2012 Annual Convention

Ethical Issues and Practicing Law in the Digital Age

Legal Ethics and Professional Conduct Committee

1.0 General, 1.0 Ethics, and .5 Substance Abuse CLE Hours

May 2-4, 2012 ♦ Cincinnati
Professor G. Kristen Barnes  
The University of Akron School of Law  
Akron, Ohio  
Professor Barnes received her JD from Harvard Law School and her PhD from Duke University. She is an assistant professor of law at The University of Akron School of Law and teaches in the areas of property, real estate transactions, and comparative law. Prior to entering academia, Professor Barnes practiced commercial real estate law at both large and small firms in Chicago. She has represented public and private businesses, developers, REITs, and lenders in various projects involving municipal development, retail, office buildings, leasing, commercial sales and acquisitions, and financing. For additional information, please visit www.uakron.edu/law.

Jonathan E. Coughlan  
Supreme Court of Ohio Disciplinary Counsel  
Columbus, Ohio  
Mr. Coughlan received his BA from Hobart College and his JD from Case Western Reserve University School of Law. His professional memberships include the American Bar Association, Ohio State Bar Association, Columbus Bar Association, National Organization of Bar Counsel, and the Association of Judicial Disciplinary Counsel. Mr. Coughlan has been Disciplinary Counsel of the Supreme Court of Ohio since 1997. He is a frequent lecturer on attorney and judicial ethics for the Judicial College of the Supreme Court of Ohio, OSBA CLE, and local bar associations. In addition, Mr. Coughlin serves as an instructor at the National College of District Attorneys in South Carolina and is an adjunct professor at The Ohio State University Michael E. Moritz College of Law, where he teaches professional responsibility. For additional information, please visit www.supremecourt.ohio.gov.

William C. Mann  
Mitchell Pencheff Fraley Catalano & Boda Co. LPA  
Columbus, Ohio  
Mr. Mann received his BA from Case Western Reserve University and his JD from the University of Dayton School of Law. His professional memberships include Ohio State Bar Association (Legal Ethics and Professional Conduct Committee); Million Dollar Advocates Forum. Mr. Mann is of counsel to his firm and a Board Certified Civil Trial Advocate by the National Board of Trial Advocacy. He concentrates his practice in the areas of professional defense, including the defense of judges and lawyers in legal ethics matters, plaintiff's personal injury matters, and general civil litigation. Mr. Mann is a frequent lecturer and contributor to seminars and programs on personal injury litigation, civil litigation, and professionalism issues. For additional information, please visit www.mitchell-lawyers.com.

Professor John P. Sahl  
The University of Akron School of Law  
Akron, Ohio  
Professor Sahl received his BA from Boston College, his JD from University of Vermont College of Law, and his LLM from Yale Law School. His professional memberships include the Ohio State Bar Association (Legal Ethics and Professional Conduct Committee), Cleveland Bar Association (Ethics and Professionalism Committee), Scanlon Inn of Court, and American Bar Association Center for Professional Responsibility (Chair, Publications Board; Coordinating Council; Franck Award Selection Committee). Professor Sahl is a Professor of Law and the Deputy Director of the Miller Institute of Professional Responsibility at The University of Akron School of Law. He teaches courses on professional responsibility, evidence, law, morals and professional responsibility, media reflections of lawyers and judges—myths and reality, sports law, and entertainment law. Professor Sahl also teaches a law course each summer at Yale University and a contracts and commercial seminar at the CEIDA Institute in Rome, Italy. He has served as a consultant and an expert witness in professional responsibility cases and has helped to investigate disciplinary
matters for the state bar involving lawyers and judges. Professor Sahl is the author of numerous articles on professional ethics and discipline for lawyers and judges and frequently speaks on those topics both nationally and internationally. For additional information, please visit www.uakron.edu/law.

Donald A. Wochna  
Vestige Digital Investigations  
Medina, Ohio

Mr. Wochna received his BS from California State University, Long Beach, and his JD from the University of Chicago Law School. He is one of the founding partners of his firm and a pioneer in the development and application of computer forensics to the practice of law. Mr. Wochna was a trial lawyer for 22 years when he left a private practice in 1999 to focus on computer forensics and the emerging field of digital investigations. As a member of the Board of Directors of the Institute of Computer Forensic Professionals and the Chief Legal Officer for Vestige Digital Investigations, he has earned a reputation as an engaging and articulate electronic evidence expert in civil discovery, internal investigations, and criminal defense litigation. Mr. Wochna has served as an expert witness in several state and federal court cases and is a widely sought CLE speaker on electronic evidence and the impact on the law and the discovery process of the proliferation of devices that can become “electronic witnesses” in all types of cases. He has educated attorneys and, through the Ohio Judicial College, provided seminars to Ohio common pleas judges in the area of electronic evidence. Mr. Wochna’s formidable knowledge relating to the application of computer forensics is illustrated by his critical testimony before the Judicial Conference of the United States on proposed changes to the Rules of Civil Procedure that affect computer forensics and litigators. At a local level, he has drafted changes to the Ohio Rules of Civil Procedure and is working with the committees of the Ohio State Bar Association to modify these Rules to accommodate discovery of electronically stored information in civil litigation. Mr. Wochna is a frequent speaker on topics related to computer forensics and use within litigation. For additional information, please visit www.vestigeltd.com.
Chapter 1

Donald A. Wochna

What Paradigm Can Be Used to Discuss the Performance of Legal, Law-Related, and Non-Legal Services Performed and Integrated Into a Single Case? .......................... 1.1
What Is the Motivation of Law Firm to Perform Discovery Services? .................................. 1.2
What Are the Ethical Issues Raised by Law Firms Performing E-Discovery Services? .......... 1.5
McDermott Law Suit as an Example of Law Firm Negligently Performing E-Discovery Service—Manual Review of Documents for a Particular Purpose ..................... 1.8
Conclusions ....................................................................................................................... 1.10
Attachments .................................................................................................................... 1.13
J-M Mfg. Co., Inc. v. McDermott Will & Emery—First Amended Complaint for Damages .................................................. 1.15
Opinion and Order ......................................................................................................... 1.27
Electronic Data, Electronic Searching, Inadvertent Production of Privileged Data: A Perfect Storm ................................................................. 1.77

Chapter 2

Social Media and Ethics
Jonathan E. Coughlan & Professor G. Kristen Barnes

Introduction to the Impact and Reach of Social Media ....................................................... 2.1
Social Media in the Courtroom ....................................................................................... 2.1
Social Media: Employers and Employees ...................................................................... 2.3
Social Media and Schools .............................................................................................. 2.3
The Future—The Legal System and Managing Changing Social Media Practices .......... 2.3

Chapter 3

Substance Abuse from the Respondent Lawyer’s Perspective in a Nutshell
William C. Mann

Introduction ....................................................................................................................... 3.1
The Client ......................................................................................................................... 3.2
The Initial Lawyer-Client Meeting ............................................................................... 3.2
What the Client Needs to Know About Substance Abuse ............................................... 3.4
Intervention by Lawyer .................................................................................................. 3.7
The Law of Substance Abuse in a Legal Ethics Case ..................................................... 3.9
To Sum Up ....................................................................................................................... 3.9
The Most Common Mistake Lawyers Make .................................................................. 3.10

Donald A. Wochna
Vestige Digital Investigations
Medina, Ohio

This seminar reflects ongoing research into three areas of technology and law: (a) the technical nature of the protocols and processes needed to identify, preserve, extract, parse, analyze, produce, and authenticate electronically stored information, (b) the ethical and regulatory restrictions impacting e-discovery conducted by law firms as legal services, and (c) the evolution of discovery as a third party, law related, expert service. This research is not complete and the thought leadership contained in this seminar has not reached a firm conclusion. The issues, however, associated with providing electronic discovery as legal services delivered by attorneys are well defined and ripe for discussion. This seminar is an attempt to begin the dialogue whether, in an ever-increasingly complex technological world, legal services ought to include the technical means by which electronically stored information is identified, preserved, extracted, parsed, analyzed, authenticated. It is the author’s suggestion that as electronic devices and networking environments become increasingly complex, it will become impossible for attorneys to competently offer electronic data advice as part of “e-discovery.”

I. What paradigm can be used to discuss the performance of legal, law-related, and non-legal services performed and integrated into a single case?

   i. Offshore outsourcing provides an excellent paradigm by which to analyze e-discovery.

   ii. Offshore outsourcing shares similar characteristics to e-discovery. E-discovery, in fact, is one of the services that has been outsourced to offshore companies.¹

iii. Ethical obligations in the offshore outsourcing context appear to require supervision of all work by a fully-qualified lawyer.  

II. E-Discovery services as law-related services. Are E-discovery services properly characterized as law-related services? What constitutes the practice of law and how is it different than law-related services offered during discovery? What constitutes e-discovery law-related services? Are law-related e-discovery services comparable to other law-related services that are outsourced overseas, such as patent work? Do law firms offer e-discovery services as law-related services; or do firms integrate e-discovery into the “practice of law”? Do e-discovery law related services enjoy protection as attorney-client/work product? Is litigation hold, preservation instructions, search strategies, etc. protected from discovery because they were performed by attorneys? Because they were performed by a law firm under circumstances in which the client reasonably believed the e-discovery services were protected as lawyer services? Does a firm offering e-discovery services as the “practice of law” insulate e-discovery protocols and processes from discovery?

a. The manner in which different law firms have integrated e-discovery into their practice.
   i. Ignore it: Judge from PA comments on attorneys agreeing to ignore e-discovery
   ii. Perform all e-discovery services as “the practice of law”
   iii. Outsource some e-discovery services based upon the expertise needed to complete the services
   iv. Outsource all e-discovery to a “electronic discovery” expert

b. Similar to the manner in which law firms have outsourced libraries, data processing, and copying. Outsourcing has expanded to include legal and law-related services. Outsourcing legal and law-related services implicates ethical and regulatory issues as applied to services that require judgment, nuance and experience in the legal field.

---


i. A common area of outsourcing that involves legal-related judgment includes the preparation of patent applications and document review.

ii. Outsourcing legal-related activities requires (a) defining with particularity the functions that will be performed by a third party; (b) integrating the completed work into to matter by client’s counsel. The third party providing the legal-related activities (including activities that require “legal judgment”) disclaim any relationship with the client that might create an attorney-client relationship between them.

c. How are legal-related functions separated into discrete assignments that can be performed by third parties?

i. Using patent law and outsourcing as an example, the initial steps require breaking down the project into separate tasks, not all of which must be performed by client’s counsel. A patent application typically consists of the following activities: (a) prior art searching, (b) drafting background, (c) drafting specifications, (d) drafting claims, (e) drafting summary, (f) preparing drawings, and (f) a final review, modifications and filing. Although the last activity has to be always performed by a USPTO registered attorney or agent, who usually also becomes the attorney of record, other activities can be either done by the IP professionals in a remote location or by the IP professional located in the U.S.⁵

ii. Based upon the engineering experience needed, each of these steps can be performed by an engineer or another professional (including an attorney). But none of the professionals performing steps a through e would create a circumstance in which they were delivering legal services or engaged in the unauthorized practice of law.

iii. By avoiding the creation of an attorney-client relationship, many non-lawyers offer outsourced activities that are integrated into a matter by the client’s attorney. These include: researching law, preparation of briefs for appellate and Supreme court, and certain documents related to merger and acquisition work.⁶

iv. Ohio Rules of Professional Conduct evidence a schema to protect a client from an assumption that all services performed by the client’s counsel must necessarily be encompassed by the attorney-client privilege.

---


⁶
v. RULE 5.7: RESPONSIBILITIES REGARDING LAW-RELATED SERVICES: (a) A lawyer shall be subject to the Ohio Rules of Professional Conduct with respect to the provision of law-related services, as defined in division (e) of this rule, if the law-related services are provided in either of the following circumstances: (1) by the lawyer in circumstances that are not distinct from the lawyer’s provision of legal services to clients; (2) in other circumstances by an entity controlled or owned by the lawyer individually or with others, unless the lawyer takes reasonable measures to ensure that a person obtaining the law-related services knows that the services are not legal services and that the protections of the client-lawyer relationship do not exist.

vi. It is obvious, by reason of this schema, that the definition of “law related” services is important. The schema defines the ethical and regulatory schema applicable to the attorney who performs law-related services without differentiating these services from legal services.

vii. The term “law-related services” denotes services that might reasonably be performed in conjunction with the provision of legal services and that are not prohibited as unauthorized practice of law when provided by a nonlawyer.\(^7\)

viii. E-discovery services can and are performed by vendors and experts under circumstances that do not constitute the unauthorized practice of law. The single most compelling indicia of a relationship in which e-discovery functions would be considered the practice of law is the rendering of services to an entity that reasonably believed that he/she/it was dealing with a lawyer who was engaged to protect their interest.\(^8\) It appears

---

\(^7\) The definition of “practice of law” rests upon the definition of “legal services”: The unauthorized practice of law, as defined by this court, “is the rendering of legal services for another by any person not admitted to practice in Ohio under Rule I and not granted active status under Rule VI, or certified under Rule II, Rule IX, or Rule XI of the Supreme Court Rules for the Government of the Bar of Ohio.” *Lorain Cty. Bar Assn. v. Kocak*, 121 Ohio St.3d 396, 2009-Ohio-1430, 904 N.E.2d 885, ¶17, quoting Gov.Bar R. VII(2)(A).

\(^8\) The leading case in this state relating to the unauthorized practice of law is *Land Title Abstract & Trust Co. v. Dworken*, supra. There, in the first paragraph of the syllabus, the following rule is announced: “1. The practice of law is not limited to the conduct of cases in court. It embraces the preparation of pleadings and other papers incident to actions and special proceedings and the management of such actions and proceedings on behalf of clients before judges and courts, and in addition conveyancing, the preparation of legal instruments of all kinds, and in general all advice to clients and all action taken for them in matters connected with the law.” The acts stressed in the above definition as constituting the practice of law are *the performance of legal services for others*. Comment Rule 5.7, Ohio Rules of Professional Conduct provides when a lawyer performs law-related services, sometimes referred to as “ancillary business,” or controls an organization that does so, there exists the potential for ethical problems. Principal among these is the possibility that the person for whom the law-related services are performed fails to understand that the services may not carry
that “legal services” and “practice of law” are concepts that are closely related to the relationship created between the lawyer and the client. In order for a non-lawyer to perform services that do not constitute the unauthorized practice of law, the non-lawyer must be careful not to create circumstances in which the client reasonably believed that he/she/it was dealing with a lawyer. Similarly, when a lawyer performs non-lawyer/law related activities and does NOT wish to have these activities subject to the ethical and regulatory schema applicable to the practice of law, the attorney must take reasonable measures to assure that the client knows that the non-lawyer/law related activities are NOT legal services and are NOT protected by the attorney client relationship.

ix. Alternatively, when a law firm performs e-discovery services either as part of the “legal services” provided or without taking any reasonable measures to inform the client as described above, then the lawyer is subject to the Ohio Rules of Professional Conduct in the performance of those e-discovery services. This is significant because the Ohio Rules of Professional Conduct require that attorneys be (a) competent to render services; (b)

III. What is the motivation of law firm to perform e-discovery services?

a. It may be hypothesized that law firms purport to conduct e-discovery services as legal services for the reasons that they do NOT outsource legal services. These firms sell judgment and experience more than legal research. Offshore outsourcing is perceived as undermining their reputation for quality legal services based upon the perception that offshore legal services are delivered by foreign attorneys with a lower quality of education, training, and experience.9

b. Clients get what they pay for, so high rates charged for e-discovery services tend to be interpreted as necessary due to the complexity of the services. Outsourcing tends to require firms to concede that there is no rational basis for high-end fee structures for e-discovery services; especially where those services can be standardized, scaled, and defended from attack by experts who design and implement the protocols.

with them the protections normally afforded as part of the client-lawyer relationship. The recipient of the law-related services may expect, for example, that the protection of client confidences, prohibitions against representation of persons with conflicting interests, and obligations of a lawyer to maintain professional independence apply to the provision of law-related services when that may not be the case.

9 Carole Silver and Mary C. Daly, Flattening the World of Legal Services” The Ethical and Liability Minefields of Offshoring Legal and Law Related Services, Georgetown Journal of International Law, Vol 38 2007 at page 412
IV. What are the ethical issues raised by law firms performing e-discovery services? Competence, attorney-client

a. Comment Rule 5.7, Ohio Rules of Professional conduct provides that when a lawyer is obliged to accord the recipients of [non-legal and law-related services] the protections of those rules that apply to the client-lawyer relationship, the lawyer must take special care to heed the proscriptions of the rules addressing conflict of interest [Rules 1.7 to 1.11, especially Rules 1.7(a)(2) and 1.8(a), (b) and (f)], and scrupulously adhere to the requirements of Rule 1.6 relating to disclosure of confidential information. The promotion of the law-related services must also in all respects comply with Rules 7.1 to 7.3, dealing with advertising and solicitation. In that regard, lawyers should take special care to identify the obligations that may be imposed as a result of a jurisdiction’s decisional law.

b. Law firms that represent themselves as “experts” in e-discovery may run afoul of Rule 7.1 regulating communications about the lawyer’s services.

i. RULE 7.1: COMMUNICATIONS CONCERNING A LAWYER’S SERVICES

ii. A lawyer shall not make or use a false, misleading, or non-verifiable communication about the lawyer or the lawyer’s services. A communication is false or misleading if it contains a material misrepresentation of fact or law or omits a fact necessary to make the statement considered as a whole not materially misleading

iii. Like patent application search and preparation, e-discovery has several functions, some of which require specialized knowledge, training, and language to explain. This renders portions of e-discovery an “expert” function demanding Rule 702 evidence to defend e-discovery decisions from attack. 10

1. Counsel should be prepared to explain what keywords, search protocols, and alternative search methods were used to generate a set of documents, including ones made subject to subsequent manual searches for responsiveness and privilege. This explanation may best come from a technical “IT” expert, a statistician, or an expert in search and retrieval technology. Counsel must be prepared to answer questions, and indeed, to prove the reasonableness and good faith of their method. 11

---


11 Id at 212
2. Whether search terms or “keywords” will yield the information sought is a complicated question involving the interplay, at least, of the sciences of computer technology, statistics and linguistics. Indeed a special project team of the Working Group on Electronic Discovery of the Sedona Conference is studying that subject and their work indicates how difficult this question is. Given this complexity, for lawyers and judges to dare opine that a certain search term or terms would be more likely to produce information than the term or terms used is truly to go where angels fear to tread. This topic is clearly beyond the ken of a layman and requires that any such conclusion be based upon evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence. Accordingly if defendants are going to contend that the search terms used by the government were insufficient, they will have to specifically so contend in a motion to compel and their contention must be based on evidence that meets the requirements of Rule 702 of the Federal Rules of Evidence.\(^{12}\)

3. Whether a law firm or individual attorney is an “expert” in all aspects of e-Discovery is doubtful. While an attorney may understand the case law decisions, the exact protocols and processes that are required in a particular matter may reflect the operating systems, file systems, and data applications used by a client. Touting one’s expertise in communications and advertisements; but failing to be able to defend particular processes and protocols in a matter could create an ethical violation.  

   c. When e-discovery protocols are designed and executed by lawyers as legal services, they can become the channel through which (a) the attorney-client privilege is attacked and (b) clients’ counsel can be disqualified.

      i. By arguing that e-discovery protocols (such as the contents of a litigation hold, or the preservation instructions issued to clients) are protected from discovery or disclosure, counsel appears to be conceding that e-discovery has been conducted under circumstances in which the client believed that the e-discovery services were protected by the attorney-client relationship, rather than being conducted by a third party, electronic evidence expert company.

      ii. Once spoliation is demonstrated, however, the litigation hold instructions are generally discoverable.\(^{13}\)

---


\(^{13}\) E-Discovery practices in a particular matter are subject to cooperation amongst counsel and are discoverable. The case Major Tours v. Colorel, 2009 U.S. Dist. LEXIS 68128 (DNJ Aug. 4, 2009), is
iii. Moreover, searching protocols by which relevant information has been identified and produced may be expert functions requiring Rule 702 evidence to explain and defend.¹⁴

iv. When lawyers perform the expert functions needed to perform e-discovery services they will become fact witnesses and depending upon circumstances, could be disqualified pursuant to Rule 3.7, Ohio Rules of Professional Conduct. Serving as an advocate and e-discovery witness presents combines both proof and analysis of proof which could trigger an objection by the tribunal.

v. Additionally, it appears that if the performance of the e-discovery services have led to court-ordered sanctions, then an attorney may have additional conflict issues trying to remain the client’s counsel while simultaneously protecting the client from the attorney’s negligent performance of e-discovery functions.

