

Zubulake v. UBS Warburg LLC

02 Civ. 1243 (SAS)

**UNITED STATES DISTRICT COURT FOR THE SOUTHERN DISTRICT OF
NEW YORK***217 F.R.D. 309; 2003 U.S. Dist. LEXIS 7939; 55 Fed. R. Serv. 3d (Callaghan) 622; 91
Fair Empl. Prac. Cas. (BNA) 1574***May 13, 2003, Decided****May 13, 2003, Filed**

SUBSEQUENT HISTORY: Motion denied by *Zubulake v. UBS Warburg LLC*, 2003 U.S. Dist. LEXIS 7940 (S.D.N.Y., May 13, 2003)

DISPOSITION: Court instructed defendants to provide emails from all optical disks and servers and from select backup tapes.

COUNSEL: **[**1]** For Plaintiff: James A. Batson, Esq., Christina J. Kang, Esq., Liddle & Robinson, LLP, New York, New York.

For Defendants: Kevin B. Leblang, Esq., Norman C. Simon, Esq., Kramer Levin Naftalis & Frankel LLP, New York, New York.

JUDGES: Shira A. Scheindlin, U.S.D.J.

OPINION BY: SHIRA A. SCHEINDLIN

OPINION**[*311] OPINION AND ORDER**

SHIRA A. SCHEINDLIN, U.S.D.J.:

The world was a far different place in 1849, when Henry David Thoreau opined (in an admittedly broader context) that "the process of discovery is very simple."¹ That hopeful maxim has given way to rapid technological advances, requiring new solutions to old problems. The issue presented here is one such problem, recast in light of current technology: To what extent is inaccessible electronic data discoverable, and who should pay for its

production?

¹ Henry David Thoreau, *A Week on the Concord and Merrimack Rivers* (1849).

I. INTRODUCTION

The Supreme Court recently reiterated that our "simplified notice pleading standard relies on liberal **[**2]** discovery rules and summary judgment motions to define disputed facts and issues and to dispose of unmeritorious claims."² Thus, it is now beyond dispute that "broad discovery is a cornerstone of the litigation process contemplated by the Federal Rules of Civil Procedure."³ The Rules contemplate a minimal burden to bringing a claim; that claim is then fleshed out through vigorous and expansive discovery.⁴

² *Swierkiewicz v. Sorema, N.A.*, 534 U.S. 506, 512, 152 L. Ed. 2d 1, 122 S. Ct. 992 (2002).

³ *Jones v. Goord*, 2002 U.S. Dist. LEXIS 8707, No. 95 Civ. 8026, 2002 WL 1007614, at *1 (S.D.N.Y. May 16, 2002).

⁴ See *Hickman v. Taylor*, 329 U.S. 495, 500-01, 91 L. Ed. 451, 67 S. Ct. 385 (1947).

In one context, however, the reliance on broad discovery has hit a roadblock. As individuals and corporations increasingly do business electronically⁵ -- using computers to create and store documents, make deals, and exchange e-mails -- the universe of discoverable material has expanded **[**3]** exponentially.⁶ The more information there is to discover, the more expensive it is to discover all the relevant information until, in the end, "discovery is not just about uncovering

the truth, but also about how much of the truth the parties can afford to disinter." ⁷

⁵ See Wendy R. Liebowitz, *Digital Discovery Starts to Work*, Nat'l L.J., Nov. 4, 2002, at 4 (reporting that in 1999, ninety-three percent of all information generated was in digital form).

⁶ *Rowe Entm't, Inc. v. William Morris Agency, Inc.*, 205 F.R.D. 421, 429 (S.D.N.Y. 2002) (explaining that electronic data is so voluminous because, unlike paper documents, "the costs of storage are virtually nil. Information is retained not because it is expected to be used, but because there is no compelling reason to discard it"), *aff'd*, 2002 U.S. Dist. LEXIS 8308, 2002 WL 975713 (S.D.N.Y. May 9, 2002).

⁷ *Rowe*, 205 F.R.D. at 423.

This case provides a textbook example of the difficulty of balancing the competing **[**4]** needs of broad discovery and manageable costs. Laura Zubulake is suing UBS Warburg LLC, UBS Warburg, and UBS AG (collectively, "UBS" or the "Firm") under Federal, State and City law for gender discrimination and illegal retaliation. Zubulake's case is certainly not frivolous ⁸ and if she prevails, her damages may be substantial. ⁹ She contends that key evidence is located **[*312]** in various e-mails exchanged among UBS employees that now exist only on backup tapes and perhaps other archived media. According to UBS, restoring those e-mails would cost approximately \$ 175,000.00, exclusive of attorney time in reviewing the e-mails. ¹⁰ Zubulake now moves for an order compelling UBS to produce those e-mails at its expense. ¹¹

⁸ Indeed, Zubulake has already produced a sort of "smoking gun": an e-mail suggesting that she be fired "ASAP" after her EEOC charge was filed, in part so that she would not be eligible for year-end bonuses. See 8/21/01 e-mail from Mike Davies to Rose Tong ("8/21/01 e-Mail"), Ex. G to the 3/17/03 Affirmation of James A. Batson, counsel for Zubulake ("Batson Aff.").

