realm. Routine processes can delete or "overwrite" information in the ordinary course, without human intervention. Similarly, a database might update itself several times per second, once again as a result of routine operation, not conscious human decision.

Another distinguishing characteristic of the new realm is that, again because of the machine processes involved, there is much more information now. Organizations have thousands if not tens of thousands of times as much information within their boundaries as they did 20 years ago. We are awash in information.

An example of the new quantity of information has been illustrated by ExxonMobil:

> Exxon generates 5.2 million e-mails daily. . . . Its employees have 65,000 desktop computers and 30,000 laptop computers. The storage capacity of the desktop and laptop computers it is now issuing to employees is 40 gigabytes each. (One gigabyte equates to 500,000 typewritten pages. Forty gigabytes equates to 20 million typewritten pages). . . .
>
> In the U.S., Exxon's disaster recovery systems generate 121,000 back-up tapes a month (1.45 million back-up tapes per year).[3]

But more important, let's examine a small business. In the physical information paradigm, there might have been four to eight file cabinets of paper constituting the business's records. But today, even the simplest system consists of, at minimum, one computer. In 2005, the average hard drive size for a new computer is 40 gigabytes. The minimal operating system and application software use 10 gigabytes (GB).[4] That leaves 30 GB for document storage. If one assumes the

3. *See* comment of ExxonMobil (04-CV-002).
4. Gigabyte—A unit of information equal to one billion (1,073,741,824) bytes or 1,024 megabytes.

average word-processing document is 150 kilobytes (KB),[5] a user could store 200,000 documents on his computer. However, even the smallest of users may have an external hard drive for backup with 250 GB of storage capacity, and a thumb drive[6] that may hold up to 1 GB. Therefore, a user could have another 1.7 million documents on her external drive and another 6,700 documents on her thumb drive. So given just three sources of electronically stored information (sometimes "ESI") on the smallest of systems, a user could have as many as 1.9 *million* documents, or between 20 million and 100 million pages of information. The average four-drawer file cabinet holds about 10,000 sheets of paper.[7] Therefore, the smallest of businesses has the capacity, conservatively, to fill the equivalent of 2,000 four-drawer file cabinets!

It should be clear from these two examples—from the largest corporations to the smallest of computer users—that the increase in societal information has implications that are almost unfathomable. No one seems to be able to devise an effective procedure to control the explosion of information.

But there is more to this than sheer quantity. This abundance is *distributed,* meaning it is not found in a centralized and easily identified location, as is the old document storage culture. The information is "smeared" around and, indeed, is constantly flowing in many different systems operated within the enterprise, and found at any moment in hard-to-account-for "storage devices."

For example, some business information is found on servers, but there may be hundreds of servers journaling all sorts of different transactions. Information is found on client/desktops and also in laptop computers, which are periodically connected to and disconnected from a network. The tiny thumb drives can be used to easily transport information from computer to computer, or simply used as a backup de-

5. Kilobyte—A unit of information equal to one thousand (1,024) bytes.
6. Thumb drive—A small, compact, rewritable storage drive, roughly the size of an adult thumb, that interfaces with a computer via the USB port, usually carried in a pocket or on a lanyard. They are sometimes found in a smaller, more durable format to fit on a keychain. The thumb drive is also known as a keychain drive, micro hard drive, pen drive, pocket drive, jump drive, USB flash drive, USB flash memory drive, USB key, USB memory key, and USB stick.
7. Gary K. Starkweather, *Optical Technology and Paper Management in the 21st Century*, Optics & Photonics News, Sept. 2002, at 29, *available at* http://www.osa-opn.org/abstract.cfm?URI=OPN-13-9-28 (last visited May 30, 2005).

vice. What about PDAs and home computers? The information is moving all around, some as a result of human decision and some as a result of automated processing. Accordingly, because of this distribution, controlling and discovering the information contained in a business is burdened with new challenges. There is not one system. There are many systems in a business.