V. McDermott law suit as an example of law firm negligently performing e-discovery service—manual review of documents for a particular purpose

i. J-M Manufacturing Company v. McDermott Will & Emery, Case No. BC 462832 Superior Court, Calif, County of Los Angeles, Staley Mosk Courthouse.

ii. J-M former employee, John Hendrix, in 2006 filed a non-public complaint under seal pursuant to False Claims Act (Quit Tam suit) alleging that J-M used piping in sewer systems across US that was made of low quality materials and sloppy manufacturing processes

iii. J-M hired McDermott to defend J-M in the FCA suit

1. McDermott claimed on its website to be knowledgeable in Quit Tam and in e-discovery


1. McDermott and J-M identify 160 custodians

¹⁴ Donald Wochna, Electronic Data, Electronic Searching, Inadvertent Production of Privileged Data: A Perfect Storm, University of Akron, Vol.43, No. 3, 2010
2. “McDermott then made electronic copies of the custodians’ data and transferred the data to third party electronic discovery vendors, Navigant Consulting, Inc., and Stratify.

   a. Vendors are generally not experts and do not take responsibility for the protocols and processes used to achieve a specific objective. This is especially important when the search terms used to identify relevant and/or privileged documents fail to satisfactorily achieve that goal.

   b. Navigant and Stratify were to run a search term filter through the collected documents.

   c. The search term filter, containing keywords, was negotiated with the federal government and contained terms designed to produce documents relevant to the federal government and state subpoenas.

   d. Additionally, Navigant and Stratify were supposed to run a privilege filter through the collected documents in order to separate out privileged documents including attorney-client privileged documents.

   e. McDermott then produced documents containing the keywords to the federal government.


   g. Government RETURNED the documents, advised McDermott of the privileged documents and asked McDermott to try again.

   h. McDermott hired contract attorneys from Hudson Global Resources in New York at $61.00 per hour. Hudson contract attorneys were trained by McDermott to identify potentially privileged documents and segregate documents into three categories: responsive but privileged; responsive and not privileged; non-responsive.

   i. Contract attorneys negligently performed; McDermott negligently performed limited spot-checking of the contract attorneys’ work.
j. 250,000 electronic documents produced to governmental entities, including approximately 3,900 privileged documents.

k. Federal government declined to intervene in the Quit Tam action

l. But federal government delivered to Hendrix, as relator, all privileged documents received.

m. March 2010, McDermott replaced as counsel. Fee dispute along with refusal to provide new counsel the file until paid

n. June 2010, new counsel learns from Relator's counsel about the 3900 privileged documents. Did not learn from McDermott. Relator claims waiver of privilege by McDermott

VI. Conclusions

a. Issue: Why are law firms providing e-discovery services as if the services were legal services?

   i. E-Discovery services are technical services requiring an understanding of computer technology, linguistics, and statistics.

   ii. Preservation techniques include cloning of sources of discoverable information. Explanations of cloning as a process, including explanations of methods for determining reliability through the use of message digests, constitute expert testimony under Rule 702.

   iii. Searching techniques related to unstructured data require Rule 702 testimony to explain and defend.

b. Issue: If lawyers insist on providing electronic discovery solutions, whether lawyer and law firm competently discharged their obligation to provide electronic discovery solutions as legal services given the client’s data architecture, data flow, and device characteristics.

c. Warning signs to counsel

   i. You do not understand the collateral searching issues related to structured versus unstructured data

   ii. You are surprised at the amount of time it takes to make modest, inconsequential changes in data or in the manner in which you want information displayed.
iii. You build entire cases only upon the Content of User Documents

iv. You do not know the difference between User documents and other documents
SUPERIOR COURT OF THE STATE OF CALIFORNIA
COUNTY OF LOS ANGELES - STANLEY MOSK COURTHOUSE

J-M MANUFACTURING COMPANY, INC., a Delaware Corporation, Plaintiff,

vs.

MCDERMOTT WILL & EMERY, a Business Entity, form unknown; and DOES 1 through 100, inclusive,

Defendants.

Case No. BC 462 832
FIRST AMENDED COMPLAINT FOR DAMAGES
Dept. 33
[Assigned to the Honorable Charles F. Palmer for All Purposes]

COMES NOW, plaintiff J-M MANUFACTURING COMPANY, INC., a Delaware Corporation doing business as JM Eagle ("JME" or "PLAINTIFF"), complaining of defendants MCDERMOTT WILL & EMERY, a business entity, form unknown ("MWE" or "McDermott"), and DOES 1 through 100, inclusive, and each of them, (collectively referred to herein as "DEFENDANTS") as follows:

FIRST AMENDED COMPLAINT FOR DAMAGES
GENERAL ALLEGATIONS

A. Identity of the parties and venue.

1. PLAINTIFF JME is now, and at all times herein was, a Delaware Corporation, in good standing and licensed to do business in Los Angeles County, California. JME’s corporate headquarters is in the County of Los Angeles, City of Los Angeles. The factual events set forth herein occurred within Los Angeles County, State of California; accordingly, venue is proper in the above-entitled court.

2. PW Eagle, Inc. was a leading extruder of PVC pipe products and its wholly-owned subsidiary, USPoly Company, LLC, was a leading manufacturer of polyethylene pipe and fittings. Together, PW Eagle and USPoly operated 12 manufacturing facilities across the United States.

3. In January of 2007, JME and PW Eagle, Inc. signed a definitive merger agreement under which JME acquired all of the outstanding common shares of PW Eagle, Inc. for $33.50 per share in cash which represented an implied total equity value of approximately $400 million. One of the main reasons that JME entered into this transaction was to acquire the ownership of patents that were then owned by PW Eagle, Inc.

4. Defendant McDermott is now, and at all times herein was, a business entity, form unknown. PLAINTIFF is informed and believes and thereon alleges that this defendant resides in Los Angeles County, California, and has been doing business in Los Angeles County, California.

5. The true names or capacities, whether individual, corporate, associate, or otherwise of defendants named in this action as Does 1-100 (“DOES”) are unknown to PLAINTIFF, who therefore sues such defendants by such fictitious names. PLAINTIFF will seek leave of this Honorable Court to amend this Complaint when the true names and/or capacities of said defendants have been ascertained. Said fictitiously named defendants are in some manner proximately responsible for the damages suffered by PLAINTIFF herein.
B. **Agency allegations.**

6. Each of the defendants, including the fictitiously named DOE defendants, was the agent, or employee, partner or officer, director or joint venturer of defendants herein, and in doing the things herein alleged acted within the course and scope of said agency, employment, partnership, joint venture, or association and under the direction of, and with the consent and permission, advance knowledge and/or ratification of the other defendants.

7. At all times relevant, defendants including the fictitiously named DOE defendants, and each of them, formed and operated under a common plan and agreement, with the resulting injuries and damages to PLAINTIFF arising from acts done in furtherance of the common design.

C. **Venue.**

8. Venue is proper in Los Angeles County pursuant to California Code of Civil Procedure section 395(a) because PLAINTIFF entered into a contract with defendants in Los Angeles County; and because defendants were to perform legal services for PLAINTIFF herein in "the Underlying Action."

D. **"The Case Within The Case."**

9. JME retained McDermott, which held itself out as knowledgeable in the area of Qui Tam actions and e-discovery, to represent it regarding subpoenas received in 2006 and 2007 from the federal, California, and Tennessee governments which called for production of copies of paper documents as well as electronic documents. The subpoenas sought information regarding the False Claims Act allegations made in the case of United States ex rel. Hendrix v. J-M Manufacturing Co., which was then a non-public complaint filed under seal. Each of the subpoenas required JME to produce paper and electronic documents. During the course of its representation of JME, McDermott billed JME on an hourly basis and at rates as high as $925.00 an hour.

---

FIRST AMENDED COMPLAINT FOR DAMAGES
10. McDermott worked with JME to identify about 160 custodians who likely had responsive electronic information. McDermott then made electronic copies of the custodians' data and transferred the data to third-party electronic discovery vendors, Navigant Consulting, Inc. (“Navigant”) and Stratify, Inc. (“Stratify”) to run a search term filter through the collected documents. The search term filter, containing keywords, was negotiated with the federal government and contained terms designed to produce documents relevant to the federal government and State subpoenas. Additionally, Navigant and Stratify were supposed to run a privilege filter through the collected documents in order to separate out privileged documents including attorney-client privileged documents.

11. McDermott then produced documents containing the keywords to the federal government; however, in addition to responsive documents, McDermott produced to the federal government documents that were not responsive to the subpoenas and were also attorney-client privileged. In response to the disclosure of attorney-client privileged documents to it, the federal government requested McDermott to conduct a further privilege review and then re-submit a new production to it. It is presently unknown to JME what further review was conducted by McDermott, Navigant and/or Stratify, if any, and a second production was sent to the federal government, which, in turn, disclosed this production to the real parties in interest (the “Relators”). As in the first production, McDermott turned over attorney-client information, and other privileged documents, which was also non-responsive to the subpoenas. In May and June of 2011, JME through its new counsel, discovered that approximately 3,900 privileged and non-responsive documents were turned over to the federal government in the second production, and it was these documents which were then, in turn, disclosed to Relators.

12. Prior to the second production of privileged documents to the federal government, McDermott retained Hudson Global Resources (“Hudson”) which provided contract lawyers in the New York area at the rate of $61.00 an hour to review documents that were identified as potentially privileged and to divide them into three categories: responsive but privileged; responsive and not privileged; and, non-responsive. McDermott’s lawyers participated in the hiring of Hudson’s contract lawyers and also assisted in their training. The contract lawyers
negligently performed their duties, and McDermott’s lawyers, in turn, negligently performed
limited spot-checking of the contract attorneys’ work. The combined efforts of the contract
lawyers, and the McDermott lawyers who supervised them, fell below the applicable standard of
care for lawyers because McDermott did not properly supervise the contract lawyers and failed to
thoroughly review the documents that they reviewed to determine whether any or a large number
of privileged documents were being disclosed.

13. About 250,000 electronic documents were produced to the governmental entities,
including many based on Hudson’s assessment of responsiveness and privilege; and included in
the second production were approximately 3,900 privileged documents. Although the federal
government declined to intervene in United States ex rel. Hendrix v. J-M Manufacturing Co., the
documents produced by McDermott were nonetheless turned over to the counsel for the Relators
in that case.

Hampton, LLP ("Sheppard") to replace McDermott. In or about June 2010, counsel for the
Relators informed Sheppard that it had received JME’s document production from the federal
government. Like the predecessor United States Attorney’s Office, which represented the federal
government, Relators’ counsel discovered that clearly privileged documents were contained in
McDermott’s production of documents. However, unlike the predecessor United States
Attorney’s Office, which immediately returned the privileged documents to McDermott, counsel
for the Relator did not immediately return the privileged documents to JME. Thereafter,
Sheppard asked Relators’ counsel to destroy or return the privileged documents. Relators’
counsel refused, alleging that JME waived the attorney-client privilege with respect to the subject
matter of the documents because McDermott twice conducted privilege reviews before producing
the documents to the federal government. JME is informed and believes and thereon alleges that
approximately 3,900 privileged documents were produced by McDermott without JME’s
informed consent and that such 3,900 documents should not have been produced by McDermott.

15. On March 15, 2010, McDermott was substituted out of the case of United States
ex rel. Hendrix v. J-M Manufacturing Co.; however, McDermott continued to provide legal

FIRST AMENDED COMPLAINT FOR DAMAGES
services in regard to United States ex rel. Hendrix v. J-M Manufacturing Co. at least until November 2010.

E. The Intentional Delay in Turning Over JME’s File.

16. After March 15, 2010, and in violation of California Rule of Professional Conduct 3-700(D), McDermott refused to turn over JME’s file until JME paid McDermott’s outstanding bill and, on information and belief, McDermott delayed turning over JME’s file in order to hide its negligent disclosure of privileged documents. McDermott’s conduct was, in part, reflected in the following emails:

   (A.) On October 21, 2010, Joel Rubinstein, a partner with MWE, wrote the following in an e-mail to Mr. Wang, the Chief Executive Officer and President of JME:

   “Subject: RE: Qui Tam

   Walter,

   I’m told that our firm policy is not to release all files until full payment is made.

   If you’d like all the files now, please send a check for the entire $530,477 and we’ll get them all out to you promptly. I think that’s the best course anyway, since this has taken so long already, and there’s no need to prolong it further.” [Emphasis added.]

   (B.) Shortly after receiving Mr. Rubinstein’s e-mail, Mr. Wang responded in an e-mail as follows:

   “Subject: RE: Qui Tam

   JOEL:

   SO YOU, AND YOUR FIRM IS HOLDING THESE FILES HOSTAGE WHEN WE HAVE AGREED TO THESE PAYMENTS. WHY DO THIS WHEN YOU CAN GET ALL THIS SETTLED?

   WALTER WANG” [Emphasis added.]
Later, on December 27, 2010, Claudia Herrarte, of JME, sent the following e-mail to Mr. Rubinstein:

“Subject: Invoices

Dear Joel,

Please advise as to when we can expect a release for JM Eagle's liabilities.

Claudia”

(D.) McDermott then responded to the above e-mail from JME, also on December 27, 2010, as follows:

“From: Rubinstein, Joel jrubinstein@mwe.com
Subject: RE: Invoices

Once we receive the final payment on the 31st, we will be in touch to coordinate sending all of the files to you and will confirm that JM does not owe anything further. Thanks.

Joel L. Rubinstein” [Emphasis added.]

17. Based in part on McDermott’s conduct, which violated the California Rules of Professional Conduct, JME paid McDermott’s outstanding bills, which were greater than the value of the negligent services it provided to JME. Additionally, the delay by McDermott in turning over JME’s file prevented and delayed JME from learning the true nature and extent of the negligent disclosure of privileged information and materials by McDermott until shortly before the filing of the complaint in this case.

18. McDermott represented JME in its acquisition of PW Eagle, Inc. and, on information and belief, JME alleges that McDermott was responsible for conducting the due diligence in regard to this transaction. After completion of the acquisition, JME discovered that some of the patents it acquired did not comply with then-existing specifications.
19. JME refers to and incorporates paragraphs 1 through 18 of this Complaint as though fully set forth herein.

20. Defendants owed PLAINTIFF a duty to render legal services competently. Defendants breached that duty by, *inter alia*, producing privileged documents to parties adverse to JME in litigation without obtaining its informed consent, failing to supervise attorneys and vendors McDermott contracted with to perform the review and production of documents, charging JME fees and costs for performance of work not properly performed, or not performed at all, billing JME for work that was unnecessary, approving for payment third-party vendor billing for work not competently handled, and refusing to turn over JME’s file until its outstanding and inflated bills were paid.

21. JME alleges, upon information and belief, that McDermott failed to inform PLAINTIFF that attorney-client privileged documents and other privileged documents had been disclosed, failed to disclose the scope of such release of privileged documents and failed to take any steps to obtain the return of such privileged documents.

22. JME alleges on information and belief that McDermott negligently represented JME in its acquisition of PW Eagle, Inc. because it failed to discover and inform JME that some of the patents it acquired did not comply with then-existing specifications which rendered the patents worthless.

23. Defendants’ breach caused JME damages in an amount to be proven at trial, but in no event less than the jurisdictional minimum of this Court.

**SECOND CAUSE OF ACTION**

**Breach of Fiduciary Duty against all defendants**

24. PLAINTIFF refers to and incorporates paragraphs 1 through 23 of this Complaint as though fully set forth herein.

25. Defendants owed PLAINTIFF a fiduciary duty. Defendants breached that duty by, *inter alia*, producing privileged documents to parties adverse to JME in several lawsuits, failing...
to supervise attorneys and vendors Mc Dermott contracted with to perform the review and
production of documents, charging JME fees and costs for performance of such work that was
not properly performed, or not performed at all, billing JME for work that was unnecessary,
approving for payment third-party vendor billing for work not competently handled, refusing to
turn over JME’s file until its outstanding bills were paid and failing to inform JME that its
refusal to turn over JME’s file until its outstanding bills were paid violated the California Rules
of Professional Conduct.

26. Defendants breach of their fiduciary duty to their client caused JME damages in
an amount to be proven at trial, but in no event less than the jurisdictional minimum of this
Court. On information and belief, PLAINTIFF alleges that Mc Dermott’s actions were
intentional, fraudulent, oppressive and malicious, and were committed with the specific intent of
caus ing PLAINTIFF injury and damage and/or were in conscious disregard of Plaintiffs’ rights.
Accordingly, Plaintiffs are entitled to recover exemplary and punitive damages.

27. The actions of defendants were (1) fraudulent, meaning an intentional
misrepresentation, deceit, or concealment of a material fact known to the defendants with the
intention on the part of the defendant of thereby depriving a person of property or legal rights or
otherwise causing injury; (2) malicious, meaning conduct which is intended by the defendants to
cause injury to PLAINTIFF or despicable conduct which is carried on by the defendants with a
willful and conscious disregard of the rights or safety of others; (3) and/or oppressive, meaning
despicable conduct that subjects a person to cruel and unjust hardship in conscious disregard of
that person’s rights; and done with the intention of depriving PLAINTIFF of substantial rights.
PLAINTIFF is therefore entitled to punitive damages in a sufficient amount to make an example
of, punish defendants, and deter future fraudulent, oppressive and malicious misconduct in an
amount according to proof at trial.

FIRST AMENDED COMPLAINT FOR DAMAGES
THIRD CAUSE OF ACTION
(Accounting Against All Defendants)

28. PLAINTIFF refers to and incorporates paragraphs 1 through 27 of this Complaint as though fully set forth herein.

29. JME is informed and believes and thereon alleges that Defendants billed JME for work not necessary, not done and not done competently, all of which was unlawful.

30. The amount of such billings is unknown to JME and cannot be ascertained without an accounting of the amounts paid for such billing.

WHEREFORE, PLAINTIFF prays for judgment against defendants, as follows:

ON THE FIRST CAUSE OF ACTION:

1. Damages according to proof at trial.

ON THE SECOND CAUSE OF ACTION:

1. Damages according to proof at trial.

2. Punitive damages in an amount according to proof at trial.

ON THE THIRD CAUSE OF ACTION:

1. An accounting of the fees and costs paid to Defendants.

ON ALL CAUSES OF ACTION:

1. For costs of suit incurred herein.

2. For a reasonable attorneys' fees.

3. Such other and further relief as the Court may deem just and proper.

FIRST AMENDED COMPLAINT FOR DAMAGES
Respectfully submitted,

Dated: July 20, 2011.

HOBSON, DUNGOG, BERNARDINO & DAVIS, LLP

By: Rafael Bernardino, Jr.
Attorneys for PLAINTIFF
J-M MANUFACTURING COMPANY, INC.
UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

MONIQUE DA SILVA MOORE, et al., : 11 Civ. 1279 (ALC) (AJP)

-against- : OPINION AND ORDER

PUBLICIS GROUPE & MSL GROUP, :

Defendants. :

------------------------------ X

ANDREW J. PECK, United States Magistrate Judge:

In my article Search, Forward: Will manual document review and keyword searches be replaced by computer-assisted coding?, I wrote:

To my knowledge, no reported case (federal or state) has ruled on the use of computer-assisted coding. While anecdotally it appears that some lawyers are using predictive coding technology, it also appears that many lawyers (and their clients) are waiting for a judicial decision approving of computer-assisted review.

Perhaps they are looking for an opinion concluding that: "It is the opinion of this court that the use of predictive coding is a proper and acceptable means of conducting searches under the Federal Rules of Civil Procedure, and furthermore that the software provided for this purpose by [insert name of your favorite vendor] is the software of choice in this court." If so, it will be a long wait.

....

Until there is a judicial opinion approving (or even critiquing) the use of predictive coding, counsel will just have to rely on this article as a sign of judicial approval. In my opinion, computer-assisted coding should be used in those cases where it will help "secure the just, speedy, and inexpensive" (Fed. R. Civ. P. 1) determination of cases in our e-discovery world.
Andrew Peck, Search, Forward, L. Tech. News, Oct. 2011, at 25, 29. This judicial opinion now recognizes that computer-assisted review is an acceptable way to search for relevant ESI in appropriate cases.\(^{2}\)

**CASE BACKGROUND**

In this action, five female named plaintiffs are suing defendant Publicis Groupe, "one of the world's 'big four' advertising conglomerates," and its United States public relations subsidiary, defendant MSL Group. (See Dkt. No. 4: Am. Compl. ¶¶ 1, 5, 26-32.) Plaintiffs allege that defendants have a "glass ceiling" that limits women to entry level positions, and that there is "systemic, company-wide gender discrimination against female PR employees like Plaintiffs." (Am. Compl. ¶¶ 4-6, 8.) Plaintiffs allege that the gender discrimination includes

(a) paying Plaintiffs and other female PR employees less than similarly-situated male employees; (b) failing to promote or advance Plaintiffs and other female PR employees at the same rate as similarly-situated male employees; and (c) carrying out discriminatory terminations, demotions and/or job reassignments of female PR employees when the company reorganized its PR practice beginning in 2008 . . . .

(Am. Compl. ¶ 8.)

Plaintiffs assert claims for gender discrimination under Title VII (and under similar New York State and New York City laws) (Am. Compl. ¶¶ 204-25), pregnancy discrimination under Title VII and related violations of the Family and Medical Leave Act (Am. Compl. ¶¶ 239-71), as

---

\(^{2}\) To correct the many blogs about this case, initiated by a press release from plaintiffs' vendor – the Court did not order the parties to use predictive coding. The parties had agreed to defendants' use of it, but had disputes over the scope and implementation, which the Court ruled on, thus accepting the use of computer-assisted review in this lawsuit.

OSIPA/DA SILVA MOORE - ESI

The complaint seeks to bring the Equal Pay Act/FLSA claims as a "collective action" (i.e., opt-in) on behalf of all "current, former, and future female PR employees" employed by defendants in the United States "at any time during the applicable liability period" (Am. Compl. ¶¶ 179-80, 190-203), and as a class action on the gender and pregnancy discrimination claims and on the New York Labor Law pay claim (Am. Compl. ¶¶ 171-98). Plaintiffs, however, have not yet moved for collective action or class certification at this time.

Defendant MSL denies the allegations in the complaint and has asserted various affirmative defenses. (See generally Dkt. No. 19: MSL Answer.) Defendant Publicis is challenging the Court's jurisdiction over it, and the parties have until March 12, 2012 to conduct jurisdictional discovery. (See Dkt. No. 44: 10/12/11 Order.)

**COMPUTER-ASSISTED REVIEW EXPLAINED**

My Search, Forward article explained my understanding of computer-assisted review, as follows:

By computer-assisted coding, I mean tools (different vendors use different names) that use sophisticated algorithms to enable the computer to determine relevance, based on interaction with (i.e., training by) a human reviewer.

Unlike manual review, where the review is done by the most junior staff, computer-assisted coding involves a senior partner (or [small] team) who review and code a "seed set" of documents. The computer identifies properties of those documents that it uses to code other documents. As the senior reviewer continues to code more sample documents, the computer predicts the reviewer's coding. (Or, the computer codes some documents and asks the senior reviewer for feedback.)
When the system's predictions and the reviewer's coding sufficiently coincide, the system has learned enough to make confident predictions for the remaining documents. Typically, the senior lawyer (or team) needs to review only a few thousand documents to train the computer.

Some systems produce a simple yes/no as to relevance, while others give a relevance score (say, on a 0 to 100 basis) that counsel can use to prioritize review. For example, a score above 50 may produce 97% of the relevant documents, but constitutes only 20% of the entire document set.

Counsel may decide, after sampling and quality control tests, that documents with a score of below 15 are so highly likely to be irrelevant that no further human review is necessary. Counsel can also decide the cost-benefit of manual review of the documents with scores of 15-50.


My article further explained my belief that Daubert would not apply to the results of using predictive coding, but that in any challenge to its use, this Judge would be interested in both the process used and the results:

[I]f the use of predictive coding is challenged in a case before me, I will want to know what was done and why that produced defensible results. I may be less interested in the science behind the "black box" of the vendor's software than in whether it produced responsive documents with reasonably high recall and high precision.