⁹ At the time she was terminated, Zubulake's annual salary was approximately \$ 500,000. Were she to receive full back pay and front pay, Zubulake estimates that she may be entitled to as much as \$ 13,000,000 in damages, not including any punitive damages or attorney's fees. See Memorandum of Law in Support of Plaintiff's

Motion for an Order Compelling Defendants to Produce E-mails, Permitting Disclosure of Deposition Transcript and Directing Defendants to Bear Certain Expenses ("Pl. Mem.") at 2-3.

[5]**

¹⁰ See 3/26/03 Oral Argument Transcript ("3/26/03 Tr.") at 14, 44-45.

¹¹ Zubulake also moves for an order (1) directing UBS to pay for the cost of deposing Christopher Behny, UBS's information technology expert and (2) permitting her to disclose the transcript of Behny's deposition to certain securities regulators. Those motions are denied in a separate Opinion and Order issued today.

II. BACKGROUND

A. Zubulake's Lawsuit

UBS hired Zubulake on August 23, 1999, as a director and senior salesperson on its U.S. Asian Equities Sales Desk (the "Desk"), where she reported to Dominic Vail, the Desk's manager. At the time she was hired, Zubulake was told that she would be considered for Vail's position if and when it became vacant.

In December 2000, Vail indeed left his position to move to the Firm's London office. But Zubulake was not considered for his position, and the Firm instead hired Matthew Chapin as director of the Desk. Zubulake alleges that from the outset Chapin treated her differently than the other members of the Desk, all of whom were male. In particular, **[**6]** Chapin "undermined Ms. Zubulake's ability to perform her job by, *inter alia*: (a) ridiculing and belittling her in front of co-workers; (b) excluding her from work-related outings with male co-workers and clients; (c) making sexist remarks in her presence; and (d) isolating her from the other senior salespersons on the Desk by seating her apart from them." ¹² No such actions were taken against any of Zubulake's male co-workers.

¹² Pl. Mem. at 2.

Zubulake ultimately responded by filing a Charge of (gender) Discrimination with the EEOC on August 16, 2001. On October 9, 2001, Zubulake was fired with two weeks' notice. On February 15, 2002, Zubulake filed the instant action, suing for sex discrimination and retaliation under *Title VII*, the *New York State Human Rights Law*,

and the *Administrative Code of the City of New York*. UBS timely answered on March 12, 2002, denying the allegations. UBS's argument is, in essence, that Chapin's conduct was not unlawfully discriminatory because he treated everyone equally [**7] badly. On the one hand, UBS points to evidence that Chapin's anti-social behavior was not limited to women: a former employee made allegations of national origin discrimination against Chapin, and a number of male employees on the Desk also complained about him. On the other hand, Chapin was responsible for hiring three new female employees to the Desk.¹³

13 See Defendants' Memorandum of Law in Opposition to Plaintiff's Motion for an Order Compelling Defendants to Produce E-Mails, Permitting Disclosure of Deposition Transcript and Directing Defendants to Bear Certain Expenses ("Def. Mem.") at 2.

B. The Discovery Dispute

Discovery in this action commenced on or about June 3, 2002, when Zubulake served UBS with her first document request. At issue here is request number twenty-eight, for "all documents concerning any communication by or between UBS employees concerning Plaintiff."¹⁴ The term document in Zubulake's request "includes, without limitation, electronic or computerized data compilations. [**8] " On July 8, 2002, UBS responded by producing approximately 350 pages of documents, including approximately 100 [*313] pages of e-mails. UBS also objected to a substantial portion of Zubulake's requests.¹⁵

14 Plaintiff's First Request for Production of Documents P 28, Ex. E to the Declaration of Kevin B. Leblang, counsel to UBS ("Leblang Dec.").