Affiliated with this distribution is the issue of *outsourcing*. Enterprises increasingly are contracting with outside companies to handle information. Thus, sometimes investigators need to look at outside companies to discover sought-after information. This can provide challenges, since the custodian of the information is not even a client, and may not cooperate enthusiastically, or at all.

The New "Information Complex"

What is apparent is that a new, societal information *complex* has evolved in the past few years. This is an interrelated set of constantly evolving subsystems—technologies that are networked in their operation, with holistic properties greater than the sum of their parts.[8] Critically, with such systems, no one person or even small group of people can explain the operation of the system.[9] It has emergent, self-organizing properties, like that of an ecosystem or economy, that simply defy explanations of the Cartesian paradigm. Indeed, such "complexity," in the technical sense of the word, is the prime reason why there is currently a need for new legal concepts—such as "not reasonably accessible" information and the "safe harbors" you will read about in this book.

Keeping one's eye on the evidentiary side for a moment, because artifacts are now unnecessary, information can be duplicated exactly, seemingly without limit. An unknown number of identical records can exist simultaneously without any having claim to originality. When

---

8. See Robert Rycroft & Don Kash, *The Complexity Challenge: Technological Challenge for the 21st Century* (1999); Mitchell Waldrop, *Complexity: The Emerging Science at the Edge of Order and Chaos* (1992); Ilya Prigogine & Isabelle Stengers, *Order Out of Chaos: Man's New Dialogue with Nature* (1984); and Arthur W. Brian, *Self-Reinforcing Mechanisms in Economics*, *in* THE ECONOMY AS AN EVOLVING COMPLEX SYSTEM: SANTA FE INSTITUTE STUDIES IN THE SCIENCES OF COMPLEXITY, Vol. 5 (Philip Anderson, Kenneth J. Arrow & David Pines eds.1988).
9. Rycroft & Kash (1999), *supra* note 8.

combined with the fact that software is almost always designed to allow users to easily and seamlessly edit files, this stretches the concept of authenticity past the breaking point.[10] The authenticity issue transcends law and now also affects accounting and finance as the nature of digital evidence erodes the foundations of auditing.[11] As far back as the Best Evidence Rule of the early Eighteenth Century,[12] the key to authenticity has been a quest for the "original." Yet this concept of "original" is now absent. The authenticity crisis is relevant to discovery, because it will bear on the "form" of information one might want to request in discovery.

What makes information so radically different now is that lawyers must deal with complex events, such as a portion of a system suddenly becoming disconnected from the larger organism, while other parts continue to grow. We are dealing with systems, whether it is the difficulty of extracting information from the system or simply how to stop it from "doing its thing." The disheartening fact is that usually, no one at the client can give an accurate status report on the beast. Only by approximation can we describe our information systems.

The Impact on the Practice of Law

The new realm of information has had a substantial impact on the practice of law. First, in contrast to earlier times when it was possible to ask for a file and receive it, now our complexity in information means it is often necessary to involve an entirely new set of specialists to even access the relevant information. Will this always be necessary? Hasn't litigation fundamentally changed when specialists are now required to do what lawyers used to do routinely—namely, to collect and analyze run-of-the-mill informational records?

---

10. *See* George L. Paul, *The Authenticity Crisis in Real Evidence*, SCIENTIFIC EVIDENCE REVIEW, Monograph No. 5, Cynthia Cwik & John L. North eds., *reprinted in* PRAC. LITIG., Vol. 16, No. 6 (November 2004), *available at* https://www.ali-aba.org/aliaba/PLIT0411-Paul.pdf (last visited June 7, 2005); Paul, *The Authenticity Crisis: Problems with Electronic Evidence*, *in* CORPORATE COUNSEL (October 2004), *available at* http://www.discoveryresources.org/pdfFiles/TheAuthenticityCrisis.pdf (last visited June 7, 2005).
11. *See* Bruce H. Nearon, *Foundations in Auditing, Skepticism, and Digital Evidence*, CPA J. (January 2005).
12. *See* LORD GEOFFREY GILBERT, LAW OF EVIDENCE (Dublin, S. Cotter) (1754).