That may mean allowing the requesting party to see the documents that were used to train the computer-assisted coding system. (Counsel would not be required to explain why they coded documents as responsive or non-responsive, just what the coding was.) Proof of a valid "process," including quality control testing, also will be important.

2

2 From a different perspective, every person who uses email uses predictive coding, even if they do not realize it. The "spam filter" is an example of predictive coding.
Of course, the best approach to the use of computer-assisted coding is to follow the Sedona Cooperation Proclamation model. Advise opposing counsel that you plan to use computer-assisted coding and seek agreement; if you cannot, consider whether to abandon predictive coding for that case or go to the court for advance approval.

Id.

THE ESI DISPUTES IN THIS CASE AND THEIR RESOLUTION

After several discovery conferences and rulings by Judge Sullivan (the then-assigned District Judge), he referred the case to me for general pretrial supervision. (Dkt. No. 48: 11/28/11 Referral Order.) At my first discovery conference with the parties, both parties’ counsel mentioned that they had been discussing an “electronic discovery protocol,” and MSL’s counsel stated that an open issue was “plaintiff’s reluctance to utilize predictive coding to try to cull down the” approximately three million electronic documents from the agreed-upon custodians. (Dkt. No. 51: 12/2/11 Conf. Tr. at 7-8.) ² Plaintiffs’ counsel clarified that MSL had "over simplified [plaintiffs’] stance on predictive coding," i.e., that it was not opposed but had "multiple concerns . . . on the way in which [MSL] plan to employ predictive coding" and plaintiffs wanted "clarification." (12/2/11 Conf. Tr. at 21.)

The Court did not rule but offered the parties the following advice:

Now, if you want any more advice, for better or for worse on the ESI plan and whether predictive coding should be used, . . . I will say right now, what should

² When defense counsel mentioned the disagreement about predictive coding, I stated that: "You must have thought you died and went to Heaven when this was referred to me," to which MSL’s counsel responded: "Yes, your Honor. Well, I'm just thankful that, you know, we have a person familiar with the predictive coding concept." (12/2/11 Conf. Tr. at 8-9.)
not a surprise, I wrote an article in the October Law Technology News called Search Forward, which says predictive coding should be used in the appropriate case.

Is this the appropriate case for it? You all talk about it some more. And if you can't figure it out, you are going to get back in front of me. Key words, certainly unless they are well done and tested, are not overly useful. Key words along with predictive coding and other methodology, can be very instructive.

I'm also saying to the defendants who may, from the comment before, have read my article. If you do predictive coding, you are going to have to give your seed set, including the seed documents marked as nonresponsive to the plaintiff's counsel so they can say, well, of course you are not getting any [relevant] documents, you're not appropriately training the computer.

(12/2/11 Conf. Tr. at 20-21.) The December 2, 2011 conference adjourned with the parties agreeing to further discuss the ESI protocol. (12/2/11 Conf. Tr. at 34-35.)

The ESI issue was next discussed at a conference on January 4, 2012. (Dkt. No. 71: 1/4/12 Conf. Tr.) Plaintiffs' ESI consultant conceded that plaintiffs "have not taken issue with the use of predictive coding or, frankly, with the confidence levels that they [MSL] have proposed . . . ."

(1/4/12 Conf. Tr. at 51.) Rather, plaintiffs took issue with MSL's proposal that after the computer was fully trained and the results generated, MSL wanted to only review and produce the top 40,000 documents, which it estimated would cost $200,000 (at $5 per document). (1/4/12 Conf. Tr. at 47-48, 51.) The Court rejected MSL's 40,000 documents proposal as a "pig in a poke." (1/4/12 Conf. Tr. at 51-52.) The Court explained that "where [the] line will be drawn [as to review and production] is going to depend on what the statistics show for the results," since "[p]roportionality requires consideration of results as well as costs. And if stopping at 40,000 is going to leave a tremendous number of likely highly responsive documents unproduced, [MSL's proposed cutoff] doesn't work." (1/4/12 Conf. Tr. at 51-52; see also id. at 57-58; Dkt. No. 88: 2/8/12 Conf. Tr. at 84.)

G/A/P/D/A SILVA MOORE - ESI

1.32 · Ethical Issues and Practicing Law
The parties agreed to further discuss and finalize the ESI protocol by late January 2012, with a conference held on February 8, 2012. (1/4/12 Conf. Tr. at 60-66; see 2/8/12 Conf. Tr.)

**Custodians**

The first issue regarding the ESI protocol involved the selection of which custodians' emails would be searched. MSL agreed to thirty custodians for a "first phase." (Dkt. No. 88: 2/8/12 Conf. Tr. at 23-24.) MSL's custodian list included the president and other members of MSL's "executive team," most of its HR staff and a number of managing directors. (2/8/12 Conf. Tr. at 24.) Plaintiffs sought to include as additional custodians seven male "comparators," explaining that the comparators' emails were needed in order to find information about their job duties and how their duties compared to plaintiffs' job duties. (2/8/12 Conf. Tr. at 25-27.) Plaintiffs gave an example of the men being given greater "client contact" or having better job assignments. (2/8/12 Conf. Tr. at 28-30.) The Court held that the search of the comparators' emails would be so different from that of the other custodians that the comparators should not be included in the emails subjected to predictive coding review. (2/8/12 Conf. Tr. at 28, 30.) As a fallback position, plaintiffs proposed to "treat the comparators as a separate search," but the Court found that plaintiffs could not describe in any meaningful way how they would search the comparators' emails, even as a separate search. (2/8/12 Conf. Tr. at 30-31.) Since the plaintiffs likely could develop the information needed through depositions of the comparators, the Court ruled that the comparators' emails would not be included in phase one. (2/8/12 Conf. Tr. at 31.)

Plaintiffs also sought to include MSL’s CEO, Olivier Fleuriot, located in France and whose emails were mostly written in French. (2/8/12 Conf. Tr. at 32-34.) The Court concluded that
because his emails with the New York based executive staff would be gathered from those custodians, and Fleuriot's emails stored in France likely would be covered by the French privacy and blocking laws. Fleuriot should not be included as a first-phase custodian. (2/8/12 Conf. Tr. at 35.)

Plaintiffs sought to include certain managing directors from MSL offices at which no named plaintiff worked. (2/8/12 Conf. Tr. at 36-37.) The Court ruled that since plaintiffs had not yet moved for collective action status or class certification, until the motions were made and granted, discovery would be limited to offices (and managing directors) where the named plaintiffs had worked. (2/8/12 Conf. Tr. at 37-39.)

The final issue raised by plaintiffs related to the phasing of custodians and the discovery cut-off dates. MSL proposed finishing phase-one discovery completely before considering what to do about a second phase. (See 2/8/12 Conf. Tr. at 36.) Plaintiffs expressed concern that there would not be time for two separate phases, essentially seeking to move the phase-two custodians back into phase one. (2/8/12 Conf. Tr. at 35-36.) The Court found MSL's separate phase approach to be more sensible and noted that if necessary, the Court would extend the discovery cutoff to allow the parties to pursue discovery in phases. (2/8/12 Conf. Tr. at 36, 50.)

---

Sources of ESI

The parties agreed on certain ESI sources, including the "EMC SourceOne [Email] Archive," the "PeopleSoft" human resources information management system and certain other sources including certain HR "shared" folders. (See Dkt. No. 88: 2/8/12 Conf. Tr. at 44-45, 50-51.) As to other "shared" folders, neither side was able to explain whether the folders merely contained forms and templates or collaborative working documents; the Court therefore left those shared folders for phase two unless the parties promptly provided information about likely contents. (2/8/12 Conf. Tr. at 47-48.)

The Court noted that because the named plaintiffs worked for MSL, plaintiffs should have some idea what additional ESI sources, if any, likely had relevant information; since the Court needed to consider proportionality pursuant to Rule 26(b)(2)(C), plaintiffs needed to provide more information to the Court than they were doing if they wanted to add additional data sources into phase one. (2/8/12 Conf. Tr. at 49-50.) The Court also noted that where plaintiffs were getting factual information from one source (e.g., pay information, promotions, etc.), "there has to be a limit to redundancy" to comply with Rule 26(b)(2)(C). (2/8/12 Conf. Tr. at 54.)

The Predictive Coding Protocol

The parties agreed to use a 95% confidence level (plus or minus two percent) to create a random sample of the entire email collection; that sample of 2,399 documents will be

---

2 The Court also suggested that the best way to resolve issues about what information might be found in a certain source is for MSL to show plaintiffs a sample printout from that source. (2/8/12 Conf. Tr. at 55-56.)
reviewed to determine relevant (and not relevant) documents for a "seed set" to use to train the predictive coding software. (Dkt. No. 88: 2/8/12 Conf. Tr. at 59-61.) An area of disagreement was that MSL reviewed the 2,399 documents before the parties agreed to add two additional concept groups (i.e., issue tags). (2/8/12 Conf. Tr. at 62.) MSL suggested that since it had agreed to provide all 2,399 documents (and MSL's coding of them) to plaintiffs for their review, plaintiffs can code them for the new issue tags, and MSL will incorporate that coding into the system. (2/8/12 Conf. Tr. at 64.) Plaintiffs' vendor agreed to that approach. (2/8/12 Conf. Tr. at 64.)

To further create the seed set to train the predictive coding software, MSL coded certain documents through "judgmental sampling." (2/8/12 Conf. Tr. at 64.) The remainder of the seed set was created by MSL reviewing "keyword" searches with Boolean connectors (such as "training and Da Silva Moore," or "promotion and Da Silva Moore") and coding the top fifty hits from those searches. (2/8/12 Conf. Tr. at 64-66, 72.) MSL agreed to provide all those documents (except privileged ones) to plaintiffs for plaintiffs to review MSL's relevance coding. (2/8/12 Conf. Tr. at 66.) In addition, plaintiffs provided MSL with certain other keywords, and MSL used the same process with plaintiffs' keywords as with the MSL keywords, reviewing and coding an additional 4,000 documents. (2/8/12 Conf. Tr. at 68-69, 71.) All of this review to create the seed set was done by senior attorneys (not paralegals, staff attorneys or junior associates). (2/8/12 Conf. Tr. at 92-93.) MSL reconfirmed that "all of the documents that are reviewed as a function of the seed set, whether [they] are ultimately coded relevant or irrelevant, aside from privilege, will be turned over to" plaintiffs. (2/8/12 Conf. Tr. at 73.)
The next area of discussion was the iterative rounds to stabilize the training of the software. MSL's vendor's predictive coding software ranks documents on a score of 100 to zero, i.e., from most likely relevant to least likely relevant. (2/8/12 Conf. Tr. at 70.) MSL proposed using seven iterative rounds; in each round they would review at least 500 documents from different concept clusters to see if the computer is returning new relevant documents. (2/8/12 Conf. Tr. at 73-74.) After the seventh round, to determine if the computer is well trained and stable, MSL would review a random sample (of 2,399 documents) from the discards (i.e., documents coded as non-relevant) to make sure the documents determined by the software to not be relevant do not, in fact, contain highly-relevant documents. (2/8/12 Conf. Tr. at 74-75.) For each of the seven rounds and the final quality-check random sample, MSL agreed that it would show plaintiffs all the documents it looked at including those deemed not relevant (except for privileged documents). (2/8/12 Conf. Tr. at 76.)

Plaintiffs' vendor noted that "we don't at this point agree that this is going to work. This is new technology and it has to be proven out." (2/8/12 Conf. Tr. at 75.) Plaintiffs' vendor agreed, in general, that computer-assisted review works, and works better than most alternatives. (2/8/12 Conf. Tr. at 76.) Indeed, plaintiffs' vendor noted that "it is fair to say [that] we are big proponents of it." (2/8/12 Conf. Tr. at 76.) The Court reminded the parties that computer-assisted review "works better than most of the alternatives, if not all of the [present] alternatives. So the idea is not to make this perfect, it's not going to be perfect. The idea is to make it significantly better than the alternatives without nearly as much cost." (2/8/12 Conf. Tr. at 76.)
The Court accepted MSL’s proposal for the seven iterative reviews, but with the following caveat:

But if you get to the seventh round and [plaintiffs] are saying that the computer is still doing weird things, it's not stabilized, etc., we need to do another round or two, either you will agree to that or you will both come in with the appropriate QC information and everything else and [may be ordered to] do another round or two or five or 500 or whatever it takes to stabilize the system.

(2/8/12 Conf. Tr. at 76-77; see also id. at 83-84, 88.)

On February 17, 2012, the parties submitted their "final" ESI Protocol which the Court "so ordered." (Dkt. No. 92: 2/17/12 ESI Protocol & Order;) Because this is the first Opinion dealing with predictive coding, the Court annexes hereto as an Exhibit the provisions of the ESI Protocol dealing with the predictive coding search methodology.

**OBSERVATIONS ON PLAINTIFF’S OBJECTIONS TO THE COURT’S RULINGS**

On February 22, 2012, plaintiffs filed objections to the Court's February 8, 2012 rulings. (Dkt. No. 93: PIs. Rule 72(a) Objections; see also Dkt. No. 94: Nurhussein Aff.; Dkt. No. 95: Neale Aff.) While those objections are before District Judge Carter, a few comments are in order.

---

5. Plaintiffs included a paragraph noting its objection to the ESI Protocol, as follows:

Plaintiffs object to this ESI Protocol in its entirety. Plaintiffs submitted their own proposed ESI Protocol to the Court, but it was largely rejected. The Court then ordered the parties to submit a joint ESI Protocol reflecting the Court's rulings. Accordingly, Plaintiffs jointly submit this ESI Protocol with MSL, but reserve the right to object to its use in this case.

(ESI Protocol ¶ J.1 at p. 22.)
Plaintiffs' Reliance on Rule 26(g)(1)(A) is Erroneous

Plaintiffs' objections to my February 8, 2012 rulings assert that my acceptance of MSL's predictive coding approach "provides unlawful 'cover' for MSL's counsel, who has a duty under FRCP 26(g) to 'certify' that their client's document production is 'complete' and 'correct' as of the time it was made. FRCP 26(g)(1)(A)." (Dkt. No. 93: Pls. Rule 72(a) Objections at 8 n.7; accord., id. at 2.) In large-data cases like this, involving over three million emails, no lawyer using any search method could honestly certify that its production is "complete" – but more importantly, Rule 26(g)(1) does not require that. Plaintiffs simply misread Rule 26(g)(1). The certification required by Rule 26(g)(1) applies "with respect to a disclosure." Fed. R. Civ. P. 26(g)(1)(A) (emphasis added). That is a term of art, referring to the mandatory initial disclosures required by Rule 26(a)(1). Since the Rule 26(a)(1) disclosure is information (witnesses, exhibits) that "the disclosing party may use to support its claims or defenses," and failure to provide such information leads to virtually automatic preclusion, see Fed. R. Civ. P. 37(c)(1), it is appropriate for the Rule 26(g)(1)(A) certification to require disclosures be "complete and correct."

Rule 26(g)(1)(B) is the provision that applies to discovery responses. It does not call for certification that the discovery response is "complete," but rather incorporates the Rule 26(b)(2)(C) proportionality principle. Thus, Rule 26(g)(1)(A) has absolutely nothing to do with MSL's obligations to respond to plaintiffs' discovery requests. Plaintiffs' argument is based on a misunderstanding of Rule 26(g)(1).\footnote{Rule 26(g)(1) provides: (continued...)}
Rule 702 and Daubert Are Not Applicable to Discovery Search Methods

Plaintiffs' objections also argue that my acceptance of MSL's predictive coding protocol "is contrary to Federal Rule of Evidence 702" and "violates the gatekeeping function underlying Rule 702." (Dkt. No. 93: PLS Rule 72(a) Objections at 2-3; accord, id. at 10-12.)

(...continued)

(g) Signing Disclosures and Discovery Requests, Responses, and Objections.

(1) Signature Required; Effect of Signature. Every disclosure under Rule 26(a)(1) or (a)(3) and every discovery request, response, or objection must be signed by at least one attorney of record in the attorney's own name . . . . By signing, an attorney or party certifies that to the best of the person's knowledge, information, and belief formed after a reasonable inquiry:

(A) with respect to a disclosure, it is complete and correct as of the time it is made; and

(B) with respect to a discovery request, response, or objection, it is:

(i) consistent with these rules and warranted by existing law or by a nonfrivolous argument for extending, modifying, or reversing existing law, or for establishing new law;

(ii) not interposed for any improper purpose, such as to harass, cause unnecessary delay, or needlessly increase the cost of litigation; and

(iii) neither unreasonable nor unduly burdensome or expensive, considering the needs of the case, prior discovery in the case, the amount in controversy, and the importance of the issues at stake in the action.


As part of this argument, plaintiffs complain that although both parties' experts (i.e., vendors) spoke at the discovery conferences, they were not sworn in. (pls. Rule 72(a) Objections at 12: "To his credit, the Magistrate [Judge] did ask the parties to bring [to the (continued...)

OJ/A/PDA SILVA MOORE - ESI

1.40 • Ethical Issues and Practicing Law
Federal Rule of Evidence 702 and the Supreme Court's *Daubert* decision\(^2\) deal with the trial court's role as gatekeeper to exclude unreliable expert testimony from being submitted to the jury at trial. See also Advisory Comm. Notes to Fed. R. Evid. 702. It is a rule for admissibility of evidence at trial.

If MSL sought to have its expert testify at trial and introduce the results of its ESI protocol into evidence, *Daubert* and Rule 702 would apply. Here, in contrast, the tens of thousands of emails that will be produced in discovery are not being offered into evidence at trial as the result of a scientific process or otherwise. The admissibility of specific emails at trial will depend upon each email itself (for example, whether it is hearsay, or a business record or party admission), not how it was found during discovery.

Rule 702 and *Daubert* simply are not applicable to how documents are searched for and found in discovery.

**Plaintiffs' Reliability Concerns Are, At Best, Premature**

Finally, plaintiffs' objections assert that "MSL's method lacks the necessary standards for assessing whether its results are accurate; in other words, there is no way to be certain if MSL's method is reliable." (Dkt. No. 93: Pls. Rule 72(a) Objections at 13-18.) Plaintiffs' concerns may be appropriate for resolution during or after the process (which the Court will be closely

\(^2\) (...continued)

conference] the ESI experts they had hired to advise them regarding the creation of an ESI protocol. These experts, however, were never sworn in, and thus the statements they made in court at the hearings were not sworn testimony made under penalty of perjury.") Plaintiffs never asked the Court to have the experts testify to their qualifications or be sworn in.

supervising), but are premature now. For example, plaintiffs complain that "MSL’s method fails to include an agreed-upon standard of relevance that is transparent and accessible to all parties. . . . Without this standard, there is a high-likelihood of delay as the parties resolve disputes with regard to individual documents on a case-by-case basis." (Id. at 14.) Relevance is determined by plaintiffs’ document demands. As statistics show, perhaps only 5% of the disagreement among reviewers comes from close questions of relevance, as opposed to reviewer error. (See page 18 n.11 below.)

The issue regarding relevance standards might be significant if MSL’s proposal was not totally transparent. Here, however, plaintiffs will see how MSL has coded every email used in the seed set (both relevant and not relevant), and the Court is available to quickly resolve any issues.

Plaintiffs complain they cannot determine if "MSL’s method actually works" because MSL does not describe how many relevant documents are permitted to be located in the final random sample of documents the software deemed irrelevant. (Pls. Rule 72(a) Objections at 15-16.) Plaintiffs argue that "without any decision about this made in advance, the Court is simply kicking the can down the road." (Id. at 16.) In order to determine proportionality, it is necessary to have more information than the parties (or the Court) now has, including how many relevant documents will be produced and at what cost to MSL. Will the case remain limited to the named plaintiffs, or will plaintiffs seek and obtain collective action and/or class action certification? In the final sample of documents deemed irrelevant, are any relevant documents found that are "hot," "smoking gun" documents (i.e., highly relevant)? Or are the only relevant documents more of the same thing? One hot document may require the software to be re-trained (or some other search method employed), while several documents that really do not add anything to the case might not matter. These types
of questions are better decided "down the road," when real information is available to the parties and the Court.

FURTHER ANALYSIS AND LESSONS FOR THE FUTURE

The decision to allow computer-assisted review in this case was relatively easy – the parties agreed to its use (although disagreed about how best to implement such review). The Court recognizes that computer-assisted review is not a magic, Staples-Easy-Button, solution appropriate for all cases. The technology exists and should be used where appropriate, but it is not a case of machine replacing humans: it is the process used and the interaction of man and machine that the courts needs to examine.

The objective of review in ediscovery is to identify as many relevant documents as possible, while reviewing as few non-relevant documents as possible. Recall is the fraction of relevant documents identified during a review; precision is the fraction of identified documents that are relevant. Thus, recall is a measure of completeness, while precision is a measure of accuracy or correctness. The goal is for the review method to result in higher recall and higher precision than another review method, at a cost proportionate to the "value" of the case. See, e.g., Maura R. Grossman & Gordon V. Cormack, Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review, Rich. J.L. & Tech., Spring 2011, at 8-9, available at http://jolt.richmond.edu/v17i3/article11.pdf.
The slightly more difficult case would be where the producing party wants to use computer-assisted review and the requesting party objects. The question to ask in that situation is what methodology would the requesting party suggest instead? Linear manual review is simply too expensive where, as here, there are over three million emails to review. Moreover, while some lawyers still consider manual review to be the "gold standard," that is a myth, as statistics clearly show that computerized searches are at least as accurate, if not more so, than manual review. Herb Roitblatt, Anne Kershaw, and Patrick Oot of the Electronic Discovery Institute conducted an empirical assessment to "answer the question of whether there was a benefit to engaging in a traditional human review or whether computer systems could be relied on to produce comparable results," and concluded that "[o]n every measure, the performance of the two computer systems was at least as accurate (measured against the original review) as that of human re-review." Herbert L. Roitblatt, Anne Kershaw & Patrick Oot, Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review, 61 J. Am. Soc'y for Info. Sci. & Tech. 70, 79 (2010).

The tougher question, raised in Klein Prods. LLC v. Packaging Corp. of Am. before Magistrate Judge Nan Nolan in Chicago, is whether the Court, at plaintiffs' request, should order the defendant to use computer-assisted review to respond to plaintiffs' document requests.