15 See Defendants' Response to Plaintiff's First Request for Production of Documents, Ex. F to the Leblang Dec.

On September 12, 2002 -- after an exchange of angry letters¹⁶ and a conference before United States Magistrate Judge Gabriel W. Gorenstein -- the parties reached an agreement (the "9/12/02 Agreement"). With respect to document request twenty-eight, the parties reached the following agreement, in relevant part:

Defendants will [] ask UBS about how to retrieve e-mails that are saved in the firm's computer system *and will produce responsive e-mails if retrieval is possible* and Plaintiff names a few individuals.¹⁷

Pursuant to the [**9] 9/12/02 Agreement, UBS agreed unconditionally to produce responsive e-mails from the accounts of five individuals named by Zubulake: Matthew Chapin, Rose Tong (a human relations representation who was assigned to handle issues concerning Zubulake), Vinay Datta (a co-worker on the Desk), Andrew Clarke (another co-worker on the Desk), and Jeremy Hardisty (Chapin's supervisor and the individual to whom Zubulake originally complained about Chapin). UBS was to produce such e-mails sent between August 1999 (when Zubulake was hired) and December 2001 (one month after her termination), to the extent possible.

16 See Exs. G and H to the Leblang Dec.

17 9/18/02 Letter from James A. Batson to Kevin B. Leblang, Ex. I to the Leblang Dec. (emphasis added). See also 9/25/02 Letter from Kevin B. Leblang to James A. Batson, Ex. K to the Leblang Dec. (confirming the above as the parties' agreement).

UBS, however, produced no additional e-mails and insisted that its initial production (the 100 pages of e-mails) was [**10] complete. As UBS's opposition to the instant motion makes clear -- although it remains unsaid -- UBS never searched for responsive e-mails on any of its backup tapes. To the contrary, UBS informed Zubulake that the cost of producing e-mails on backup tapes would be prohibitive (estimated at the time at approximately \$ 300,000.00).¹⁸

18 See 3/26/03 Tr. at 14 (Statement of Kevin B. Leblang).

Zubulake, believing that the 9/12/02 Agreement included production of e-mails from backup tapes, objected to UBS's nonproduction. In fact, Zubulake *knew* that there were additional responsive e-mails that UBS had failed to produce because she herself had produced approximately 450 pages of e-mail correspondence. Clearly, numerous responsive e-mails had been created and deleted¹⁹ at UBS, and Zubulake wanted them.

19 The term "deleted" is sticky in the context of

electronic data. "Deleting' a file does not actually erase that data from the computer's storage devices. Rather, it simply finds the data's entry in the disk directory and changes it to a 'not used' status -- thus permitting the computer to write over the 'deleted' data. Until the computer writes over the 'deleted' data, however, it may be recovered by searching the disk itself rather than the disk's directory. Accordingly, many files are recoverable long after they have been deleted -- even if neither the computer user nor the computer itself is aware of their existence. Such data is referred to as 'residual data.'" Shira A. Scheindlin & Jeffrey Rabkin, *Electronic Discovery in Federal Civil Litigation: Is Rule 34 Up to the Task?*, 41 B.C. L. Rev. 327, 337 (2000) (footnotes omitted). Deleted data may also exist because it was backed up before it was deleted. Thus, it may reside on backup tapes or similar media. Unless otherwise noted, I will use the term "deleted" data to mean residual data, and will refer to backed-up data as "backup tapes."

[**11] On December 2, 2002, the parties again appeared before Judge Gorenstein, who ordered UBS to produce for deposition a person with knowledge of UBS's e-mail retention policies in an effort to determine whether the backup tapes contained the deleted e-mails and the burden of producing them. In response, UBS produced Christopher Behny, Manager of Global Messaging, who was deposed on January 14, 2003. Mr. Behny testified to UBS's e-mail backup protocol, and also to the cost of restoring the relevant data.

C. UBS's E-Mail Backup System

In the first instance, the parties agree that e-mail was an important means of communication at UBS during the relevant time period. Each salesperson, including the salespeople on the Desk, received approximately [*314] 200 e-mails each day.²⁰ Given this volume, and because Securities and Exchange Commission regulations require it,²¹ UBS implemented extensive email backup and preservation protocols. In particular, e-mails were backed up in two distinct ways: on backup tapes and on optical disks.

²⁰ See 3/26/03 Tr. at 14 (Statement of Kevin B. Leblang).

²¹ SEC Rule 17a-4, promulgated pursuant to Section 17(a) of the Securities Exchange Act of

1934, provides in pertinent part:

Every [] broker and dealer shall preserve for a period of not less than 3 years, the first two years in an accessible place . . . originals of all communications received and copies of all communications sent by such member, broker or dealer (including inter-office memoranda and communications) relating to his business as such.

17 C.F.R. § 240.17a-4(b) and (4).

[**12] 1. Backup Tape Storage

UBS employees used a program called HP OpenMail, manufactured by Hewlett-Packard,²² for all work-related e-mail communications.²³ With limited exceptions, all e-mails sent or received by any UBS employee are stored onto backup tapes. To do so, UBS employs a program called Veritas NetBackup,²⁴ which creates a "snapshot" of all e-mails that exist on a given server at the time the backup is taken. Except for scheduling the backups and physically inserting the tapes into the machines, the backup process is entirely automated.

²² Hewlett-Packard has since discontinued sales of HP OpenMail, although the company still supports the product and permits existing customers to purchase new licenses. See <http://www.openmail.com/>.

²³ See 1/14/03 Deposition of Christopher Behny ("Behny Dep."), Ex. M to the Leblang Dec. Unless otherwise noted, all information about UBS's e-mail systems is culled from the Behny Dep. Because that document has been sealed, repeated pin cites are unnecessary and thus omitted.

²⁴ See generally VERITAS NetBackup Release 4.5 Technical Overview, available at <http://www.veritas.com>.

[**13] UBS used the same backup protocol during the entire relevant time period, from 1999 through 2001. Using NetBackup, UBS backed up its e-mails at three intervals: (1) daily, at the end of each day, (2) weekly, on Friday nights, and (3) monthly, on the last business day of the month. Nightly backup tapes were kept for twenty

working days, weekly tapes for one year, and monthly tapes for three years. After the relevant time period elapsed, the tapes were recycled.²⁵

25 Of course, periodic backups such as UBS's necessarily entails the loss of certain e-mails. Because backups were conducted only intermittently, some e-mails that were deleted from the server were never backed up. For example, if a user both received *and* deleted an e-mail on the same day, it would not reside on any backup tape. Similarly, an e-mail received and deleted within the span of one month would not exist on the monthly backup, although it might exist on a weekly or daily backup, if those tapes still exist. As explained below, if an e-mail was to or from a "registered trader," however, it may have been stored on UBS's optical storage devices.

[**14] Once e-mails have been stored onto backup tapes, the restoration process is lengthy. Each backup tape routinely takes approximately five days to restore, although resort to an outside vendor would speed up the process (at greatly enhanced costs, of course). Because each tape represents a snapshot of one server's hard drive in a given month, each server/month must be restored separately onto a hard drive. Then, a program called Double Mail is used to extract a particular individual's e-mail file. That mail file is then exported into a Microsoft Outlook data file, which in turn can be opened in Microsoft Outlook, a common e-mail application. A user could then browse through the mail file and sort the mail by recipient, date or subject, or search for key words in the body of the e-mail.

Fortunately, NetBackup also created indexes of each backup tape. Thus, Behny was able to search through the tapes from the relevant time period and determine that the e-mail files responsive to Zubulake's requests are contained on a total of ninety-four backup tapes.

2. Optical Disk Storage

In addition to the e-mail backup tapes, UBS also stored certain e-mails on optical [*315] disks. For certain "registered [*15] traders," probably including the members of the Desk,²⁶ a copy of *all* e-mails sent to or received from outside sources (*i.e.*, e-mails from a "registered trader" at UBS to someone at another entity, or vice versa) was simultaneously written onto a series of

optical disks. Internal e-mails, however, were not stored on this system.

26 In using the phrase "registered trader," Behny referred to individuals designated to have their e-mails archived onto optical disks. Although Behny could not be certain that such a designation corresponds to Series 7 or Series 63 broker-dealers, he indicated that examples of registered traders include "equity research people, [and] equity traders type people." *See* Behny Dep. at 35. He admitted that members of the Desk were probably "registered" in that sense:

Q: Do you know whether the Asian Equities Sales desk was registered to keep a secondary copy in 1999?

A: I can't say conclusively.

Q: Do you have an opinion?

A: My opinion is yes.

Id. at 36. *See also id.* (admitting that the same was probably true in 2000 and 2001).

[**16] UBS has retained each optical disk used since the system was put into place in mid-1998. Moreover, the optical disks are neither erasable nor rewritable. Thus, UBS has *every* e-mail sent or received by registered traders (except internal emails) during the period of Zubulake's employment, even if the e-mail was deleted instantaneously on that trader's system.

The optical disks are easily searchable using a program called Tumbleweed.²⁷ Using Tumbleweed, a user can simply log into the system with the proper credentials and create a plain language search. Search criteria can include not just "header" information, such as the date or the name of the sender or recipient, but can also include terms within the text of the email itself. For example, UBS personnel could easily run a search for e-mails containing the words "Laura" or "Zubulake" that were sent or received by Chapin, Datta, Clarke, or Hardisty.²⁸

27 *See generally* <http://www.tumbleweed.com/en/products/solutions/archive.html>.

28 Rose Tong, the fifth person designated by Zubulake's document request, would probably *not* have been a "registered trader" as she was a human resources employee.