The Roitblatt, Kershaw, Oot article noted that "[t]he level of agreement among human reviewers is not strikingly high," around 70-75%. They identify two sources for this variability: fatigue ("A document that they [the reviewers] might have categorized as responsive when they were more attentive might then be categorized [when the reviewer is distracted or fatigued] as non-responsive or vice versa."), and differences in "strategic judgment." Id. at 77-78. Another study found that responsiveness "is fairly well defined, and that disagreements among assessors are largely attributable to human error," with only 5% of reviewer disagreement attributable to borderline or questionable issues as to relevance. Maura R. Grossman & Gordon V. Cormack, Inconsistent Assessment of (continued...)

1.44 · Ethical Issues and Practicing Law
Likewise, Wachtell, Lipton, Rosen & Katz litigation counsel Maura Grossman and University of Waterloo professor Gordon Cormack, studied data from the Text Retrieval Conference Legal Track (TREC) and concluded that: "[T]he myth that exhaustive manual review is the most effective – and therefore the most defensible – approach to document review is strongly refuted. Technology-assisted review can (and does) yield more accurate results than exhaustive manual review, with much lower effort." Maura R. Grossman & Gordon V. Cormack, Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review, Rich. J.L. & Tech., Spring 2011, at 48. The technology-assisted reviews in the Grossman-Cormack article also demonstrated significant cost savings over manual review: "The technology-assisted reviews require, on average, human review of only 1.9% of the documents, a fifty-fold savings over exhaustive manual review." Id. at 43.

Because of the volume of ESI, lawyers frequently have turned to keyword searches to cull email (or other ESI) down to a more manageable volume for further manual review. Keywords have a place in production of ESI – indeed, the parties here used keyword searches (with Boolean connectors) to find documents for the expanded seed set to train the predictive coding software. In too many cases, however, the way lawyers choose keywords is the equivalent of the

\(\ldots\text{continued}\)


Grossman and Cormack also note that "not all technology-assisted reviews . . . are created equal" and that future studies will be needed to "address which technology-assisted review process(es) will improve most on manual review." Id.
child's game of "Go Fish." The requesting party guesses which keywords might produce evidence to support its case without having much, if any, knowledge of the responding party's "cards" (i.e., the terminology used by the responding party's custodians). Indeed, the responding party's counsel often does not know what is in its own client's "cards."

Another problem with keywords is that they often are over-inclusive, that is, they find responsive documents but also large numbers of irrelevant documents. In this case, for example, a keyword search for "training" resulted in 165,208 hits; Da Silva Moore's name resulted in 201,179 hits; "bonus" resulted in 40,756 hits; "compensation" resulted in 55,602 hits; and "diversity" resulted in 38,315 hits. (Dkt. No. 92: 2/17/12 ESI Protocol Ex. A.) If MSL had to manually review all of the keyword hits, many of which would not be relevant (i.e., would be false positives), it would be quite costly.

Moreover, keyword searches usually are not very effective. In 1985, scholars David Blair and M. Maron collected 40,000 documents from a Bay Area Rapid Transit accident, and instructed experienced attorney and paralegal searchers to use keywords and other review techniques to retrieve at least 75% of the documents relevant to 51 document requests. David L. Blair & M. E. Maron, An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System, 28 Comm. ACM 289 (1985). Searchers believed they met the goals, but their average recall was just 20%. Id. This result has been replicated in the TREC Legal Track studies over the past few years.

See Ralph C. Losey, "Child's Game of 'Go Fish' is a Poor Model for e-Discovery Search," in Adventures in Electronic Discovery 209-10 (2011).

This Opinion should serve as a wake-up call to the Bar in this District about the need for careful thought, quality control, testing, and cooperation with opposing counsel in designing search terms or "keywords" to be used to produce emails or other electronically stored information ("ESI").

Electronic discovery requires cooperation between opposing counsel and transparency in all aspects of preservation and production of ESI. Moreover, where counsel are using keyword searches for retrieval of ESI, they at a minimum must carefully craft the appropriate keywords, with input from the ESI's custodians as to the words and abbreviations they use, and the proposed methodology must be quality control tested to assure accuracy in retrieval and elimination of "false positives." It is time that the Bar – even those lawyers who did not come of age in the computer era – understand this.


Computer-assisted review appears to be better than the available alternatives, and thus should be used in appropriate cases. While this Court recognizes that computer-assisted review is not perfect, the Federal Rules of Civil Procedure do not require perfection. See, e.g., Pension Comm. of Univ. of Montreal Pension Plan v. Banc of Am. Sec., 685 F. Supp. 2d 456, 461 (S.D.N.Y.)
2010). Courts and litigants must be cognizant of the aim of Rule 1, to "secure the just, speedy, and inexpensive determination" of lawsuits. Fed. R. Civ. P. 1. That goal is further reinforced by the proportionality doctrine set forth in Rule 26(b)(2)(C), which provides that:

On motion or on its own, the court must limit the frequency or extent of discovery otherwise allowed by these rules or by local rule if it determines that:

(i) the discovery sought is unreasonably cumulative or duplicative, or can be obtained from some other source that is more convenient, less burdensome, or less expensive;

(ii) the party seeking discovery has had ample opportunity to obtain the information by discovery in the action; or

(iii) the burden or expense of the proposed discovery outweighs its likely benefit, considering the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the action, and the importance of the discovery in resolving the issues.


In this case, the Court determined that the use of predictive coding was appropriate considering: (1) the parties' agreement, (2) the vast amount of ESI to be reviewed (over three million documents), (3) the superiority of computer-assisted review to the available alternatives (i.e., linear manual review or keyword searches), (4) the need for cost effectiveness and proportionality under Rule 26(b)(2)(C), and (5) the transparent process proposed by MSL.

This Court was one of the early signatories to The Sedona Conference Cooperation Proclamation, and has stated that "the best solution in the entire area of electronic discovery is cooperation among counsel. This Court strongly endorses The Sedona Conference Proclamation (available at www.TheSedonaConference.org)." William A. Gross Constr. Assocs., Inc. v. Am.
Mrs. Mut. Ins. Co., 256 F.R.D. at 136. An important aspect of cooperation is transparency in the discovery process. MSL's transparency in its proposed ESI search protocol made it easier for the Court to approve the use of predictive coding. As discussed above on page 10, MSL confirmed that "[a]ll of the documents that are reviewed as a function of the seed set, whether [they] are ultimately coded relevant or irrelevant, aside from privilege, will be turned over to" plaintiffs. (Dkt. No. 88: 2/8/12 Conf. Tr. at 73; see also 2/17/12 ESI Protocol at 14: "MSL will provide Plaintiffs' counsel with all of the non-privileged documents and will provide, to the extent applicable, the issue tag(s) coded for each document . . . . If necessary, counsel will meet and confer to attempt to resolve any disagreements regarding the coding applied to the documents in the seed set.") While not all experienced ESI counsel believe it necessary to be as transparent as MSL was willing to be, such transparency allows the opposing counsel (and the Court) to be more comfortable with computer-assisted review, reducing fears about the so-called "black box" of the technology.14 This Court highly recommends that counsel in future cases be willing to at least discuss, if not agree to, such transparency in the computer-assisted review process.

Several other lessons for the future can be derived from the Court's resolution of the ESI discovery disputes in this case.

First, it is unlikely that courts will be able to determine or approve a party's proposal as to when review and production can stop until the computer-assisted review software has been trained and the results are quality control verified. Only at that point can the parties and the Court

14 It also avoids the GIGO problem, i.e., garbage in, garbage out.

G/A/P/D/A, SILVA MOORE • ESI
see where there is a clear drop off from highly relevant to marginally relevant to not likely to be relevant documents. While cost is a factor under Rule 26(b)(2)(C), it cannot be considered in isolation from the results of the predictive coding process and the amount at issue in the litigation.

Second, staging of discovery by starting with the most likely to be relevant sources (including custodians), without prejudice to the requesting party seeking more after conclusion of that first stage review, is a way to control discovery costs. If staging requires a longer discovery period, most judges should be willing to grant such an extension. (This Judge runs a self-proclaimed "rocket docket," but informed the parties here of the Court's willingness to extend the discovery cutoff if necessary to allow the staging of custodians and other ESI sources.)

Third, in many cases requesting counsel's client has knowledge of the producing party's records, either because of an employment relationship as here or because of other dealings between the parties (e.g., contractual or other business relationships). It is surprising that in many cases counsel do not appear to have sought and utilized their client's knowledge about the opposing party's custodians and document sources. Similarly, counsel for the producing party often is not sufficiently knowledgeable about their own client's custodians and business terminology. Another way to phrase cooperation is "strategic proactive disclosure of information," i.e., if you are knowledgeable about and tell the other side who your key custodians are and how you propose to search for the requested documents, opposing counsel and the Court are more apt to agree to your approach (at least as phase one without prejudice).

Fourth, the Court found it very helpful that the parties' ediscovery vendors were present and spoke at the court hearings where the ESI Protocol was discussed. (At ediscovery
programs, this is sometimes jokingly referred to as "bring your geek to court day.") Even where as here counsel is very familiar with ESI issues, it is very helpful to have the parties' ediscovery vendors (or in-house IT personnel or in-house ediscovery counsel) present at court conferences where ESI issues are being discussed. It also is important for the vendors and/or knowledgeable counsel to be able to explain complicated ediscovery concepts in ways that make it easily understandable to judges who may not be tech-savvy.

CONCLUSION

This Opinion appears to be the first in which a Court has approved of the use of computer-assisted review. That does not mean computer-assisted review must be used in all cases, or that the exact ESI protocol approved here will be appropriate in all future cases that utilize computer-assisted review. Nor does this Opinion endorse any vendor (the Court was very careful not to mention the names of the parties' vendors in the body of this Opinion, although it is revealed in the attached ESI Protocol), nor any particular computer-assisted review tool. What the Bar should take away from this Opinion is that computer-assisted review is an available tool and should be seriously considered for use in large-data-volume cases where it may save the producing party (or both parties) significant amounts of legal fees in document review. Counsel no longer have to worry about being the "first" or "guinea pig" for judicial acceptance of computer-assisted review. As with keywords or any other technological solution to ediscovery, counsel must design an appropriate process, including use of available technology, with appropriate quality control testing, to review
and produce relevant ESI while adhering to Rule 1 and Rule 26(b)(2)(C) proportionality. Computer-assisted review now can be considered judicially-approved for use in appropriate cases.

SO ORDERED.

Dated: New York, New York
February 24, 2012

Andrew J. Peck
United States Magistrate Judge

Copies by ECF to: All Counsel
Judge Andrew L. Carter, Jr.
EXHIBIT
A. Scope

1. **General.** The procedures and protocols outlined herein govern the production of electronically stored information ("ESI") by MSLGROUP Americas, Inc. ("MSL") during the pendency of this litigation. The parties to this protocol will take reasonable steps to comply with this agreed-upon protocol for the production of documents and information existing in electronic format. Nothing in this protocol will be interpreted to require disclosure of documents or information protected from disclosure by the attorney-client privilege, work-product product doctrine or any other applicable privilege or immunity. It is Plaintiffs’ position that nothing in this protocol will be interpreted to waive Plaintiffs’ right to object to this protocol as portions of it were mandated by the Court over Plaintiffs’ objections, including Plaintiffs’ objections to the predictive coding methodology proposed by MSL.

2. **Limitations and No-Waiver.** This protocol provides a general framework for the production of ESI on a going forward basis. The Parties and their attorneys do not intend by this protocol to waive their rights to the attorney work-product privilege, except as specifically required herein, and any such waiver shall be strictly and narrowly construed and shall not
extend to other matters or information not specifically described herein. All Parties preserve their attorney client privileges and other privileges and there is no intent by the protocol, or the production of documents pursuant to the protocol, to in any way waive or weaken these privileges. All documents produced hereunder are fully protected and covered by the Parties’ confidentiality and clawback agreements and orders of the Court effectuating same.


B. ESI Preservation

1. MSL has issued litigation notices to designated employees on February 10, 2010, March 14, 2011 and June 9, 2011.

C. Sources

1. The Parties have identified the following sources of potentially discoverable ESI at MSL. Phase I sources will be addressed first, and Phase II sources will be addressed after Phase I source searches are complete. Sources marked as “N/A” will not be searched by the Parties.

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Description</th>
<th>Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>a EMC SourceOne Archive</td>
<td>Archiving System used to capture and store all incoming and outbound e-mails and selected instant message conversations saved through IBM Sametime (see below).</td>
<td>I</td>
</tr>
<tr>
<td>b Lotus Notes E-mail</td>
<td>Active corporate system that provides e-mail communication and calendaring functions</td>
<td>N/A</td>
</tr>
<tr>
<td>c GroupWise E-mail</td>
<td>Legacy corporate system that provided e-mail communication and calendaring functions.</td>
<td>N/A</td>
</tr>
<tr>
<td>d IBM Sametime</td>
<td>Lotus Notes Instant Messaging and collaboration application.</td>
<td>N/A</td>
</tr>
<tr>
<td>e Home Directories</td>
<td>Personal network storage locations on the file server(s)</td>
<td>II</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>f</strong></td>
<td>Shared Folders</td>
<td>Shared network storage locations on the file server(s) that are accessible by individual users, groups of users or entire departments. (With the exception of the following Human Resources shared folders which will be in Phase I: Corporate HR, North America HR and New York HR.)</td>
</tr>
<tr>
<td><strong>g</strong></td>
<td>Database Servers</td>
<td>Backend databases (e.g. Oracle, SQL, MySQL) used to store information for front end applications or other purposes.</td>
</tr>
<tr>
<td><strong>h</strong></td>
<td>Halogen Software</td>
<td>Performance management program provided by Halogen to conduct performance evaluations.</td>
</tr>
<tr>
<td><strong>i</strong></td>
<td>Noovoo</td>
<td>Corporate Intranet site.</td>
</tr>
<tr>
<td><strong>j</strong></td>
<td>Corporate Feedback</td>
<td>E-mail addresses that employees may utilize to provide the company with comments, suggestions and overall feedback.</td>
</tr>
<tr>
<td><strong>k</strong></td>
<td>Hyperion Financial Management (&quot;HFM&quot;)</td>
<td>Oracle application that offers global financial consolidation, reporting and analysis.</td>
</tr>
<tr>
<td><strong>l</strong></td>
<td>Vurv/Taleo</td>
<td>Talent recruitment software.</td>
</tr>
<tr>
<td><strong>m</strong></td>
<td>ServiceNow</td>
<td>Help Desk application used to track employee computer related requests.</td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>PeopleSoft</td>
<td>Human resources information management system.</td>
</tr>
<tr>
<td><strong>o</strong></td>
<td>PRISM</td>
<td>PeopleSoft component used for time and billing management.</td>
</tr>
<tr>
<td><strong>p</strong></td>
<td>Portal</td>
<td>A project based portal provided through Oracle/BEA Systems.</td>
</tr>
<tr>
<td><strong>q</strong></td>
<td>Desktops/Laptops</td>
<td>Fixed and portable computers provided to employees to perform work related activities. (With the exception of 2 desktop/laptop hard drives for which MSL will collect and analyze the data to determine the level of duplication as compared to the EMC SourceOne Archive. The parties will meet and confer regarding the selection of the two custodians.)</td>
</tr>
<tr>
<td><strong>r</strong></td>
<td>Publicis Benefits Connection</td>
<td>Web based site that maintains information about employee benefits and related information.</td>
</tr>
<tr>
<td><strong>s</strong></td>
<td>GEAR S</td>
<td>Employee expense reporting system.</td>
</tr>
<tr>
<td><strong>t</strong></td>
<td>MS&amp;L City</td>
<td>Former corporate Intranet.</td>
</tr>
<tr>
<td><strong>u</strong></td>
<td>Adium</td>
<td>Application which aggregates instant messages.</td>
</tr>
<tr>
<td><strong>v</strong></td>
<td>Pidgin</td>
<td>Application which aggregates instant message.</td>
</tr>
<tr>
<td><strong>w</strong></td>
<td>IBM Lotus</td>
<td>Mobile device synchronization and security system.</td>
</tr>
<tr>
<td>Traveler and MobileIron</td>
<td>Portable PDAs, smart phones, tablets used for communication.</td>
<td>N/A</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------------------</td>
<td>-----</td>
</tr>
<tr>
<td>y Mobile Communication Devices</td>
<td>Social media and collaboration portal.</td>
<td>N/A</td>
</tr>
<tr>
<td>z Yammer</td>
<td>Web-based customer relationship management application.</td>
<td>N/A</td>
</tr>
<tr>
<td>aa SalesForce.com</td>
<td>Portable storage media, external hard drives, thumb drives, etc. used to store copies of work related ESI.</td>
<td>N/A</td>
</tr>
</tbody>
</table>

a. **EMC SourceOne** - MSL uses SourceOne, an EMC e-mail archiving system that captures and stores all e-mail messages that pass through the corporate e-mail system. In addition, if a user chooses to save an instant messaging chat conversation from IBM Sametime (referenced below), that too would be archived in SourceOne. Defendant MSL also acknowledges that calendar items are regularly ingested into the SourceOne system. The parties have agreed that this data source will be handled as outlined in section E below.

b. **Lotus Notes E-mail** - MSL currently maintains multiple Lotus Notes Domino servers in various data centers around the world. All e-mail communication and calendar items are journaled in real time to the EMC SourceOne archive. The parties have agreed to not collect any information from this data source at this time.

c. **GroupWise E-mail** – Prior to the implementation of the Lotus Notes environment, GroupWise was used for all e-mail and calendar functionality. Before the decommissioning of the GroupWise servers, MSL created backup tapes of all servers that housed the GroupWise e-mail databases. The parties have agreed to not collect any information from this data source at this time.

d. **IBM Sametime** – MSL provides custodians with the ability to have real time chat conversations via the IBM Sametime application that is part of the Lotus Notes suite of products.
e. **Home Directories** – Custodians with corporate network access at MSL also have a dedicated and secured network storage location where they are able to save files. MSL will collect the home directory data for 2 custodians and analyze the data to determine the level of duplication of documents in this data source against the data contained in the EMC SourceOne archive for the same custodians. (The parties will meet and confer regarding the selection of the two custodians.) The results of the analysis will be provided to Plaintiffs so that a determination can be made by the parties as to whether MSL will include this data source in its production of ESI to Plaintiffs. If so, the parties will attempt to reach an agreement as to the approach used to collect, review and produce responsive and non-privileged documents.

f. **Shared Folders** – Individual employees, groups of employees and entire departments at MSL are given access to shared network storage locations to save and share files. As it relates to the Human Resources related shared folders (i.e., North America HR Drive (10.2 GB), Corporate HR Drive (440 MB), NY HR Drive (1.9 GB), Chicago HR Drive (1.16 GB), Boston HR Drive (43.3 MB), and Atlanta HR Drive (6.64 GB)), MSL will judgmentally review and produce responsive and non-privileged documents from the North America HR Drive, Corporate HR Drive, and NY HR Drive. MSL will produce to Plaintiffs general information regarding the content of other Shared Folders. The parties will meet and confer regarding the information gathered concerning the other Shared Folders and discuss whether any additional Shared Folders should be moved to Phase I.

g. **Database Servers** – MSL has indicated that it does not utilize any database servers, other than those that pertain to the sources outlined above in C, which are likely to contain information relevant to Plaintiffs’ claims.

h. **Halogen Software** – MSL utilizes a third party product, Halogen, for performance management and employee evaluations. The parties will meet and confer in order to
exchange additional information and attempt to reach an agreement as to the scope of data and the approach used to collect, review and produce responsive and non-privileged documents.

i. **Noovoo** – MSL maintains a corporate Intranet site called “Noovoo” where employees are able to access Company-related information. MSL will provide Plaintiffs with any employment-related policies maintained within Noovoo.

j. **Corporate Feedback** – MSL has maintained various e-mail addresses that employees may utilize to provide the company with comments, suggestions and overall feedback. These e-mail addresses include “powerofone@mslworldwide.com”, “poweroftheindividual@mslworldwide.com”, “townhall@mslworldwide.com” and “whatsonyourmind@mslworldwide.com”. The parties have agreed that all responsive and non-privileged ESI will be produced from these e-mail accounts and any other e-mail accounts that fall under this category of information. At present, MSL intends to manually review the contents of each of these e-mail accounts. However, if after collecting the contents of each of the e-mail accounts MSL determines that a manual review would be impractical, the parties will meet and confer as to the approach used to collect, review and produce responsive and non-privileged documents.

k. **Hyperion Financial Management (“HFM”)** – MSL uses an Oracle application called HFM that offers global financial consolidation, reporting and analysis capabilities.

l. **Vurv/Taleo** – Since approximately 2006, MSL used an application known as Vurv as its talent recruitment software. As of August 31, 2011, as a result of Vurv being purchased by Taleo, MSL has been using a similar application by Taleo as its talent recruitment software. The application, which is accessed through MSL’s public website, allows users to search for open positions as well as input information about themselves. To the extent Plaintiffs
contend they were denied any specific positions, they will identify same and the Parties will meet and confer to discuss what, if any, information exists within Veriv/Taleo regarding the identified position. If information exists in Veriv/Taleo or another source regarding these positions, MSL will produce this information, to the extent such information is discoverable.

m. **ServiceNow** – MSL utilizes ServiceNow as its Help Desk application. This system covers a wide variety of requests by employees for computer-related assistance (e.g., troubleshoot incidents, install software, etc.).

n. **PeopleSoft** – MSL utilizes PeopleSoft, an Oracle-based software product, to record employee data such as date of hire, date of termination, promotions, salary increases, transfers, etc. MSL has produced data from this source and will consider producing additional data in response to a specific inquiry from Plaintiffs.

o. **PRISM** – MSL utilizes PRISM for tracking time and billing. It is used primarily to track an employee's billable time. MSL will consider producing additional data in response to a specific inquiry from Plaintiffs.

p. **Portal** – MSL maintains a portal provided through Oracle/BEA Systems. The portal is web-based and is used for light workflow activities (such as reviewing draft documents).

q. **Desktops/Laptops** – MSL provided employees with desktop and/or laptop computers to assist in work related activities. MSL will collect the desktop/laptop hard drive data for 2 custodians and analyze the data to determine the level of duplication of documents in this data source against the data contained in the EMC SourceOne archive for the same custodians. (The parties will meet and confer regarding the selection of the two custodians.) The results of the analysis will be provided to Plaintiffs so that a determination can be made by the parties as to whether MSL will include this data source in its production of ESI to Plaintiffs.
If so, the Parties will attempt to reach an agreement as to the approach used to collect, review and produce responsive and non-privileged documents.

r. **Publicis Benefits Connection** – Plaintiffs understand that MSL provides employees with access to a centralized web based site that provides access to corporate benefits information and other related content.

s. **GEARS** – MSL maintains a centralized web-based expense tracking and reporting system called “GEARS” where users are able to enter expenses and generate reports.

t. **MS&L City** – MSL maintained a corporate web-based intranet prior to migrating to Noovoo.

u. **Adium** – This is a free and open source instant messaging client for Mac OS X users.

v. **Pidgin** – Pidgin is a chat program which lets users log into accounts on multiple chat networks simultaneously. However, the data resides with a third party messaging provider (e.g. AIM, Yahoo!, Google Talk, MSN Messenger, etc.).

w. **IBM Lotus Traveler and MobileIron** – MSL maintains these systems for e-mail device sync and security features for employees’ mobile devices, including Blackberry devices, iPhones, iPads, Android phones, and Android tablets.

x. **Mobile Communication Devices** – MSL provides mobile devices and/or connectivity including Blackberry devices, iPhones, iPads, Android phones, and Android tablets to designated employees.

y. **Yammer** – This is an instant messaging application hosted externally, used for approximately one year in or around 2008 through 2009.

z. **SalesForce.com** – This is a web-based customer relationship management application but it was not widely used.
aa. Removable Storage Devices – MSL does not restrict authorized employees from using removable storage devices.

D. Custodians

1. The Parties agree that MSL will search the e-mail accounts of the following individuals as they exist on MSL's EMC SourceOne archive. (Except where a date range is noted, the custodian's entire e-mail account was collected from the archive.)

<table>
<thead>
<tr>
<th>Custodian Name</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Lund, Wendy</td>
<td>Executive VP of Global Client and Business Development</td>
</tr>
<tr>
<td>2. Fite, Vicki</td>
<td>Managing Director, MSL Los Angeles</td>
</tr>
<tr>
<td>3. Wilson, Renee</td>
<td>President, NE Region, Managing Director NY</td>
</tr>
<tr>
<td>4. Brennan, Nancy (1/1/08 to 5/31/08)</td>
<td>SVP/Director Corporate Branding</td>
</tr>
<tr>
<td>5. Lilien (Lillien, Kashanian), Tara</td>
<td>SVP, North America Human Resources</td>
</tr>
<tr>
<td>6. Miller, Peter</td>
<td>Executive Vice President, CFO</td>
</tr>
<tr>
<td>7. Masini, Rita</td>
<td>Chief Talent Officer</td>
</tr>
<tr>
<td>8. Tsokanos, Jim</td>
<td>President of the Americas</td>
</tr>
<tr>
<td>9. Da Silva Moore, Monique</td>
<td>Director Healthcare Practice, Global</td>
</tr>
<tr>
<td>10. O’Kane, Jeanine (2/8/10 to 2/24/11)</td>
<td>Director of Healthcare North America</td>
</tr>
<tr>
<td>11. Perlman, Carol</td>
<td>Senior VP</td>
</tr>
<tr>
<td>12. Mayers, Laurie</td>
<td>SVP MS&amp;L Digital</td>
</tr>
<tr>
<td>13. Wilkinson, Kate</td>
<td>Account Executive</td>
</tr>
<tr>
<td>14. Curran, Joel (5/1/08 to 5/31/10)</td>
<td>Managing Director MSL Chicago</td>
</tr>
<tr>
<td>15. Shapiro, Maury</td>
<td>North American CFO</td>
</tr>
<tr>
<td>16. Baskin, Rob (1/1/08 to 12/31/08)</td>
<td>Managing Director</td>
</tr>
<tr>
<td>17. Pierce, Heather</td>
<td>VP</td>
</tr>
<tr>
<td>18. Branim, Jud (1/1/08 to 1/31/10)</td>
<td>Managing Director, MS&amp;L Digital</td>
</tr>
<tr>
<td>19. McDonough, Jenni (1/1/08 to 12/31/08)</td>
<td>VP, Director of Human Resources</td>
</tr>
<tr>
<td>20. Hannaford, Donald (1/1/08 to 3/1/08)</td>
<td>Managing Director</td>
</tr>
<tr>
<td>21. Orr, Bill (1/1/08 to 2/24/11)</td>
<td>Managing Director</td>
</tr>
<tr>
<td>22. Dhillon, Neil (9/8/08 to 5/31/10)</td>
<td>Managing Director MSL Washington DC</td>
</tr>
<tr>
<td>23. Hubbard, Zaneta</td>
<td>Account Supervisor</td>
</tr>
<tr>
<td>24. Morgan, Valerie (1/1/08 to 2/24/11)</td>
<td>HR Director</td>
</tr>
<tr>
<td>25. Daversa, Kristin (1/1/08 to 2/24/11)</td>
<td>HR Director</td>
</tr>
<tr>
<td>26. Vosk, Lindsey (1/1/08 to 2/24/11)</td>
<td>HR Manager</td>
</tr>
<tr>
<td>27. Carberry, Joe (1/1/08 to 2/24/11)</td>
<td>President, Western Region</td>
</tr>
<tr>
<td>28. Sheffield, Julie (1/1/08 to 2/24/11)</td>
<td>HR/Recruiting Associate</td>
</tr>
<tr>
<td>29. MaryEllen O'Donohue</td>
<td>SVP (2010)</td>
</tr>
</tbody>
</table>
E. Search Methodology

1. General. The Parties have discussed the methodologies or protocols for the search and review of ESI collected from the EMC SourceOne archive and the following is a summary of the Parties’ agreement on the use of Predictive Coding. This section relates solely to the EMC SourceOne data source (hereinafter referred to as the “e-mail collection”).

2. General Overview of Predictive Coding Process. MSL will utilize the Axcelerate software by Recommend to search and review the e-mail collection for production in this case.

The process begins with Jackson Lewis attorneys developing an understanding of the entire e-mail collection while identifying a small number of documents, the initial seed set, that is representative of the categories to be reviewed and coded (relevance, privilege, issue-relation). It is the step when the first seed sets are generated which is done by use of search and analytical tools, including keyword, Boolean and concept search, concept grouping, and, as needed, up to 40 other automatically populated filters available within the Axcelerate system. This assists in the attorneys’ identification of probative documents for each category to be reviewed and coded.

Plaintiffs’ counsel will be provided with preliminary results of MSL’s hit counts using keyword searches to create a high priority relevant seed set, and will be invited to contribute their own proposed keywords. Thereafter, Plaintiffs’ counsel will be provided with the non-privileged keyword hits – both from MSL’s keyword list and Plaintiffs’ keyword list – which were reviewed and coded by MSL. Plaintiffs’ counsel will review the documents produced and promptly provide defense counsel with their own evaluation of the initial coding applied to the documents, including identification of any documents it believes were incorrectly coded. To the

\(^7\) As noted in Paragraphs A(1) and J of this Protocol, Plaintiffs object to the predictive coding methodology proposed by MSL.
extent the parties disagree regarding the coding of a particular document, they will meet and confer in an effort to resolve the dispute prior to contacting the Court for resolution. The irrelevant documents so produced shall be promptly returned after review and analysis by Plaintiffs’ counsel and/or resolution of any disputes by the Court.

The seed sets are then used to begin the Predictive Coding process. Each seed set of documents is applied to its relevant category and starts the software “training” process. The software uses each seed set to identify and prioritize all substantively similar documents over the complete corpus of the e-mail collection. The attorneys then review and code a judgmental sample of at least 500 of the “computer suggested” documents to ensure their proper categorization and to further calibrate the system by recoding documents into their proper categories. Accelerate learns from the new corrected coding and the Predictive Coding process is repeated.

Attorneys representing MSL will have access to the entire e-mail collection to be searched and will lead the computer training, but they will obtain input from Plaintiffs’ counsel during the iterative seed selection and quality control processes and will share the information used to craft the search protocol as further described herein. All non-privileged documents reviewed by MSL during each round of the iterative process (i.e., both documents coded as relevant and irrelevant) will be produced to Plaintiffs’ counsel during the iterative seed set selection process. Plaintiffs’ counsel will review the documents produced and promptly provide defense counsel with its own evaluation of the initial coding applied to the documents, including identification of any documents it believes were incorrectly coded. To the extent the Parties disagree regarding the coding of a particular document, they will meet and confer in an effort to resolve the dispute prior to contacting the Court for resolution. Again, the irrelevant documents
so produced shall be promptly returned after review and analysis by Plaintiffs' counsel and/or resolution of any disputes by the Court.

At the conclusion of the iterative review process, all document predicted by Accelerate to be relevant will be manually reviewed for production. However, depending on the number of documents returned, the relevancy rating of those documents, and the costs incurred during the development of the seed set and iterative reviews, MSL reserves the right to seek appropriate relief from the Court prior to commencing the final manual review.

The accuracy of the search processes, both the systems' functions and the attorney judgments to train the computer, will be tested and quality controlled by both judgmental and statistical sampling. In statistical sampling, a small set of documents is randomly selected from the total corpus of the documents to be tested. The small set is then reviewed and an error rate calculated therefrom. The error rates can then be reliably projected on the total corpus, having a margin of error directly related to the sample size.

3. **Issue Tags.** The parties agree that, to the extent applicable, as part of the seed set training described above, as well as during the iterative review process, all documents categorized as relevant and not privileged, to the extent applicable, also shall be coded with one or more of the following agreed-upon issue tags:

   a. Reorganization.
   b. Promotion/Assignments.
   c. Work/Life Balance.
   d. Termination.
   e. Compensation.
   f. Maternity/Pregnancy.
   g. Complaints/HR.
h. Publicis Groupe/Jurisdiction.

This issue coding will take place during the initial random sample, creation of the seed set and initial and iterative training (see paragraphs 4, 5 and 6 below). This input shall be provided to Plaintiffs' counsel along with the initial document productions. Plaintiffs' counsel shall promptly report any disagreements on classification, and the parties shall discuss these issues in good faith, so that the seed set training may be improved accordingly. This issue-tagging and disclosure shall take place during the described collaborative seed set training process. The disclosures here made by MSL on its issue coding are not required in the final production set.

4. Initial Random Sample. Using the Axcelerate software to generate a random sample of the entire corpus of documents uploaded to the Axcelerate search and review platform, MSL's attorneys will conduct a review of the random sample for relevance and to develop a baseline for calculating recall and precision. To the extent applicable, any relevant documents also will be coded with one or more of the issue tags referenced in paragraph E.3 above. The random sample consists of 2,399 documents, which represents a 95% confidence level with a confidence estimation of plus or minus 2%. The Parties agree to utilize the random sample generated prior to the finalization of this protocol. However, during Plaintiffs' counsel's review of the random sample, they may advise as to whether they believe any of the documents should be coded with one or more of the subsequently added issue codes (i.e., Complaints/HR and Publicis Groupe/Jurisdiction) and will, as discussed above, indicate any disagreement with MSL's classifications.

5. Seed Set.

a. Defendant MSL. To create the initial seed set of documents that will be used to "train" the Axcelerate software as described generally above, MSL primarily utilized keywords listed on Exhibits A and B to this protocol, but also utilized other judgmental analysis
and search techniques designed to locate highly relevant documents, including the Boolean, concept search and other features of Axcelerate. Given the volume of hits for each keyword (Exhibit A), MSL reviewed a sampling of the hits and coded them for relevance as well as for the following eight preliminary issues: (i) Reorganization; (ii) Promotion; (iii) Work/Life Balance; (iv) Termination; (v) Compensation; and (vi) Maternity. Specifically, except for keywords that were proper names, MSL performed several searches within each set of key word hits and reviewed a sample of the hits. The Axcelerate software ranked the hits in order of relevance based on the software’s analytical capabilities and the documents were reviewed in decreasing order of relevance (i.e., each review of the sample of supplemental searches started with the highest ranked documents). Exhibit B identifies the supplemental searches conducted, the number of hits, the number of documents reviewed, the number of documents coded as potentially responsive and general comments regarding the results. In addition, to the extent applicable, documents coded as responsive also were coded with one or more issue tags. MSL will repeat the process outlined above and will include the newly defined issues and newly added custodians. MSL will provide Plaintiffs’ counsel with all of the non-privileged documents and will provide, to the extent applicable, the issue tag(s) coded for each document, as described above. Plaintiffs’ counsel shall promptly review and provide notice as to any documents with which they disagree where they do not understand the coding. If necessary, counsel will meet and confer to attempt to resolve any disagreements regarding the coding applied to the documents in this seed set.

b. Plaintiffs. To help create the initial seed set of documents that will be used to “train” the Axcelerate software, Plaintiffs provided a list of potential key words to MSL. MSL provided Plaintiffs with a hit list for their proposed key words. This process was repeated twice with the hit list for Plaintiffs’ most recent set of keywords attached as Exhibit C. MSL
will review 4,000 randomly sampled documents from Plaintiffs’ supplemental list of key words to be coded for relevance and issue tags. MSL will provide Plaintiffs’ counsel with all non-privileged documents and will provide, to the extent applicable, the issue tag(s) coded for each document. Plaintiffs’ counsel shall promptly review and provide notice as to any documents with which they disagree with or where they do not understand the coding. If necessary, the Parties’ counsel will meet and confer to attempt to resolve any disagreements regarding the coding applied to the documents in this seed set.

c. Judgmental Sampling. In addition to the above, a number of targeted searches were conducted by MSL in an effort to locate documents responsive to several of Plaintiffs’ specific discovery requests. Approximately 578 documents have already been coded as responsive and produced to Plaintiffs. In addition, several judgmental searches were conducted which resulted in approximately 300 documents initially being coded as responsive and several thousand additional documents coded as irrelevant. The documents coded as relevant and non-privileged also will be reviewed by Plaintiffs’ counsel and, subject to their feedback, included in the seed set. An explanation shall be provided by MSL’s attorneys for the basis of the bulk tagging of irrelevant documents (primarily electronic periodicals and newsletters that were excluded in the same manner as spam junk mail is excluded). The explanation shall include the types of documents bulk tagged as irrelevant as well as the process used to identify those types of documents and other similar documents that were bulk tagged as irrelevant.

6. Initial And Iterative Training. Following the creation of the first seed set, the Axcelerate software will review the entire data set to identify other potentially relevant documents. MSL will then review and tag a judgmental based sample, consisting of a minimum of 500 documents, including all documents ranked as highly relevant or hot, of the new
“Computer Suggested” documents, which were suggested by the Accelerate software. MSL’s attorneys shall act in consultation with the Accelerate software experts to make a reasonable, good faith effort to select documents in the judgmental sample that will serve to enhance and increase the accuracy of the predictive coding functions. The results of this first iteration, both the documents newly coded as relevant and not relevant for particular issue code or codes, will be provided to Plaintiffs’ counsel for review and comment. (All documents produced by the parties herein to each other, including, without limitation, these small seed set development productions, shall be made under the Confidentiality Stipulation in this matter as well as any clawback agreement that shall be reduced to an order acceptable to the Court. Any documents marked as irrelevant shall be returned to counsel for MSL at the conclusion of the iterative training phase, unless the relevancy of any documents are disputed, in which case they may be submitted to the Court for review.)

Upon completion of the initial review, and any related meet and confer sessions and agreed upon coding corrections, the Accelerate software will be run again over the entire data set for suggestions on other potentially relevant documents following the same procedures as the first iteration. The purpose of this second and any subsequent iterations of the Predictive Coding process will be to further refine and improve the accuracy of the predictions on relevance and various other codes. The results of the second iteration shall be reviewed and new coding shared with Plaintiffs’ counsel as described for the first iteration. This process shall be repeated five more times, for a total of seven iterations, unless the change in the total number of relevant documents predicted by the system as a result of a new iteration, as compared to the last iteration, is less than five percent (5%), and no new documents are found that are predicted to be hot (aka highly relevant), at which point MSL shall have the discretion to stop the iterative process and begin the final review as next described. If more than 40,000 documents are
returned in the final iteration, then MSL reserves the right to apply to the Court for relief and limitations in its review obligations hereunder. Plaintiffs reserve the right, at all times, to challenge the accuracy and reliability of the predictive coding process and the right to apply to the Court for a review of the process.

7. Final Search and Production. All of the documents predicted to be relevant in the final iteration described in paragraph six above will be reviewed by MSL, unless it applies to the court for relief hereunder. All documents found by MSL's review to be relevant and non-privileged documents will be promptly produced to Plaintiffs. If more than 40,000 documents are included in the final iteration, then MSL reserves its right to seek payment from Plaintiffs for all reasonable costs and fees MSL incurred related to the attorney review and production of more 40,000 documents. This provision is not intended as a waiver by MSL to also seek an award of all discovery costs incurred, including costs related to the first 40,000 documents, at the conclusion of this action under 28 U.S.C. §1920(4) and Rule 54(d)(1) Federal Rules of Civil Procedure. [Plaintiffs object to the inclusion of MSL's proposed cost-shifting language as premature and argumentative, in contravention of the Court's January 4, 2012 and February 8, 2012 orders. Plaintiffs believe costs should be subject to a separate hearing.]

8. Quality Control by Random Sample of Irrelevant Documents. In addition, at the conclusion of this search protocol development process described above, and before the final search and production described in Paragraph 7 above, MSL will review a random sample of 2,399 documents contained in the remainder of the database that were excluded as irrelevant. The results of this review, both the documents coded as relevant and not relevant, but not privileged, will be provided to Plaintiffs' counsel for review. (Any documents initially coded as “not relevant” will be provided subject to the Confidentiality Stipulation and any clawback agreements entered in this matter will be returned to counsel for MSL within 60 days of their
production.) The purpose for this review is to allow calculation of the approximate degree of recall and precision of the search and review process used. If Plaintiffs object to the proposed review based on the random sample quality control results, or any other valid objection, they shall provide MSL with written notice thereof within five days of the receipt of the random sample. The parties shall then meet and confer in good faith to resolve any difficulties, and failing that shall apply to the Court for relief. MSL shall not be required to proceed with the final search and review described in Paragraph 7 above unless and until objections raised by Plaintiffs have been adjudicated by the Court or resolved by written agreement of the Parties.

F. Costs

1. MSL proposes to limit the costs of its final review and production of responsive ESI from the MSL email collection to an additional $200,000, above and beyond the approximately $350,000 it has already paid or is anticipated to pay in e-discovery related activities as previously described and disclosed to Plaintiffs. Specifically, although MSL potentially will conduct and pay for review of more than 40,000 documents, if that is required under the predictive coding process described in paragraphs 6 and 7 above, MSL reserves its right to seek relief from the Court (e.g., a cost shifting award and/or ruling that MSL need to review more than a specified number of documents) pursuant to the principles of proportionality. See Rule 1, Rule 26(b)(2)(C), Rule 26(b)(2)(B), and Rule 26(g), Federal Rules of Civil Procedure; Commentary on Proportionality in Electronic Discovery, 11 SEDONA CONF. J. 289 (2010); Oot, et al, Mandating Reasonableness in a Reasonable Inquiry, Denver University Law Review, 87:2, 522-559 (2010); Also see Rule 403 of the Federal Evidence Code (inadmissibility of cumulative evidence). [Plaintiffs object to the inclusion of MSL’s proposed cost-shifting language as premature and argumentative, in contravention of the Court’s January 4, 2012]
and February 8, 2012 orders. Plaintiffs believe costs should be subject to a separate hearing.

2. Plaintiffs agree to bear all of the costs associated with their compliance with the terms of this protocol and with the receipt and review of ESI produced hereunder including the costs associated with its ESI experts at DOAR Litigation Consulting who will be involved with Plaintiffs in all aspects of this ESI protocol. Plaintiffs propose that MSL bear all of the costs associated with its obligations under the terms of this protocol and do not agree to limit the amount of information subject to the review and production of ESI by MSL.

G. Format of Production For Documents Produced From Accelerate

1. TIFF/Native File Format Production. Documents will be produced as single-page TIFF images with corresponding multi-page text and necessary load files. The load files will include an image load file as well as a metadata (.DAT) file with the metadata fields identified on Exhibit D. Defendant MSL will produce spreadsheets (.xls files) and PowerPoint presentations (.ppt files) in native form as well as any documents that cannot be converted to TIFF format (e.g., audio or video files, such as mp3s, wavds, megs, etc.). In addition, for any redacted documents that are produced, the documents’ metadata fields will be redacted where required. For the production of ESI from non-email sources, the parties will meet and confer to attempt to reach an agreement of the format of production.

2. Appearance. Subject to appropriate redaction, each document’s electronic image will convey the same information and image as the original document. Documents that present imaging or formatting problems will be promptly identified and the parties will meet and confer in an attempt to resolve the problems.

3. Document Numbering. Each page of a produced document will have a legible, unique page identifier “Bates Number” electronically “burned” onto the image at a location
that does not obliterate, conceal or interfere with any information from the source document. The Bates Number for each page of each document will be created so as to identify the producing party and the document number. In the case of materials redacted in accordance with applicable law or confidential materials contemplated in any Confidentiality Stipulation entered into by the parties, a designation may be “burned” onto the document’s image at a location that does not obliterate or obscure any information from the source document.

4. **Production Media.** The producing party will produce documents on readily accessible, computer or electronic media as the parties may hereafter agree upon, including CD-ROM, DVD, external hard drive (with standard PC compatible interface), (the “Production Media”). Each piece of Production Media will be assigned a production number or other unique identifying label corresponding to the date of the production of documents on the Production Media (e.g., “Defendant MSL Production April 1, 2012”) as well as the sequence of the material in that production (e.g. “-001”, “-002”). For example, if the production comprises document images on three DVDs, the producing party may label each DVD in the following manner “Defendant MSL Production April 1, 2012”, “Defendant MSL Production April 1, 2012-002”, “Defendant MSL Production April 1, 2012-003.” Additional information that will be identified on the physical Production Media includes: (1) text referencing that it was produced in *da Silva Moore v. Publicis Groupe SA, et al.*; and (2) the Bates Number range of the materials contained on the Production Media. Further, any replacement Production Media will cross-reference the original Production Media and clearly identify that it is a replacement and cross-reference the Bates Number range that is being replaced.

5. **Write Protection and Preservation.** All computer media that is capable of write-protection should be write-protected before production.
6. **Inadvertent Disclosures.** The terms of the Parties' Clawback Agreement and Court Order shall apply to this protocol.

7. **Duplicate Production Not Required.** A party producing data in electronic form need not produce the same document in paper format.

**H. Timing.**

1. To the extent a timeframe is not specifically outlined herein, the parties will use their reasonable efforts to produce ESI in a timely manner consistent with the Court's discovery schedule.