[**17] III. LEGAL STANDARD

Federal Rules of Civil Procedure 26 through 37 govern discovery in all civil actions. As the Supreme Court long ago explained,

The pre-trial deposition-discovery mechanism established by *Rules 26 to 37* is one of the most significant innovations of the Federal Rules of Civil Procedure. Under the prior federal practice, the pre-trial functions of notice-giving issue-formulation and fact-revelation were performed primarily and inadequately by the pleadings. Inquiry into the issues and the facts before trial was narrowly confined and was often cumbersome in method. The new rules, however, restrict the pleadings to the task of general notice-giving and invest the deposition-discovery process with a vital role in the preparation for trial. 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 [**18] with recognized privileges, for the parties to obtain the *fullest possible* knowledge of the issues and facts before trial.²⁹

Consistent with this approach, *Rule 26(b)(1)* specifies that,

Parties may obtain discovery regarding *any matter*, not privileged, *that is relevant* to the claim or defense of any party, including the existence, description, nature, custody, condition, and location of any books, documents, or other tangible

things and the identity and location of persons having knowledge of any discoverable matter. For good cause, the court may order discovery of any matter relevant to the subject matter involved in the action. Relevant information need not be admissible [**316] at the trial if the discovery appears reasonably calculated to lead to the discovery of admissible evidence. All discovery is subject to the limitations imposed by Rule 26(b)(2)(i), (ii), and (iii).³⁰

29 *Hickman*, 329 U.S. at 500-01 (emphasis added).

30 *Fed. R. Civ. P. 26(b)(1)* (emphasis added).

[**19] In turn, *Rule 26(b)(2)* imposes general limitations on the scope of discovery in the form of a "proportionality test":

The frequency or extent of use of the discovery methods otherwise permitted under these rules and by any local rule shall be limited by the court if it determines that: (i) the discovery sought is unreasonably cumulative or duplicative, or is obtainable from some other source that is more convenient, less burdensome, or less expensive; (ii) the party seeking discovery has had ample opportunity by discovery in the action to obtain the information sought; or (iii) the burden or expense of the proposed discovery outweighs its likely benefit, taking into account the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues.

31

Finally, "under [the discovery] rules, the presumption is that the responding party must bear the expense of complying with discovery requests, but [it] may invoke the district court's discretion under *Rule 26(c)* to grant orders protecting [it] from 'undue burden or expense' in doing so, [**20] including orders conditioning

discovery on the requesting party's payment of the costs of discovery." ³²

31 *Fed. R. Civ. P. 26(b)(2)*.

32 *Oppenheimer Fund, Inc. v. Sanders*, 437 U.S. 340, 358, 57 L. Ed. 2d 253, 98 S. Ct. 2380 (1978).

The application of these various discovery rules is particularly complicated where electronic data is sought because otherwise discoverable evidence is often only available from expensive-to-restore backup media. That being so, courts have devised creative solutions for balancing the broad scope of discovery prescribed in *Rule 26(b)(1)* with the cost-consciousness of *Rule 26(b)(2)*. By and large, the solution has been to consider cost-shifting: forcing the requesting party, rather than the answering party, to bear the cost of discovery.

By far, the most influential response to the problem of cost-shifting relating to the discovery of electronic data was given by United States Magistrate Judge James C. Francis IV of this district in *Rowe Entertainment*. Judge **[**21]** Francis utilized an eight-factor test to determine whether discovery costs should be shifted. Those eight factors are:

(1) the specificity of the discovery requests; (2) the likelihood of discovering critical information; (3) the availability of such information from other sources; (4) the purposes for which the responding party maintains the requested data; (5) the relative benefits to the parties of obtaining the information; (6) the total cost associated with production; (7) the relative ability of each party to control costs and its incentive to do so; and (8) the resources available to each party. ³³

Both Zubulake and UBS agree that the eight-factor *Rowe* test should be used to determine whether cost-shifting is appropriate. ³⁴

33 *205 F.R.D. at 429*.

34 Zubulake mistakenly identifies the *Rowe* test as a "marginal utility" test. In fact, "marginal utility" -- a common term among economists, see Istvan Meszaros, *Beyond Capital* § 3.2 (1995) (describing the intellectual history of marginal utility) -- refers only to the second *Rowe* factor, the likelihood of discovering critical information.

See *Rowe*, 205 F.R.D. at 430 (quoting *McPeck v. Ashcroft*, 202 F.R.D. 31, 34 (D.D.C. 2001)).

[22] IV. DISCUSSION**

A. Should Discovery of UBS's Electronic Data Be Permitted?

Under *Rule 34*, a party may request discovery of any document, "including writings, drawings, graphs, charts, photographs, phonorecords, and other data compilations. . . ." ³⁵ The "inclusive description" of the term document "accords with changing [***317**] technology." ³⁶ "It makes clear that *Rule 34* applies to *electronics* [sic] data compilations." Thus, "electronic documents" are no less subject to disclosure than paper records." ³⁷ This is true not only of electronic documents that are currently in use, but also of documents that may have been deleted and now reside only on backup disks. ³⁸

35 *Fed. R. Civ. P. 34(a)*.