2. The parties will produce ESI on a rolling basis.

**I. General Provisions.**

1. Any practice or procedure set forth herein may be varied by agreement of the parties, and first will be confirmed in writing, where such variance is deemed appropriate to facilitate the timely and economical exchange of electronic data.

2. Should any party subsequently determine it cannot in good faith proceed as required by this protocol, the parties will meet and confer to resolve any dispute before seeking Court intervention.

3. The Parties agree that e-discovery will be conducted in phases and, at the conclusion of the search process described in Section E above, the Parties will meet and confer regarding whether further searches of additional custodians and/or the Phase II sources is warranted and/or reasonable. If agreement cannot be reached, either party may seek relief from the Court.
J. Plaintiffs’ Objection

1. Plaintiffs object to this ESI Protocol in its entirety. Plaintiffs submitted their own proposed ESI Protocol to the Court, but it was largely rejected. The Court then ordered the parties to submit a joint ESI Protocol reflecting the Court’s rulings. Accordingly, Plaintiffs jointly submit this ESI Protocol with MSL, but reserve the right to object to its use in this case.

This protocol may be executed in counterparts. Each counterpart, when so executed, will be deemed and original, and will constitute the same instrument.

By: JANETTE WIPPER, ESQ.
DEEPIKA BAINS, ESQ.
SIHAM NURHUSSIEIN, ESQ.
SANFORD WITTELS & HEISLER, LLP
Attorneys for Plaintiffs and Class
555 Montgomery Street, Ste. 1206
San Francisco, CA 94111
Telephone: (415) 391-6900

By: BRETT M. ANDERS, ESQ.
VICTORIA WOODIN CHAVEY, ESQ.
JEFFREY W. BRECHER, ESQ.
JACKSON LEWIS LLP
Attorneys for Defendant MSLGROUP
58 South Service Road, Suite 410
Melville, NY 11747
Telephone: (631) 247-0404

Date: ___________________________, 2012

So Ordered: ____________________________
United States Magistrate Judge

Date: ___________________________, 2012

BY ECF
ELECTRONIC DATA, ELECTRONIC SEARCHING, INADVERTENT PRODUCTION OF PRIVILEGED DATA: A PERFECT STORM

WHY ATTORNEYS ARE BEING FORCED TO RECOGNIZE THAT SEARCHING ELECTRONICALLY STORED INFORMATION IS AN EXPERT FUNCTION

Donald Wochna

I. Introduction ................................................................. 843
II. The Challenge of Manual Review: Volume ................................. 845
III. The Cost of Manual Review of Electronically Stored Information ....................................................... 849
IV. Nature of Electronic Searching: Expert Language and Expert Function ................................................................. 851
V. Analysis of Nature of Keyword Searching: Expert Function Combining Linguistics, Statistics, and Computer Technology ................................................................. 856
VI. Searching Electronically Stored Information Is an Expert Process ................................................................. 861
VII. Challenging Electronic Search Processes as Reasonable and Federal Evidence Rule 502(b) ................................................................. 862
VIII. Defending Electronic Search Processes as Reasonable and Federal Evidence Rule 502(b) ................................................................. 865
IX. Conclusion ........................................................................ 868

I. INTRODUCTION

Recent case law,¹ changes in civil procedural rules,² and the dramatic increase in the volume of electronically stored information³

2. See Fed. R. Civ. P. 26(f) (“Under Rule 26(f), parties must sit down together at an early ‘meet and confer’ conference to discuss a range of issues involving electronically stored information.”).
have combined to form a “perfect storm” in which to trap unwary attorneys into potentially committing malpractice. Faced with enormous volumes of client data that must be reviewed for privilege and unacceptably high costs of manual review, many attorneys are relying upon electronic searches to identify privileged documents within large client data sets. Recent case law discussed herein analyzes this type of electronic searching and concludes that it is an expert function. Attorneys who fail to treat electronic searching as an expert function may be unable to defend their electronic search protocols when challenged, and, as a consequence thereof, may incur sanctions, including loss of attorney-client privilege protection for client documents inadvertently produced in litigation. This article examines the case law analyzing electronic searching of client data and concludes that treating electronic searching as an expert function is consistent with the

information. Such a conference is intended to be broad in scope and to cover the gamut of preservation, scope, formatting, and accessibility issues.”

3. George L. Paul & Jason B. Baron, Information Inflation: Can the Legal System Adapt?, 13 RICH. J. & TECH. 10, *12-13 (2007). Probably close to 100 billion e-mails are sent daily, with approximately 30 billion e-mails created or received by federal government agencies each year. The amount of stored information continues to grow exponentially. Perhaps more easily grasped, the amount of information in business has increased by thousands, if not tens of thousands of times in the last few years.

4. The Sedona Conference, The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery, 8 SEDONA CONF. J. 189, 192 (2007) [hereinafter Best Practices] (“Discovery of relevant information gathered about a topic in dispute is at the core of the litigation process. However, the advent of e-discovery is causing a rapid transformation in how that information is gathered. While discovery disputes are not new, the huge volume of available electronically stored information poses unique challenges.”).

5. See O’Keefe, 537 F. Supp. 2d at 24. Whether search terms or “keywords” will yield the information sought is a complicated question involving the interplay, at least, of the sciences of computer technology, statistics, and linguistics. . . . Given this complexity, for lawyers and judges to dare opine that a certain search term or terms would be more likely to produce information than the terms that were used is truly to go where the angels fear to tread. This topic is clearly beyond the ken of a layman and requires that any such conclusion be based on evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence.

6. Victor Stanley, Inc. v. Creative Pipe, Inc., 250 F.R.D. 251, 261 n.10 (D. Md. 2008). [Judge Facicia] made the entirely self-evident observation that challenges to the sufficiency of keyword search methodology invariably involve scientific, technical and scientific subjects, and give due weight to pronouncements from lawyers unsupported by an affidavit or other showing that the search methodology was effective for its intended purpose are of little value to a trial judge who must decide a discovery motion aimed at either compelling a more comprehensive search or preventing one.
requirements of Evidence Rule 702. This article suggests that the practical impact of treating electronic searching as an expert function is to permit attorneys to focus and strategize on the process of electronic searching rather than on the completeness of document production. In effect, electronic searching permits attorneys to quit focusing on finding documents and begin focusing on identifying electronic sources of information on which reside relevant documents that can be extracted by means of electronic searching protocols.

II. THE CHALLENGE OF MANUAL REVIEW: VOLUME

There appears to be no serious case law discussion of the minimum competency standards required to search paper documents or physical file cabinets for data relevant to a matter or privileged as attorney-client communication. It is beyond cavil that, in a world of paper documents, it has been standard procedure for attorneys to manually review data prior to production in litigation and determine whether the data was privileged, relevant, confidential, etc. As clients migrated from typewriters to word processors to computers, not only has the media on which data resides changed, but the volume of client data has exploded.

The shift of information storage to a digital realm has, for a variety of reasons, caused an explosion in the amount of information that resides in any enterprise profoundly affecting litigation. This massive amount of electronically stored information is distributed broadly among different storage devices, from large mainframe computers, to tiny machines capable of storing information equivalent to several warehouses of documents each, all of which are or can be integrated into other systems. These systems are complex, interdependent, and evolve spontaneously, like ecosystems. It is often impossible to find one person, or even one discrete group of people, who completely understand the working of this new form of “information ecosystem.”

   Just a few years ago all information was stored on physical records such as paper. . . . It was reasonable, and indeed relatively easy in all but the exceptional case, for the legal profession to gather and then manually review all the individual items collected as part of the discovery process prior to their production.

8. Id. ("Discovery has changed. In just a few years, the review process needed to identify and produce information has evolved from one largely involving the manual review of paper documents to one involving vastly greater volumes of electronically stored information.").

9. Id.
Although client data has undergone a radical transformation from discrete pieces of paper to an “information ecosystem,” attorneys have generally continued to manually review client electronic data for privilege, treating the electronic data in the same manner as they have reviewed paper documents for generations. The impact of treating electronic information as if it were the same as paper documents is most significant in the manual review of data for privilege.

Much of the manual review of client data occurs as part of the general discovery process. The United States Supreme Court has long held that discovery of data relevant to a matter and in the possession, custody, or control of a litigant was a necessary part of litigation in order to ensure open, efficient, and fair dealings within the federal court system. Under the Hickman view of litigation, “every party to a civil action is entitled to the disclosure of all relevant information in the possession of any person, unless the information is privileged.” The goal of liberalized discovery was to avoid surprise and to “make a trial less a game of blind man’s bluff and more a fair contest with the basic issues and facts disclosed to the fullest practicable extent.” Discovery of relevant information has become the way for litigants to obtain the fullest possible knowledge of the issues before trial, while permitting attorneys and clients to preserve privileged communications.

10. See Rest Practices, supra note 4. See also supra note 7 and accompanying text.
11. See Rest Practices, supra note 4. See also supra note 8 and accompanying text.
12. Hickman v. Taylor, 329 U.S. 495, 501 (1947). The various instruments of discovery now serve (1) as a device, along with the pre-trial hearing under Rule 16, to narrow and clarify the basic issues between the parties, and (2) as a device for ascertaining the facts, or information as to the existence or whereabouts of facts, relative to those issues. Thus civil trials in the federal courts no longer need to be carried on in the dark. The way is now clear, consistent, with recognized privileges, for the parties to obtain the fullest possible knowledge of the issues and facts before trial.
13. Id. at 597-98. (The deposition-discovery rules are to be accorded a broad and liberal treatment. No longer can the time-honored cry of “fishing expedition” serve to preclude a party from inquiring into the facts underlying his opponent’s case. Mutual knowledge of all the relevant facts gathered by both parties is essential to proper litigation. To that end, either party may compel the other to disgorge whatever facts he has in his possession. The deposition-discovery procedure simply advances the stage at which the disclosure can be compelled from the time of trial to the period preceding it, thus reducing the possibility of surprise. . . . And as Rule 26(b) provides, further limitations, come into existence when the inquiry touches upon the irrelevant or immaterial issues in the recognized domains of privilege.)
As the world transitioned from paper documents to electronically stored information, the Federal Rules of Civil Procedure, generally, were interpreted to accommodate that change as part of the discovery process.\textsuperscript{16} Rule 34 of the Federal Rules of Civil Procedure and its state-law counterparts were generally interpreted to include electronic information within the definition of “data compilation.”\textsuperscript{17} As a result, “data compilations” were deemed to be documents just like traditional paper documents and subject to discovery and production.\textsuperscript{18} Just like paper documents, “data compilations” needed to be reviewed and privileged client data identified and excluded from production to a party opponent in litigation.\textsuperscript{19}

It was not long before the unique features of electronic data began to interfere with the review and production of data. The volume of electronic information compared to paper documents, the redundancy of multiple electronic copies of the same information, the lack of a coherent filing system in which electronic information may be stored, and the unique cost issues associated with electronic information storage systems and media that have become obsolete were the primary reasons that electronically stored information was difficult to review for privilege and produce.\textsuperscript{20} Although these features of electronic data

\textsuperscript{16} Roland Bernieri, Avoiding an E-Discovery Odyssey, 56 N.Y.U. L. Rev. 491, 495 (2009).

\textsuperscript{17} The legal community has attempted to address the effect of technology on discovery issues. In August 2004, an advisory committee published a proposed set of amendments for the Federal Rules of Civil Procedure designed to guide courts and attorneys on issues associated with electronic discovery. The committee passed a revised set, and ultimately these were adopted by the U.S. Supreme Court without a substantive modification.

\textsuperscript{18} See Best Practices, supra note 4. See also supra note 8 and accompanying text.

\textsuperscript{19} Sedona Principles Second Edition, supra note 17, at page iv (“This equal treatment suggested that electronic information should be searched for, processed, and produced like paper.”).


[The Court is not persuaded by the plaintiffs' attempt to equate traditional paper-based discovery with the discovery of e-mail files. Several commentators have noted important differences between the two. . . . Chief among these differences is the sheer volume of electronic information. . . . Additionally, computers have the ability to capture several copies (or drafts) of the same email, thus multiplying the volume of documents. . . . Also, unlike most paper-based discovery, archived e-mails typically lack a coherent filing system. Moreover, stored archival systems commonly store information on magnetic tapes which have become obsolete. Thus, parties incur additional costs in translating the data from the tapes into usable form.]
dramatically increased the cost of privilege review, the consequences of failing to adequately review client data continued to threaten attorneys and clients with the draconian results of privilege waiver.21

As technical challenges to the production of electronically stored information were encountered, a body of research and law began to be created giving some guidance to attorneys regarding the choices and decisions necessary to produce electronically stored information in discovery.22 Privilege review, however, has only recently been addressed by case law.

Whether electronically stored information can be reviewed by attorneys for privilege or relevancy in a manner identical to the review of paper documents has not been the subject of much research, and still less case law.23 This may be because, until recently, attorneys were manually reviewing all electronically stored information prior to production.24 It is highly doubtful, however, whether manual review of documents for privilege can survive the increase in volume of data that has occurred as the result of the ubiquitous use of computers and electronic communication networks that form the new “information ecosystem.”25 Indeed the manual review for privilege of ever-

---

21. Victor Stanley, Inc. v. Creative Pipe, Inc., 250 F.R.D. 251, 267-8 (D. MD 2008) “[T]he court finds that the Defendants waived any privilege or work product protection for the 165 documents at issue by disclosing them to the Plaintiff”; see also Fed. R. CIV. P. 26(b)(5) Advisory Committee Note (“The Committee on the Rules of Practice and Procedure] has repeatedly been advised that the risk of privilege waiver and the work necessary to avoid it, add to the costs and delay of discovery. When the review is of electronically stored information, the risk of waiver, and the time and effort required to avoid it, can increase substantially because of the volume of electronically stored information and the difficulty in ensuring that all information to be produced has in fact been reviewed.”

22. Sedona Principles Second Edition, supra note 17, at page iv. (“For from supplementing The Sedona Principles, the new Federal Rules have highlighted the many areas of electronic discovery in which there is continued and growing need for guidance.”).


24. See Best Practices, supra note 4. See also supra notes 7-8 and accompanying text.


[With the digital revolution there has also been a paradigm shift in the review process which is feasible. The shift of information storage to a digital realm has, for a variety of reasons, caused an explosion in the amount of information that resides in any enterprise- profoundly affecting litigation. This massive amount of electronically stored information is distributed among different storage devices . . . These systems are complex, interdependent, and evolve spontaneously, like ecosystems. ]
increasingly larger volumes of electronic information has become the single most costly step in the production of electronically stored information.\textsuperscript{26}

III. THE COST OF MANUAL REVIEW
OF ELECTRONICALLY STORED INFORMATION

The cost of manual review is driven initially by the sheer volume of data that can now be stored on very small devices.\textsuperscript{27} For example, manually reviewing one gigabyte of electronic documents can be estimated to cost a client about $32,000 of attorney time and labor.\textsuperscript{28} This estimate is based upon a common assumption that one gigabyte of data constitutes 80,000 to 100,000 pages of data.\textsuperscript{29} A single attorney ought to be able to review 500 pages of data per hour with acceptable accuracy.\textsuperscript{30} One gigabyte of data, therefore, will require one attorney to spend 160 to 200 hours reviewing the data and identifying whether it is privileged.\textsuperscript{31} At an average billable rate of $200 per hour, one attorney can review one gigabyte of data at a cost of between $32,000 to $40,000.\textsuperscript{32}

Given the cost of manual review, it is not surprising that the continued use of this procedure is becoming (and in many cases, has already become) and unacceptable cost of litigation. In one recent case, for example, a litigant spent eighteen months and $11.4 million to hire contract attorneys to review electronic documents culled from 127

\textsuperscript{26}. Paul & Baron, supra note 3, at 4 ("Litigators can no longer depend on manual review alone. It is too time-consuming and expensive – with costs often exceeding the amount in dispute.").

\textsuperscript{27}. Best Practices, supra note 4, at 198 ("In many organizations, the average worker maintains several gigabytes of stored data. At the same time, the cost of storage has plummeted from $20,000 per gigabyte in 1990 to less than $1 per gigabyte today.").

\textsuperscript{28}. Paul & Baron, supra note 3, at 20.

\textsuperscript{29}. See id.

\textsuperscript{30}. See id.

\textsuperscript{31}. See id.

\textsuperscript{32}. See id.
document custodians for privilege prior to production. Cases involving terabytes of data (one terabyte = 1000 gigabytes) will require tens of millions of dollars to manually review. It has become obvious to anyone that is familiar with these changes and costs that the litigation system cannot continue to operate under these strictures. Cost has gotten so significant that manual review of large datasets for privilege, relevance, or work product has been characterized by at least one group as "indefensible.

In response to the cost of manual review, attorneys are being forced to leverage technology and "use computers and not just associates, contract lawyers, or outsourced offshore workers to search [client data]. Generally, attorneys use computers to search client data by running keyword searching software programs to identify documents responsive to requests for production of documents. The most common form of electronic search tool is a software program that accepts "keywords" or phrases and identifies instances of those words or phrases in the client data. The keywords and phrases can be either simple words, word combinations, or may contain Boolean and related operators. While the use of this type of keyword searching has long been used to search for relevant case law in computerized legal libraries, its use to identify privileged and work product documents within the client data set is relatively new. Some commentators have duly noted that keyword searching case law libraries is significantly different than

34. Id. at *2 (discussing the "huge" production of data in the case).
35. Paul & Baron, supra note 3, at 20 ("The numbers add up to more than a burden than any party should assume, no matter how rich in resources, without changes being made in the way cases are litigated and to techniques used in discovery.").
36. Best Practices, supra note 4, at 199 ("Although the continued use of manual search and review methods may be indefensible in discovery involving significant amounts of electronically stored information, merely adopting sophisticated automated search tools, alone, will not necessarily lead to successful results.").
37. Paul & Baron, supra note 3, at 36.
38. Id. at 37 ("The legal profession has adopted keyword searching in light of its longtime familiarity with its use in connection with the offerings of the major online legal retrieval services, which allow for searches to be made of structured databases containing case precedents and statutory authority.").
39. Best Practices, supra note 4, at 200
   By far the most commonly used search methodology today is the use of "keyword searches" of full text and metadata as a means of filtering data for producing responsive documents in civil discovery. . . . [T]he use of the term 'keyword search' refers to text-based searching using simple words or word combinations, with or without Boolean and related operators.

Id.
40. Id. at 21.
keyword searching client data, primarily because the language in case
law is much more structured and predictable than the language used in
communications and documents created by employees in a workplace
environment.\footnote{Rich & Baron, supra note 3, at 38 ("First, and most importantly, there are profound issues
of ambiguity and indeterminacy in human language, and thus it all texts in large, heterogeneous
databases subject to discovery. . . . Furthermore, people make up words on the fly, including new
words that function as language.")}

Relying upon the results of a keyword search tool or any other form
of electronic search protocol to identify documents that a litigant claims
are privileged or work product necessarily exposes that search tool or
electronic protocol to analysis when its results are challenged.\footnote{Victor Stanley, Inc. v. Creative Pipe,
Inc., 250 F.D.R. 251, 262 (D. Md. 2008).} This analysis has only recently been the subject of a few cases in which
courts have begun to define the nature of electronic searching and the
minimum competency necessary to defend search results.\footnote{Selection of the appropriate search and information retrieval technique requires careful
advance planning by persons qualified to design effective search methodology. The
implementation of the methodology selected should be tested for quality assurance, and
the party selecting the methodology must be prepared to explain the rationale for the
method chosen to the court, demonstrate that it is appropriate for the task, and show that
it was properly implemented.} A detailed
discussion of each of these cases reveals a common thread: Configuring
legally defensible electronic search strategies is an expert function,
significantly different than the expertise needed to review paper
documents. The failure to recognize the expert nature of electronic
searching may lead attorneys to construct search strategies that cannot be
defended when challenged.

IV. Nature of Electronic Searching:
Expert Language and Expert Function

The first case to define some characteristics of electronic search
tools and protocols was a criminal matter in which the defendant
challenged the testimony of a prosecution witness.\footnote{See id.; United States v. Ganier, 468 F.3d 920 (6th Cir.
2006); United States v. O’Keeffe, 537 F. Supp. 2d 14 (D.D.C 2008).} In the Ganier case,
the prosecution sought to elicit the testimony of Special Agent Wallace
Druick regarding the electronic searches that defendant Ganier had run
on the defendant’s computer and the deletion of certain data relevant to a
grand jury investigation.\footnote{Ganier,468 F.3d 920.} Basically, the prosecution sought to introduce

\begin{itemize}
\item \footnote{Id. at 924.}
\end{itemize}
Agent Druick’s testimony to link, in time and by subject, the search activity on Ganier’s computer with the grand jury deliberations. The defense objected to Druick’s testimony on the ground that the testimony was admissible only under Evidence Rule 702, as expert testimony, and the prosecution had neither properly identified Druick as an expert witness, nor properly tendered an expert report as required by Rule 706 of the Federal Rules of Criminal Procedure. The trial court agreed with the defense and dismissed the case. On appeal, the Sixth Circuit Appellate Court examined the issue whether Agent Druick’s testimony was of such a character that it required he be admitted as an expert under Evidence Rule 702.

The prosecution argued that Agent Druick was a fact witness, not an expert witness, because Agent Druick merely launched certain special software to run over the defendant’s computer and then observed the results:

The government argues that Druick’s proposed testimony is not based on scientific, technical, or other specialized knowledge, but is simply lay testimony available by “running commercially-available software, obtaining results, and reciting them.” The government contends that this testimony is of the same type as “facts . . . that could be observed by any person reasonably proficient in the use of commonly used computer software, such as Microsoft Word and Microsoft Outlook (such as the existence and location of multiple copies of documents that are identical or virtually identical to the allegedly ‘deleted’ documents),” which Ganier previously indicated he did not consider to be expert testimony.

The Sixth Circuit Appellate Court analyzed the issue whether the proposed testimony of Agent Druick was expert testimony by reviewing the type of knowledge that Agent Druick would necessarily apply to the

46. Id. at 924-25.
47. Id. at 921. 
48. Id. 
49. United States v. Ganier, 468 F.3d 920, 925 (6th Cir. 2006).
We must first determine whether the district court erred by concluding that Druick’s proposed testimony fell within the scope of Federal Rule of Criminal Procedure 16(a)(1)(D). Rule 16(a)(1)(D) requires, in part, that “[i]f the defendant’s request, the government must give to the defendant a written summary of any testimony that the government intends to use under Rules 702, 703, or 705 of the Federal Rules of Evidence during its case-in-chief at trial.”