36 Advisory Committee Note to *Fed. R. Civ. P. 34*.

37 *Rowe*, 205 F.R.D. at 428 (collecting cases).

38 See *Antioch Co. v. Scrapbook Borders, Inc.*, 210 F.R.D. 645, 652 (D. Minn. 2002) ("It is a well accepted proposition that deleted computer files, whether they be e-mails or otherwise, are discoverable."); *Simon Property Group L.P. v. MySimon, Inc.*, 194 F.R.D. 639, 640 (S.D. Ind. 2000) ("First, computer records, including records that have been 'deleted,' are documents discoverable under *Fed. R. Civ. P. 34*").

[23]** That being so, Zubulake is entitled to discovery of the requested e-mails so long as they are relevant to her claims, ³⁹ which they clearly are. As noted, e-mail constituted a substantial means of communication among UBS employees. To that end, UBS has already produced approximately 100 pages of e-mails, the contents of which are unquestionably relevant. ⁴⁰

39 See *Fed. R. Civ. P. 26(b)(1)*.

40 See, e.g., 8/21/01 e-Mail.

Nonetheless, UBS argues that Zubulake is not entitled to any further discovery because it already produced all responsive documents, to wit, the 100 pages of e-mails. This argument is unpersuasive for two reasons. *First*, because of the way that UBS backs up its

e-mail files, it clearly could not have searched all of its e-mails without restoring the ninety-four backup tapes (which UBS admits that it has not done). UBS therefore cannot represent that it has produced all responsive emails. *Second*, Zubulake herself has produced over 450 pages of relevant e-mails, including [**24] e-mails that would have been responsive to her discovery requests but were never produced by UBS. These two facts strongly suggest that there are e-mails that Zubulake has not received that reside on UBS's backup media.⁴¹

41 UBS insists that "from the time Plaintiff commenced her EEOC action in August 2001 . . . UBS collected and produced all existing responsive e-mails sent or received between 1999 and 2001 from these and other employees' computers." Def. Mem. at 6. Even if this statement is completely accurate, a simple search of employees' computer files would not have turned up e-mails deleted prior to August 2001. Such deleted documents exist *only* on the backup tapes and optical disks, and their absence is precisely why UBS's production is not complete.

B. Should Cost-Shifting Be Considered?

Because it apparently recognizes that Zubulake is entitled to the requested discovery, UBS expends most of its efforts urging the court to shift the cost of production to "protect [it] . . . from undue [**25] burden or expense."⁴² Faced with similar applications, courts generally engage in some sort of cost-shifting analysis, whether the refined eight-factor *Rowe* test or a cruder application of *Rule 34's* proportionality test, or something in between.⁴³

42 Def. Mem. at 9 (quoting *Fed. R. Civ. P. 26(c)*).

43 See, e.g., *Byers v. Illinois State Police*, 2002 U.S. Dist. LEXIS 9861, No. 99 C. 8105, 2002 WL 1264004 (N.D. Ill. June 3, 2002); *In re Bristol-Myers Squibb Sec. Litig.*, 205 F.R.D. 437, 443 (D.N.J. 2002); *Rowe*, 205 F.R.D. 421; *McPeck*, 202 F.R.D. 31.

The first question, however, is whether cost-shifting must be considered in every case involving the discovery of electronic data, which -- in today's world -- includes virtually all cases. In light of the accepted principle, stated above, that electronic evidence is no less discoverable than paper evidence, the answer is, "No."

The Supreme Court has instructed that "the presumption is that the responding party [**26] must bear the expense of complying with discovery requests. . . ." ⁴⁴ Any principled approach to electronic evidence must respect this presumption.

44 *Oppenheimer Fund*, 437 U.S. at 358.

Courts must remember that cost-shifting may effectively end discovery, especially when private parties are engaged in litigation with large corporations. As large companies increasingly move to entirely paper-free environments, [**318] the frequent use of cost-shifting will have the effect of crippling discovery in discrimination and retaliation cases. This will both undermine the "strong public policy favoring resolving disputes on their merits," ⁴⁵ and may ultimately deter the filing of potentially meritorious claims.

45 *Pecarsky v. Galaxiworld.com, Inc.*, 249 F.3d 167, 172 (2d Cir. 2001).

Thus, cost-shifting should be considered [**27] *only* when electronic discovery imposes an "undue burden or expense" on the responding party. ⁴⁶ The burden or expense of discovery is, in turn, "undue" when it "outweighs its likely benefit, taking into account the needs of the case, the amount in controversy, the parties' resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues." ⁴⁷

46 *Fed. R. Civ. P. 26(c)*.