Id. at 925-26.
output report generated by the “commercially-available software” used by Agent Drucek.\textsuperscript{51}

The reports generated by the forensic software display a heading, a string of words and symbols, a date and time, and a list of words. The government asserts that these reports reveal three different types of searches performed with particular search terms at particular times, but such an interpretation would require Drucek to apply knowledge and familiarity with computers and the particular forensic software well beyond that of the average layperson. This constitutes “scientific, technical, or other specialized knowledge” within the scope of Rule 702.\textsuperscript{52}

In addition to the special knowledge and familiarity with computers that the court determined was an integral part of Agent Drucek’s testimony, the Sixth Circuit Appellate Court also analyzed the language that Agent Drucek would necessarily use in his testimony to explain the actions taken to search and destroy information on defendant’s computer.\textsuperscript{53} It is useful to note that the court analogized the language used by Agent Drucek to describe computer-related facts with the specialized language used by police officers to explain drug arrests.\textsuperscript{54} The Appellate Court affirmed the designation made by the trial court that Agent Drucek’s testimony was properly suppressed pursuant to Evidence Rule 702, requiring that he be identified as an expert.\textsuperscript{55}

Applying the Gnanier analytic paradigm to an explanation of the use of electronic search tools by attorneys to identify privileged, relevant, or work product documents, requires that we identify the nature of the “knowledge” and “vernacular” that would need to be used to explain the use of such tools.\textsuperscript{56} The most common type of electronic search

\textsuperscript{51} Id. at 926.

\textsuperscript{52} Id.

\textsuperscript{53} Id.

\textsuperscript{54} United States v. Gnanier, 468 F.3d 920, 926 (6th Cir. 2006).

Because the categorization of computer-related testimony is a relatively new question, comparisons with other areas of expert testimony are instructive. Software programs such as Microsoft Word and Outlook may be as commonly used as home medical thermometers, but the forensic tests Drucek ran are more akin to specialized medical tests run by physicians. The average layperson today may be able to interpret the outputs of popular software programs as easily as he or she interprets everyday vernacular, but the interpretation Drucek needed to apply to make sense of the software reports is more similar to the specialized knowledge police officers use to interpret slang and code words used by drug dealers.

\textsuperscript{55} Id. at 927.

\textsuperscript{56} Id. at 926.
methodology used today is “keyword” searches.\textsuperscript{57} Keywords can be used by themselves or combined with “operators” to construct search engines which are then applied to data sets.\textsuperscript{58} Many courts embrace keyword searching as an electronic search protocol and many attorneys attempt to agree upon the keywords that will be used for purposes of production in discovery.\textsuperscript{59}

Defending the use of specific keywords requires an analysis of the metrics by which keyword searching is measured. Keyword search measurements include specialized concepts, such as basic information retrieval metrics of “recall” and “precision.” Recall is a measure of completeness—namely, how well an electronic search protocol has identified all the potentially responsive documents from the client data set.\textsuperscript{60} It is derived by dividing the number of responsive documents retrieved by the total number of responsive documents.\textsuperscript{61} “Precision” is a measure of efficiency—namely how well an electronic search protocol has identified responsive documents as a percentage of the total number of documents retrieved, including all false positives.\textsuperscript{62} It is derived by dividing the number of responsive documents retrieved by the total number of documents retrieved.\textsuperscript{63}

Explaining the manner in which search terms were chosen may require use of specialized language. “Ambiguity” and “variation” are common characteristics of language that need to be incorporated into electronic search protocols to render them effective for particular client

\textsuperscript{57} See \textit{Best Practices}, supra note 4. See also supra note 88 and accompanying text.

\textsuperscript{58} \textit{Best Practices}, supra note 4, at 207 (“First, there are keyword based methods, ranging from the simple use of keywords alone, to the use of strings of keywords with what are known as ‘Boolean operators’ (including AND, OR, ‘AND NOT’ or ‘BUT NOT’).”)

\textsuperscript{59} Id. at 209 (“Courts have not only accepted, but in some cases ordered, the use of keyword searching to define discovery parameters and resolve discovery disputes; see also Balboa Threadworks v. Stuckey, 2006 WL 743668, at *5 (D. Kan 2006).”

As to the formulation of a search protocol, whether one using keyword searches and/or other search procedures, the parties are directed to meet and confer in an attempt to agree on an appropriate protocol, and should lean heavily on their respective computer experts in designing such a protocol. Numerous types and varieties of search protocols have been discussed and adopted by courts, and these may guide the parties in designing a search protocol to be used in this case.

\textsuperscript{60} Paul & Baron, supra note 3, at 41.

\textsuperscript{61} \textit{Best Practices}, supra note 4, at 205.

\textsuperscript{62} Paul & Baron, supra note 3, at 41.

\textsuperscript{63} \textit{Best Practices}, supra note 4, at 205.
data sets.64 “Ambiguity” refers to the tendency of words and expressions to have different meanings when in different contexts.65 Each context is a “variation.”66 One of the compelling characteristics of language is the ability to use many different words and expressions to convey content.67 Configuring electronic search protocols to effectively identify privileged, relevant, or work product documents in the client data set requires that the search protocols reflect the ambiguity and variation of language used in the client data set.68

Specialized language may be necessary to explain the manner in which ambiguity and variation were recognized in the client data set and incorporated into the search. “Taxonomies” and “ontologies” are essentially synonyms of words and relevant classes of words that are developed and included in electronic search strategies to refine the search by maximizing recall and precision.69 Specialized statistical concepts may be necessary to explain the basis upon which the size of statistically significant random samples of client data sets were calculated.70 Increasing reliance upon sampling of electronically stored information was expressly incorporated into amendments to Rule 34:

Rule 34(a)(1) is also amended to make clear that parties may request an opportunity to test or sample materials sought under the rule in addition to inspecting and copying them. That opportunity may be important for both electronically stored information and hard-copy materials. The current rule is not clear that such testing or sampling is authorized; the amendment expressly permits it. As with any other form of discovery, issues of burden and intrusiveness raised by requests to test or sample can be addressed under Rule 26(b)(2) and 26(c).71

64. Id. at 206 (“The richness of human language causes a severe challenge in identifying informational records.”).
65. Id.
66. Id.
67. Id. at 207 (“But as the Blair and Maron study demonstrates, human language is highly ambiguous and full of variation.”).
68. Id. (“In the years since Blair and Maron, the IR community has been engaged in research and development methods, tools, and techniques that compensate for endemic ambiguity and variation in human language, and thus maximize the recall and precision of searches.”).
69. Paul & Baron, supra note 3, at 43.
70. Best Practices, supra note 4, at 207 (“There are a variety of statistical techniques, which analyze word counts.”).
71. Fed. R. Civ. P. 34, Advisory Committee’s Note.
Explaining the manner in which a client data set was randomly sampled or electronic search results were applied to the client data set almost surely require specialized language of the type that the Ganiher court labeled as “expert testimony.”

Finally, explaining the type of electronic search that was used in a particular case and comparing that type of search with other searching methods will require specialized language.

Even before the emergence of the Web, information retrieval science has constituted a vast and growing field. However, broadly speaking, information retrieval methods fall into three broad classes: set-theoretic (Boolean strings, supplemented by fuzzy search capabilities), algebraic (promised on the mathematical idea that the meaning of a document can be derived from the constituent terms in a document, and thus weighting retrieval by the proximity of a document’s terms in the form of two or higher dimensional maps, as in vector space modeling), and probabilistic (using language models and Bayesian belief networks, the latter of which involves make educated inferences about the relevance of future documents based on prior experience in reviewing documents in a given collection).

The Ganiher court’s “specialized language” analysis, when applied to the testimony, language, and vernacular required to defend the use of electronic search tools, including keyword searching, by attorneys for purposes of identifying privileged, work product, or relevant data, appears to characterize such testimony as Rule 702 expert testimony. Court decisions subsequent to Ganiher have analyzed the specific types of expert testimony that will be required in order to successfully defend particular keyword searching protocols.

V. ANALYSIS OF NATURE OF KEYWORD SEARCHING: EXPERT FUNCTION COMBINING LINGUISTICS, STATISTICS, AND COMPUTER TECHNOLOGY

In US v O’Keefe, Judge Facciola analyzed a defendant’s challenge to the electronic search protocols used by the Department of State to locate all information in its possession custody or control related to O’Keefe’s indictment charging he expedited visa requests in exchange for gifts. In his analysis, Judge Facciola set forth the scope of the

72. See supra note 52 and accompanying text.
73. Paul & Baron, supra note 3, at 42
74. See supra notes 51-52 and accompanying text.
technical character and specialized features required of electronic search protocols, including keyword searches.76

As noted above, defendants protest the search terms the government uses. Whether search terms or “keywords” will yield the information sought is a complicated question involving the interplay, at least, of the sciences of computer technology, statistics and linguistics.77 Indeed, a special project team of the Working Group on Electronic Discovery of the Sedona Conference is studying that subject and their work indicates how difficult this question is.78 Given this complexity, for lawyers and judges to dare opine that a certain search term or terms would be more likely to produce information than the terms that were used is truly to go where angels fear to tread. This topic is clearly beyond the ken of a layman and requires that any such conclusion be based on evidence that, for example, meets the criteria of Rule 702 of the Federal Rules of Evidence. Accordingly, if defendants are going to contend that the search terms used by the government were insufficient, they will have to specifically so contend in a motion to compel and their contention must be based on evidence that meets the requirements of Rule 702 of the Federal Rules of Evidence.79

Judge Facciola’s analysis in O’Keefe is focused on the nature of keyword searching and concludes that keyword searching is an expert function because it relies upon the application of specialized knowledge and concepts.80 Judge Facciola incorporated into the O’Keefe decision the research and knowledge of the combined areas of computer technology, linguistics, and statistics, and determined that challenges to the use of electronic search protocols must be based on expert testimony.81

After the Ganiar and O’Keefe cases, the issue remained whether creating keyword search protocols was, itself, an expert function demanding special competencies on the part of the attorneys or law

76. Id.
77. See Paul & Baron, supra note 3.
78. See Best Practices, supra note 4.
79. Id. at 23–24.
80. Id. at 24.
81. Id. ("Whether search terms or keywords will yield the information sought is a complicated question involving the interplay, at least, of the sciences of computer technology, statistics, and linguistics.").

In the Victor Stanley matter, defendants Creative Pipe Inc. and Mark and Stephanie Pappas produced data to plaintiff Victor Stanley, Inc. in response to plaintiff’s request for production of documents. Prior to producing the data, defendants’ counsel conducted an electronic search of the client data for privileged documents. Unfortunately, the electronic search did not identify all privileged material, and 165 privileged documents were disclosed to plaintiff. Defendants requested the return of the 165 documents, but plaintiff insisted that the privilege had been waived by disclosure.

The issue before Magistrate Judge Grimm was whether defendants had waived the attorney-client privilege by reason of their inadvertent production of the privileged documents. Basically, if the defendants had acted in a reasonable manner to prevent the inadvertent disclosure, there would be no waiver of privilege. Whether defendants acted reasonably, in turn, required the court to analyze the manner in which defendants created their keyword search strategy.

To create the keyword search, counsel for Creative Pipe met and conferred with their client and with co-defendant Mark Pappas. Together they devised a keyword search strategy to locate privileged document consisting of seventy keywords that they believed ought to identify all privileged data. Counsel ran those keywords over all client documents and any document that contained one or more of the keywords was withheld from production on the ground of privilege. It is significant to note that the privilege search undertaken in the Victor Stanley case appears to be identical to the manner in which any law firm might use electronic keyword searches to identify privileged documents contained within a client data set of documents relevant to litigation.

Although counsel was in control of the client’s data set, counsel took no action other than to “guess” the keywords that ought to be used.

83. Id. at 255.
84. Id.
85. Id.
86. Id. at 257.
88. Id. at 254.
89. Id.
90. Id.
to identify privileged documents. Counsel took no actions to determine
the actual language used by the client’s employees to create the data in
the client data set. No attempt was made to identify any taxonomies or
ontologies; no sampling was done to identify the ambiguities or
variations used by the creators of the client data. It is also interesting to
note that this technique (meeting to confer and “guess” keywords) is the
manner in which most litigation counsel agrees with a party opponent
regarding search terms to be used to identify potentially relevant data as
part of discovery. Indeed, in *Victor Stanley*, the litigants had previously
met and agreed upon search terms designed to locate potentially relevant
data that would then need to be reviewed for privilege. The fact that
the seventy keywords used by counsel in *Victor Stanley* did not
completely identify all privileged documents indicates that there was
sufficient ambiguity in the client data set to defeat the keyword search.
The issue for the court, however, was whether the keyword strategy used
by counsel and client was defensible, notwithstanding its failure to
capture all privileged documents.

In order to decide whether the seventy keywords comprised a
defensible search of the client data set for privileged documents, the
court looked to the defendants to provide the court with information
about the people and the process that was used to create (or guess) the
seventy keywords. Magistrate Judge Grimm demanded that defendants
produce evidence in the nature or testimony that demonstrated the
protocols chosen by defendants were appropriate for the task. It is
significant—indeed essential—for attorneys to note that the process of
keyword searching for privileged information was not shielded by work
product or attorney-client privilege. Nor was the process of constructing

91. *Id.*
92. *Id.* at 257.
   While it is known that M. Paggas (a party) and Meh and Schid (attorneys) selected the
   keywords, nothing is known from the affidavits provided to the court regarding their
   qualifications for designing a search and information retrieval strategy that could be
   expected to produce an effective and reliable privilege review.

94. *Id.* at 262.
   Use of search and information retrieval methodology, for the purpose of identifying and
   withholding privileged or work product protected information from production, requires
   the utmost care in selecting methodology that is appropriate for the task, because the
   consequence of failing to do so, as in this case, may be the disclosure of
   privileged/protected information to an adverse party, resulting in a determination by the
   court that the privilege/protection has been waived.
keywords assumed to be a function subsumed by the ordinary practice of law—such as one might characterize the manual review of client data by attorneys in the past. Rather, the court characterized the use of electronic search tools as a process that must be defended and explained as any other process or methodology:

Selection of the appropriate search and information retrieval technique requires careful advance planning by persons qualified to design effective search methodology. The implementation of the methodology selected should be tested for quality assurance, and the party selecting the methodology must be prepared to explain the rationale for the method chosen to the court, demonstrate that it is appropriate for the task, and show that it was properly implemented.95

Applying this process-oriented analysis to the facts in the Victor Stanley case, Magistrate Judge Grimm held that counsel for defendants had:

[F]ailed to provide the court with information regarding: the keywords used; the rationale for their selection; the qualifications of M. Pappas and his attorneys to design an effective and reliable search and information retrieval method; whether the search was a simple keyword search, or a more sophisticated one, such as one employing Boolean proximity operators; or whether they analyzed the results of the search to assess its reliability, appropriateness for the task, and the quality of its implementation.96

The Victor Stanley court grounded its process-oriented analysis upon the science of information retrieval. The court demanded defendants provide testimony in support of the search protocols used—not mere legal argument.97 Citing Judge Facciola in United States v. O’Keefe, Judge Grimm noted, in the Creative Pipe case, that keyword searches may, indeed, be the proper method for searching in a matter, but “there are well-known limitations and risks associated with them, and proper selection and implementation obviously involves technical, if not scientific knowledge.”98

Judge Facciola made the entirely self-evident observation that challenges to the sufficiency of keyword search methodology

95. Id.
96. Id. at 259-60.
97. Id. at 260 (“While keyword searches have long been recognized as appropriate and helpful for ESI search and retrieval, there are well-known limitations and risks associated with them, and proper selection and implementation obviously involves technical, if not scientific knowledge.”).
98. Id. at 260.
unavoidably involve scientific, technical and scientific [sic] subjects, and *ipsa dixit* pronouncements from lawyers unsupported by an affidavit or other showing that the search methodology was effective for its intended purpose are of little value to a trial judge who must decide a discovery motion aimed at either compelling a more comprehensive search or preventing one.99

VI. SEARCHING ELECTRONICALLY STORED INFORMATION IS AN EXPERT PROCESS

The *Victor Stanley* court is not alone in characterizing electronic searching as a process that must be defended as an expert process. The Sedona Conference’s “Practice Point 7” related to Search and Information Retrieval Methods also describes electronic searching as a process or methodology based on the science of information retrieval.

Counsel should be prepared to explain what keywords, search protocols, and alternative search methods were used to generate a set of documents, including ones made subject to subsequent manual searches for responsiveness and privilege. This explanation may best come from a technical “IT” expert, a statistician, or an expert in search and retrieval technology. Counsel must be prepared to answer questions, and indeed, to prove the reasonableness and good faith of their methods.100

Characterizing electronic searching as an expert process subjects the search to analysis and challenges, requires the search process be defended, and triggers significant implications for attorney issues related to malpractice.101

100. *Best Practices, supra* note 4, at 212.

Electronic discovery requires cooperation between opposing counsel and transparency in all aspects of preservation and production of ESI. Moreover, where counsel are using keyword searches for retrieval of ESI, they at a minimum must carefully craft the appropriate keywords, with input from the ESI’s custodians as to the words and abbreviations they use; and the proposed methodology must be quality control tested to assure accuracy in retrieval and elimination of “false positives.” It is time that the Bar, even those lawyers who did not come of age in the computer era, understand this.
VII. CHALLENGING ELECTRONIC SEARCH PROCESSES AS REASONABLE

In any particular case, challenges to electronic search protocols will probably be raised by a party opponent to measure the degree to which a responding party reasonably searched client data, especially if there has been little or no discussion amongst or between counsel and little or no transparency of the search tools and protocols. Traditionally, the producing party to a discovery request has enjoyed a presumption that it knows best the location and method by which to identify and produce documents. However, this presumption may not always apply to the use of electronic search tools. While courts and litigants may be willing to accept representations and statements on the record that counsel has performed a manual review of client data in a competent manner, recent case law discussed herein suggests no similar deference is afforded electronic searching. Perhaps this is because courts and litigants intuitively understand that searching electronically stored information is an expert process much more difficult to properly design and execute than the manual review of paper documents. Additionally, the efficacy of manual review was never directly challenged but was hidden behind professional representations and assumptions of competency. When using electronic search tools to perform electronic searching functions, however, the efficacy of the search can be addressed directly and measured.

102. See Smith v. Life Investors Ins. Co. of Am., 2009 WL 2045197, at *7 (W.D. Pa. 2009). Defendant must do more than summarize list the number of pages it has produced and the time and effort it has invested. Rather, Defendant has a burden to demonstrate that its search for documents was reasonable. A thorough explanation of the search terms and procedures used would be a large step in that direction.

103. See Fed. R. Civ. P. 26. The language used 26(a)(1) and 26(a)(2) states: “[A] party must disclose.” Id. There is a presumption based on this language that a disclosing party is aware of what information they need to disclose and therefore is best suited to determine the location and method of producing the discovery items. However, 26(a) permits a court to intervene in the discovery process if necessary to facilitate disclosure of necessary information. Id.

104. Best Practices, supra note 4, at 204.


106. Best Practices, supra note 4, at 205.

One can often adjust a system to retrieve more documents, thereby increasing recall, but at the expense of retrieving more irrelevant documents, and thus decreasing precision.
Challenges to the effectiveness of electronic search methodologies can reasonably be expected to focus upon the expert nature of electronic searching and upon the reasonableness of the search and configuration decisions made while using the tools and protocols.\textsuperscript{107} Indeed, a challenging party may argue that the process used by the responding party is essentially an expert technology which has not been validated by subjecting it to peer review, and unbiased empirical testing or analysis.\textsuperscript{108}

Automated software solutions that enter the marketplace may also be challenged as a process—and some fear that these challenges could be difficult to overcome.

The probability of such a challenge is greater if the technology is patented or proprietary to a developer or vendor (i.e., in a so-called “Black Box”). In such circumstances, e-discovery and litigation support vendors that use these technologies may be several degrees of separation from the original developers. A requesting party may demand the responding party to “prove up” the use of such search technology. This could set the stage for a difficult and expensive battle of experts.\textsuperscript{109}

Perhaps future electronic searching software will need to follow the path of computer forensic software.\textsuperscript{110} Forensic software is used by experts to identify, preserve, and extract relevant electronic data (content and artifacts).\textsuperscript{111} Expert opinions based upon the results of using the

---

\textsuperscript{107} See Best Practices, supra note 4. See also supra note 109 and accompanying text.

\textsuperscript{108} Id. supra note 4, at 204.

\textsuperscript{109} Id.

\textsuperscript{110} See, e.g., United States v. Tank, 200 F.3d 627 (9th Cir. 2000) (validating the use of EnCase software to create image of Defendant’s computer and the authenticity of computer evidence in the general context of Fed. R. Evid. 901 (a)); Smith & Slifer Smith & Frantum/Vail Assocs. Real Estate, LLC, 2009 WL 482603 (D. Colo. 2009) (upholding good faith in the destruction of documents on the part of the respondent in producing computer records related to discovery).


Alcock’s proposed testimony also includes identifying “files and fragments of files previously deleted from the laptop” which he indicates involved expert reasoning since his work was “a time-consuming process due to the number of computer drives and files involved, and the complexity of retrieving files and artifacts damaged due to the attempts of erasing the drive to conceal or deprive the use of data once present on the laptop.” .

. . . Accordingly, Alcock’s proposed testimony constitutes expert testimony.

Id.
software are usually validated by independent means, thereby authenticating the results of the software.\textsuperscript{112}

In a similar manner, future electronic searching software tools may need to be validated by comparing the ability of the software to locate relevant documents in a universe of test data in which the identity of all relevant documents is known. The challenges of developing such electronic searching software, however, are formidable. They include the tremendous flexibility of the language and the creativity of humans which continues to frustrate attorneys who attempt to “rationally” guess effective keywords. For example, attorneys appear to be only 20 percent effective “at thinking up all of the different ways that document authors could refer to words, ideas, or issues in their case.”\textsuperscript{113}

The limitations on search and retrieval methodology exposed in the Blair and Maron study was not the ability of the computer to find documents that met the attorneys’ search criteria, but rather the inability of the attorneys and paralegals to anticipate all the possible ways that people could refer to the issues in the case. The richness of human language causes a severe challenge in identifying informational records.\textsuperscript{114}

If search and information retrieval methods are measured against the accuracy of attorneys “guessing” the language used by key players, then it may be relatively easy to demonstrate sufficient accuracy and precision to satisfy a Daubert/Frye challenge.