47 *Fed. R. Civ. P. 26(b)(2)(iii)*. As noted, a court is also permitted to impose conditions on discovery when it might be duplicative, *see Fed. R. Civ. P. 26(b)(2)(i)*, or when a reasonable discovery deadline has lapsed, *see id. 26(b)(2)(ii)*. Neither of these concerns, however, is likely to arise solely because the discovery sought is of electronic data.

Many courts have automatically assumed that an undue burden or expense may arise simply because electronic evidence is involved. ⁴⁸ This makes no sense. Electronic evidence is frequently cheaper and easier [**28] to produce than paper evidence because it can be searched automatically, key words can be run for privilege checks, and the production can be made in electronic form obviating the need for mass photocopying. ⁴⁹

48 See, e.g., *Murphy Oil USA, Inc. v. Fluor Daniel, Inc.*, 2002 U.S. Dist. LEXIS 3196, No. Civ.A. 99-3564, 2002 WL 246439, at *3 (E.D. La. Feb. 19, 2002) (suggesting that application of *Rowe* is appropriate whenever "a party, as does Flour [sic], contends that the burden or expense of the discovery outweighs the benefit of the discovery").

49 See generally Scheindlin & Rabkin, *Electronic Discovery*, 41 *B.C. L. Rev.* at 335-341 (describing types of discoverable electronic data and their differences from paper evidence).

In fact, whether production of documents is unduly burdensome or expensive turns primarily on whether it is kept in an *accessible or inaccessible* format (a distinction that corresponds closely to the expense of production). In the world of paper documents, for **[**29]** example, a document is accessible if it is readily available in a usable format and reasonably indexed. Examples of inaccessible paper documents could include (a) documents in storage in a difficult to reach place; (b) documents converted to microfiche and not easily readable; or (c) documents kept haphazardly, with no indexing system, in quantities that make page-by-page searches impracticable. But in the world of electronic data, thanks to search engines, any data that is retained in a machine readable format is typically accessible.⁵⁰

50 See Scheindlin & Rabkin, *Electronic Discovery*, 41 *B.C. L. Rev.* at 364 ("By comparison [to the time it would take to search through 100,000 pages of paper], the average office computer could search all of the documents for specific words or combinations of words in minute, perhaps less."); see also *Public Citizen v. Carlin*, 337 U.S. App. D.C. 320, 184 F.3d 900, 908-10 (D.C. Cir. 1999).

Whether electronic data is accessible or inaccessible **[**30]** turns largely on the media on which it is stored. Five categories of data, listed in order from most accessible to least accessible, are described in the literature on electronic data storage:

1. *Active, online data*: "On-line storage is generally provided by magnetic disk. It is used in the very active stages of an electronic records [sic] life -- when it is being created or received and processed,

as well as when the access frequency is high and the required speed of access is very fast, i.e., milliseconds."⁵¹ Examples of online data include hard drives.

2. *Near-line data*: "This typically consists of a robotic storage device (robotic library) that houses removable media, uses robotic arms to access the media, and uses multiple read/write devices to store and retrieve records. Access speeds can range from as low **[*319]** as milliseconds if the media is already in a read device, up to 10-30 seconds for optical disk technology, and between 20-120 seconds for sequentially searched media, such as magnetic tape."⁵² Examples include optical disks.

3. *Offline storage/archives*: "This is removable optical disk or magnetic tape media, which can be labeled and stored **[**31]** in a shelf or rack. Off-line storage of electronic records is traditionally used for making disaster copies of records and also for records considered 'archival' in that their likelihood of retrieval is minimal. Accessibility to off-line media involves manual intervention and is much slower than on-line or near-line storage. Access speed may be minutes, hours, or even days, depending on the access-effectiveness of the storage facility."⁵³ The principled difference between nearline data and offline data is that offline data lacks "the coordinated control of an intelligent disk subsystem," and is, in the lingo, JBOD ("Just a Bunch Of Disks").⁵⁴

4. *Backup tapes*: "A device, like a tape recorder, that reads data from and writes it onto a tape. Tape drives have data capacities of anywhere from a few hundred kilobytes to several gigabytes. Their transfer speeds also vary considerably. . . The disadvantage of tape drives is that they are sequential-access devices, which means that to read any particular block of data, you need to read

all the preceding blocks." ⁵⁵ As a result, "the data on a backup tape are not organized for retrieval of individual documents or files [because] **[**32]** . . . the organization of the data mirrors the computer's structure, not the human records management structure." ⁵⁶ Backup tapes also typically employ some sort of data compression, permitting more data to be stored on each tape, but also making restoration more time-consuming and expensive, especially given the lack of uniform standard governing data compression. ⁵⁷