The Daubert challenge raised by [Judge] Facciola, then, may be met not by judging the scientific validity of a search engine in an absolute way, but by judging how valid it is to suit the purposes of e-discovery production, an undertaking which involves many factors, such as the costs in time, money and energy to the producing party and their marginal benefit to the requesting party and the litigation, that have no bearing on the scientific validity of the search engines.\textsuperscript{115}

\\textsuperscript{112} Leonard Deutschman, \textit{When Re-Discovery Is Put to the Test}, L. TECH. NEWS 1, May 14, 2008 (discussing issues of authentication related to the proprietary nature of search engines).

\textsuperscript{113} \textit{Best Practices}, supra note 4, at 206 (citing David Blair and M.E. Maron BART study, at 1985).

\textsuperscript{114} \textit{Id}.

\textsuperscript{115} Leonard Deutschman, supra note 112.
VIII. DEFENDING ELECTRONIC SEARCH PROCESSES AS REASONABLE AND FEDERAL EVIDENCE RULE 502(B)

Challenges to electronic search methodologies will require litigants to defend their electronic search processes and will force counsel to consider evidentiary issues at the beginning of the process, when electronic searching protocols are being created or negotiated. Based upon the language analysis in Ganier, and the expert process analysis in O'Keefe and Victor Stanley, prudent attorneys will treat electronic searching as an expert function comprised of skills in the area of computer technology, linguistics, and statistics. Prudent attorneys will base their electronic search strategies upon the advice of an expert or other authoritative source that is willing and able to defend those search strategies when challenged.

It cannot credibly be denied that resolving contested issues of whether a particular search and information retrieval method was appropriate—in the context of a motion to compel or motion for protective order—includes scientific, technical or specialized information. If so, then the trial judge must decide a method’s appropriateness with the benefit of information from some reliable source—whether an affidavit from a qualified expert, a learned treatise, or, if appropriate, from information judicially noticed.

Requiring litigants to defend their search protocols with expert testimony is similar to requiring expert testimony to explain and defend random sampling protocols. The requirement is a direct and necessary result of the court’s recognition of the technical aspects of electronically searching data. While some attorneys may view the requirement as a burden, Magistrate Judge Grimm suggests that this requirement ought to benefit the discovery process by reducing costs through cooperation between or amongst litigants.

117. See United States v. Ganier II, 468 F.3d 920 (6th Cir. 2006).
119. See Victor Stanley, Inc., 250 F.R.D. at 261 n.10. See also supra note 6 and accompanying text.
120. In re Viacom Prod. Liab. Litig., 2006 WL 1726675, at *2 n.5 (5th Cir. 2006) (“By random sampling, we mean adhering to a statistically sound protocol for sampling documents... The parties must provide expert assistance to the district court in constructing any protocol.”).
121. Victor Stanley, Inc., 250 F.R.D. at 261 n.10. See also supra note 6 and accompanying text.
If attorneys are to defend the process of electronically searching client data sets, they will need to better understand the features of that process. Compiling an exhaustive analysis of the many features of electronic searching is beyond the scope of this article, but a few fundamental features ought to be noted by all attorneys. For example, a key feature of electronic searching is its iterative nature. Rarely, if ever, will an initial keyword search yield satisfactory results. Keyword searches are notoriously over or under-inclusive. Part of the problem is the ambiguity of language; another is the failure of attorneys to recognize that the client data set, itself, can be analyzed for information to accurately create keyword searches.

For example, client data can be indexed. Indexing the entire client data set identifies every word in every document, accurately states the number of times the word appears, and keeps track of the documents in which the word resides and the key player that created the document. Rather than guess the keywords that client personnel may have used to create relevant or privileged data, attorneys could use indexing tools to know the universe of words actually used in the client data set and the frequency of their use. While this method is used commonly to locate “code” words or phrases used by cliques or clans in networks, it ought to be included in every electronic search strategy to help lessen the “guesswork” from keyword searches. This technique might also be considered by courts that mandate litigants agree to search terms as part of discovery conferences. Rather than forcing litigants to “guess” at the language used by their respective clients to designate relevant, work product, or privileged matters, courts perhaps could agree upon an iterative, index-enhanced, protocol that would substantially improve keyword searching.

Successfully defending electronic search methodologies will be especially important in light of changes to the federal rules of evidence

---

122. Paul & Baron, supra note 3, at 50.
123. Rest Practices, supra note 4, at 201.
[Although basic keyword searching techniques have been widely accepted both by courts and parties as sufficient to define the scope of their obligation to perform a search for responsive documents, the experience of many litigators is that simple keyword searching alone is inadequate in at least some discovery contexts. This is because simple keyword searches wind up being both over- and under-inclusive in light of the inherent malleability and ambiguity of spoken and written English as well as all other languages.

125. Paul & Baron, supra note 3, at 50.
that prohibit the use of inadvertently produced client data only if, *inter alia*, reasonable precautions were taken to avoid the disclosure. New Evidence Rule 502(b) was designed to respond to:

widespread complaints that litigation costs for review and protection of material that is privileged or work product have become prohibitive due to the concern that any disclosure of protected information in the course of discovery (however innocent or minimal) will operate as a subject matter waiver of all protected information. 126

As amended, Rule 502(b) provides that the inadvertent disclosure of privileged information in a federal proceeding, or to a federal officer or agency, does not waive the attorney-client privilege if:

1. The disclosure is inadvertent;
2. The holder of the privilege or protection took reasonable steps to prevent disclosure, and
3. The holder promptly took reasonable steps to rectify the error, including (if applicable) following Fed. R. Civ. P. 26(b)(3)(D). 127

Whether the holder of the privilege took "reasonable" steps to prevent disclosure will be the focus of analysis on a case by case basis. 128 So long as the steps taken can be proven to be reasonable, Rule 502(b)(2) ought to be satisfied. As noted by the Advisory Committee to Rule 502, the rule does not explicitly codify the reasonable test, because the rule is really a set of non-determinative guidelines that vary from case to case. 129

Rule 502(b) clearly invites attorneys to anticipate the technical, linguistic, and statistical challenges related to the use of electronic search tools, and create an electronic search process that can be defended in any particular case. 130 Evidence Rule 502(b) is an attempt to provide attorneys some relief from the overwhelming task of manually reviewing all client documents for privilege by expressly protecting client privilege while using reasonable electronic search protocols.

Evidence Rule 502(b) appears to incorporate the “expert process analysis” set out in *O'Keefe* and *Victor Stanley* with a particular emphasis upon computer technology to derive electronic search

---

126. FED. R. EVID. 502(b), Advisory Committee’s Note.
127. FED. R. EVID. 502(b).
128. *Id*.
129. *Id*.
130. Other electronic discovery processes may also fail to be reasonable. See Ameren Illinois Co. v. PerkinElmer, Inc., 2007 WL 329290 (D. Ill. 2007) (corrupt files included in production set that were readable by receiving party not reasonably protected from disclosure).
solutions that will be reasonable. It states: “A party that uses advanced analytical software applications and linguistic tools in screening for privilege and work product may be found to have taken ‘reasonable steps’ to prevent inadvertent disclosure.” Greater use of sampling and the implementation of an efficient system of records management may also be relevant to the issue whether reasonable precautions have been taken to avoid disclosure of privileged data in any particular case.

IX. CONCLUSION

As the volume of client data increases in litigation, economic pressure to reduce or eliminate manual review of client data for privilege, work product, and relevance will increase. Attorneys will be forced to use electronic searching tools and protocols to identify privileged, work product, or relevant data. These electronic tools, however, are fundamentally different from manual review. Electronic search and information retrieval tools represent an expert process that can be properly used and defended only if attorneys recognize that these tools must be used and configured in accordance with properly designed search protocols, results measured in accordance with accepted metrics such as recall and precision, and implemented in a technically valid manner. Challenges to the use of electronic search and information retrieval protocols will focus upon their technical features and will force attorneys to recognize that electronic searching is an expert process.

By focusing upon the expert process of electronic searching, and by judging the “reasonableness” of that process, courts are properly moving away from focusing discovery on measurements of the completeness of production. This shift in focus represents a significant “relief” from the economics of manual review. By creating an electronic search and information retrieval process that is defensible for the particular case in which it is to be used, attorneys will be able to incorporate technology into discovery and “dial in” the amount of precision and recall necessary. The end result will be an a reasonable process, of sufficient scope, precision, and recall to satisfy discovery without undue burden and expense.

131. Fed. R. Evid. 502, Advisory Committee’s Note.
132. Paul & Barron, supra note 3, at 47.
Social Media and Ethics

Jonathan E. Coughlan
Supreme Court of Ohio Disciplinary Counsel
Columbus, Ohio

Professor G. Kristen Barnes
The University of Akron School of Law
Akron, Ohio

I. Introduction to the Impact and Reach of Social Media
   A. Scope of the issue
   B. Overview of ethical issues

II. Social Media in the Courtroom
   A. Judges and social media
      1. Types of use
      2. Potential consequences
         a) Fairness and bias
         b) Undermining judicial authority and position
      3. Acceptable parameters of use
      4. Legal restrictions on use
      5. Making decisions about the use of social media by others in the courtroom
   B. Jurors and social media
      1. Types of use
         a) Researching case
b) Communicating feelings about serving on the jury or thoughts about the case before being selected to serve as a juror

c) Communicating with outsiders during the course of a trial

2. Potential consequences
   a) Fairness and bias
   b) Economic costs (e.g. mistrials, waste of judicial resources, etc.)

3. Acceptable parameters of use

4. Legal restrictions on social media

5. Balancing the right to a fair trial with the First Amendment right to free speech

6. Practical questions – what’s a lawyer to do?

7. Is there a new standard for lawyers?

C. Lawyers and social media

1. Types of use
   a. Soliciting business
   b. Using the social media profiles of others to select jurors or design trial strategy
   c. Foursquare: Locating witnesses and evidence and collecting evidence
   d. Ex parte communications (“friending judges”)
   e. Staying connected

2. Ethical concerns
   a. Attorney-client privilege
   b. Fairness

3. Acceptable parameters of use
   a. Advisory opinions authorizing certain types of social media use

2.2 • Ethical Issues and Practicing Law
b. Current trends

4. Legal restrictions on use

III. Social Media: Employers and Employees

A. Government entities
   1. Job applicants
   2. Employees

B. Private sector

C. Should the legal restrictions vary depending upon the type of employment?

IV. Social Media and Schools

A. Facebook and Twitter Accounts
   1. Students
   2. Student Athletes

B. Censorship and monitoring the use of social media
   1. Illusion of privacy: passwords and privacy protected accounts
   2. Social media monitoring companies

C. Issues associated with censorship and monitoring
   1. Ethical concerns
   2. Liability concerns

V. The Future - The Legal System and Managing Changing Social Media Practices
Substance Abuse from the Respondent Lawyer’s Perspective in a Nutshell

William C. Mann
Mitchell Pencheff Fraley Catalano & Boda Co. LPA
Columbus, Ohio
SUBSTANCE ABUSE FROM THE RESPONDENT LAWYER'S PERSPECTIVE IN A NUTSHELL

William Mann

1. Introduction:

This article will discuss the general aspects of a legal ethics case, which has a substance abuse component to it, from the respondent's (defendant's) perspective. The reader is first warned that no two cases are the same, but many are alike. The reader is reminded that this is a nutshell presentation. It is hopefully reliable but necessarily succinct and general in nature.

In a nutshell, there are basically four parts to each case. One, the lawyer and the client usually go through a Kabuki dance where the lawyer seeks to discover if the client has a substance abuse problem, and the client seeks to avoid such a discovery. Two, the lawyer encourages the client to get quality professional care and treatment for his or her substance abuse problem, and the client gets that help. Third, the lawyer seeks to obtain solid documentation that the client's substance abuse problem caused or significantly contributed to the client's alleged or actual misconduct. Four, the client gets better and can return to work as a competent, ethical and professional lawyer.

This presentation is about a lofty subject. But I am not a lofty person. As such, some may find that my presentation lacks a certain aloofness. I hope it does. Others may believe I am being disdainful of, and condescending to, some of my legal ethics—substance abuse clients. I am not. I love my legal ethics clients, and try to be an effective advocate for all of them. But, I also try to follow the prime directive of Gerry Spence, the greatest trial lawyer ever. That is, I always try to tell the truth.
2. **The Client:**

   The first thing a respondent's legal ethics case needs is a respondent or potential respondent. Respondent's come in all sorts of different physical, mental and emotional sizes, shapes and models. When a client with a substance abuse problem meets with me, he or she is often in denial. Denial is the invention of a false reality, and the belief in that false reality. Denial is very common among substance abusers.

   For instance, most (almost all) substance abusers will say things such as:

   1. I don't have a substance abuse problem.
   2. Sure I drink alcohol, but I can stop anytime I feel like stopping.
   3. Okay, so I snort a little cocaine once in a while, but only after I've had a really hard day. And nobody gets hurt. And did I mention that I don't have a problem.

   I'm getting a little ahead of myself. Let me go back to my initial meeting with the client.

3. **The Initial Lawyer – Client Meeting:**

   I usually begin with, "Thank you for coming in to see me. How can I help?"

   The client who is "guilty" and/or who has a substance abuse problem often responds with, "Blah-blah-blah-blah-blah." That is often followed by "Blah-blah-blah-blah-blah-blah." By the time the client is done speaking, my clothes have gone out of style, assuming they were in style in the first place, which is probably a risky assumption. And all the client's "blah-blah-blah," had a probative value of only a little more than absolute zero. In other words, the client who gives me a long drawn out story, but hardly tells me anything of significance is probably hiding something.
My job is to figure out what is being concealed from me (it could be anything) and why (most likely because it is embarrassing and/or establishes “guilt.”) Anyway, I often respond by asking if he or she has been contacted by a disciplinary prosecutor (the Office of Disciplinary Counsel or a certified grievance committee). If the answer is yes, and in substance abuse cases it usually is, I ask to see the Letter of Inquiry that the client received from the disciplinary prosecutor.

I then review the Letter of Inquiry and any other relevant materials that client provides. In a substance abuse case – and remember, I said this presentation is necessarily succinct and general in nature – our conversation then goes something like this:

Bill

This letter claims you were late for 12 court hearings within a 60 day period. Is that true?

Client


Now, we are getting somewhere, because anytime we see a reduction of productivity at home, work or play we suspect the client might have a substance abuse problem.

After again reminding the client of the confidentiality of our communication, we focus in on the following types of questions, which of course, are asked diplomatically and respectfully:

1. Have you ever consumed ___________________________?
2. When? Why?

3. When is the last time you consumed ________________?

4. Are you consuming ________________ more now than you have in the past?

5. Are you verbally abusive to others? How often? Why?

6. Are you physically abusive to others? How often? Why?

7. How often do you lie to others about your consumption of ________________?

8. Have you ever been embarrassed by your consumption of ________________?

9. Have you ever broken a promise to anybody that you would stop consuming ________________?

10. How often do you make excuses for, or try to justify your use of ___________?

In some cases, I ask the client if he or she would be willing to take a blood or urine analysis test, or see a substance abuse counselor for an initial evaluation. If the client has a substance abuse problem that goes undetected and untreated, the client is at serious risk for future legal ethics problems.

4. What The Client Needs To Know About Substance Abuse:

A. What Is Drug (Substance) Abuse?

Drug (or substance) abuse is the use of a mind-altering drug without medical need, in an amount large enough or over a period long enough to threaten the quality of life or health and safety of the user or others.

Drug abuse occurs in all socioeconomic groups and involves highly educated people as well as those who are uneducated and unemployed.
Not all substance abusers are drug dependent. For example, a person who infrequently gets drunk, is a drug abuser, but may not necessarily be drug dependent.

**Drug (or substance) dependent** people are those who are unable to control their compulsion to consume a mind-altering substance regardless of the negative consequences of that consumption. Like drug abuse, drug dependence occurs in all socioeconomic groups and involves highly educated people as well as those who are uneducated and unemployed.

Drug dependence is usually a chronic and progressive disease. In other words, it is usually of long duration, and it usually gets worse as time passes. Because drug dependence is a disease, it usually cannot be controlled without appropriate medical care and medical treatment.

For purposes of this presentation, I will often use the terms drug abuse, substance abuse, drug dependent, and substance dependent interchangeably. However, the actual distinction is set forth above.

**B. What Causes Drug Abuse?**

Nobody has definitively established the causes of drug abuse. Some studies show it has a hereditary component. Some studies show it is related to specific personality types. Other studies show it is caused by underlying mental health conditions.


The causes of substance abuse for an individual are multifactorial and involve a complex interplay of genetic
predisposition, environmental factors, and neurobiologic factors’.

While we do not know what causes substance abuse, we do know that it is not caused by a lack of morals or by a weak character. Throughout history, people of good morals and good character have engaged in substance abuse.

C. How Substance Abusers Can Find Help:

Substance abusers often fail to realize that the disease of substance abuse has control over them. Abusers frequently invent a “false reality”, in which they are in control of the disease and can stop abusing, “anytime I want”. As stated above, this “false reality” is called denial.

Cold Turkey

Denial is usually overcome in one of three ways. In the first, the drug abuser simply realizes he has a problem, and stops abusing and/or gets medical help. It is rare that an abuser comes to terms with the disease in this fashion, but it reportedly happens on occasion.

Hitting Bottom

The second way that denial is often overcome is when the drug abuser has a negative life experience, such as a very serious accident, that causes him or her to realize that he or she is controlled by the disease. Once aware that the disease controls his or her life, the abuser seeks out professional medical help and starts to recover.
**Intervention**

The third way that denial is overcome is through a procedure known as intervention. Intervention is a process of interrupting drug abuse by presenting reality to the abuser in a way that makes it understandable to him. Intervention may be the most popular form of overcoming denial.

5. **Intervention By Lawyer:**

If possible, and when appropriate, I convince my client that he or she has a problem and needs quality medical care to overcome their medical (substance abuse) problem.

Substance abuse is a medical disease that can be treated. Often it can be treated with great success. However, it is a medical problem and, therefore, almost always requires the services of a properly trained medical professional in order to get the disease under control.

In metropolitan areas, drug abuse treatment options are usually plentiful. Family physicians and mental health professionals can often be a source of information about treatment options.

Typically, I suggest that my substance abuse clients make the Ohio Lawyers Assistance Program (OLAP) part of their recovery team. It has been my experience that OLAP usually does an outstanding job helping my clients.

In 1977, OLAP was created to help lawyers with substance abuse problems, and now also helps lawyers with mental health disorders. OLAP typically does wonderful work. It offers help in the following areas:
1. Providing confidential advice about individual problems.
2. Providing help in arranging and implementing formal interventions.
3. Providing help in deciding between treatment options.
4. Providing monitoring and aftercare services.

As stated above, OLAP does wonderful work and has saved the professional (and the private) lives of countless judges and lawyers. Further, it is my opinion that OLAP, because of its established record of success and its long history of credibility, has substantial influence with Ohio's disciplinary authorities.

Per Ohio Rule of Professional Conduct 8.3(c), any knowledge obtained by OLAP (or another properly constituted bar association committee designed to assist lawyers with substance abuse or mental health problems) is privileged and confidential, provided the knowledge was obtained within the scope of official OLAP activities.

You can contact OLAP at:

Ohio Lawyers Assistance Program, Inc.
1650 Lake Shore Drive
Suite 375
Columbus, OH 43204-4991
Tel: 800-348-4343
Or 614-596-0621

Please note that when someone enters an OLAP program, OLAP will require him or her to sign a contract and it will insist that that person fulfill his or her obligations under the contract.
6. **The Law Of Substance Abuse In A Legal Ethics Case:**

According to Rule 10 of the Rules & Regulations Governing Procedure On Complaints and Hearings Before The Board Of Commissioners On Grievances and Discipline Of The Supreme Court of Ohio:

(B) In determining the appropriate sanction, the Board shall consider all relevant factors; precedent established by the Supreme Court of Ohio; and the following:

(2) Mitigation. The following shall not control the Board’s discretion, but may be considered in favor of recommending a less severe sanction:

***

  g) Chemical dependency or mental disability when there has been all of the following:

  (i) A diagnosis of a chemical dependency or mental disability by a qualified health care professional or alcohol/substance abuse counselor;

  (ii) A determination that the chemical dependency or mental disability contributed to cause the misconduct.

  (iii) In the event of chemical dependency, a certification of successful completion of an approved treatment program or in the event of mental disability, a sustained period of successful treatment.

  (iv) A prognosis from a qualified health care professional or alcohol-substance abuse counselor that the attorney will be able to return to competent, ethical professional practice under specified conditions.

  (h) other interim rehabilitation.

7. **To Sum Up:**

There are basically four parts to each substance abuse case for the respondent's lawyer. They are:
1. The lawyer and the client go through a Kabuki dance where the lawyer seeks to discover if the client has a substance abuse problem and the client seeks to avoid such discovery.

2. The lawyer encourages the client to get quality professional care and treatment for his or her substance abuse problem, and the client gets that help.

3. The lawyer seeks to obtain solid documentation that the client's substance abuse problem caused or significantly contributed to the client's alleged or actual misconduct.

4. The client gets better and can return to work as a competent, ethical and professional lawyer.

   Please note that the Supreme Court of Ohio has typically taken an enlightened approach to substance abuse – disciplinary cases provided the client has earned an enlightened approach.

8. **The Most Common Mistake Lawyers Make:**

   Some lawyers, typically those with little ethics experience, do not try to help their client recognize his or her substance abuse problem, and don't assist their client in getting quality professional help. All too often, these lawyers become enablers who make it easy for their clients to avoid dealing with their substance abuse problem. As indicated, this puts the client at serious risk.

   Good lawyers take the opposite approach. Once again, we turn to the great Gerry Spence for guidance. Spence was asked to represent Randy Weaver, a white
supremacist and anti-Semitic bigot with highly unusual fundamentalist religious beliefs, as a result of the Ruby Ridge incident.

Spence first met Randy Weaver when he visited him in Weaver's jail cell.

"My name is Gerry Spence," I began. "I'm the lawyer you've been told about. Before we begin to talk I want you to understand that I do not share your political or religious beliefs. Many of my dearest friends are Jews. My daughter is married to a Jew. My sister is married to a black man. She has adopted a black child. I deplore what the Nazis stand for. If I defend you I will not defend your political beliefs or your religious beliefs, but your rights as an American citizen to a fair trial. Spence, From Freedom To Slavery (1993), pages 4-5. (Emphasis added).