5. *Erased, fragmented or damaged data*: "When a file is first created and saved, it is laid down on the [storage media] in contiguous clusters. . . As files are erased, their clusters are made available again as free space. Eventually, some newly created files become larger than the remaining contiguous free space. These files are then broken up and randomly placed throughout the disk." ⁵⁸ Such broken-up files are said to be "fragmented," and along with damaged and erased data can only be accessed after significant processing. ⁵⁹

Of these, the first three categories are typically identified as accessible, and the latter **[*320]** two as inaccessible. ⁶⁰ The difference between the two classes is easy to appreciate. Information deemed "accessible" is stored in a readily usable format. Although the time it **[**33]** takes to actually access the data ranges from milliseconds to days, the data does not need to be restored or otherwise manipulated to be usable. "Inaccessible" data, on the other hand, is not readily usable. Backup tapes must be restored using a process similar to that previously described, fragmented data must be de-fragmented, and erased data must be reconstructed, all before the data is usable. That makes such data inaccessible. ⁶¹

51 Cohasset Associates, Inc., *White Paper: Trustworthy Storage and Management of Electronic Records: The Role of Optical Storage Technology* 10 (April 2003) ("White Paper").

52 *Id.* at 11.

53 *Id.*

54 CNT, *The Future of Tape 2*, available at <http://www.cnt.com/literature/documents/pl556.pdf>.

55 *Webopedia*, at http://inews.webopedia.com/TERM/t/tape_drive.html.

56 Kenneth J. Withers, *Computer-Based Discovery in Federal Civil Litigation* (unpublished manuscript) at 15.

57 *See generally* SDLT, Inc., *Making a Business Case for Tape*, at http://quantum.treehousei.com/Surveys/publishing/survey_148/pdfs/making_a_business_case_for_tape.pdf (June 2002); Jerry Stern, *The Perils of Backing Up*, at http://www.grsoftware.net/backup/article/s/jerry_perils.html (last visited May 5, 2003).

[34]**

58 Sunbelt Software, Inc., *White Paper: Disk Defragmentation for Windows NT/2000: Hidden Gold for the Enterprise 2*, at <http://www.sunbelt-software.com/evaluation/455/web/documents/idcwhite-paper-english.pdf> (last visited May 5, 2003).

59 *See* Executive Software, Inc., *Identifying Common Reliability/Stability Problems Caused by File Fragmentation*, at http://www.execsoft.com/Reliability_Stability_Whitepaper.pdf (last visited May 1, 2003) (identifying problems associated with file fragmentation, including file corruption, data loss, crashes, and hard drive failures); Stan Miastkowski, *When Good Data Goes Bad*, PC World, Jan. 2000, available at <http://www.pcworld.com/resource/printable/article/0,aid,13859,00.asp>.

60 *See generally* White Paper 10-13.

61 A report prepared by the Sedona Conference recently propounded "Best Practices" for electronic discovery. *See* The Sedona Conference, *The Sedona Principles: Best Practices Recommendations & Principles for Addressing Electronic Document Production* (March 2003), ("Sedona Principles"), available at http://www.thesedonaconference.org/publications_html. Although I do not endorse or indeed agree with all of the *Sedona Principles*, they do recognize the difference between "active data" and data stored on backup tapes or "deleted, shadowed, fragmented or residual data," *see id.*

(Principles 8 and 9), a distinction very similar to the accessible/inaccessible test employed here.

[**35] The case at bar is a perfect illustration of the range of accessibility of electronic data. As explained above, UBS maintains e-mail files in three forms: (1) active user e-mail files; (2) archived e-mails on optical disks; and (3) backup data stored on tapes. The active (HP OpenMail) data is obviously the most accessible: it is online data that resides on an active server, and can be accessed immediately. The optical disk (Tumbleweed) data is only slightly less accessible, and falls into either the second or third category. The e-mails are on optical disks that need to be located and read with the correct hardware, but the system is configured to make searching the optical disks simple and automated once they are located. For these sources of e-mails -- active mail files and e-mails stored on optical disks -- it would be wholly inappropriate to even consider cost-shifting. UBS maintains the data in an accessible and usable format, and can respond to Zubulake's request cheaply and quickly. Like most typical discovery requests, therefore, the producing party should bear the cost of production.

E-mails stored on backup tapes (via NetBackup), however, are an entirely different matter. [**36] Although UBS has already identified the ninety-four potentially responsive backup tapes, those tapes are not currently accessible. In order to search the tapes for responsive e-mails, UBS would have to engage in the costly and time-consuming process detailed above. It is therefore appropriate to *consider* cost shifting.

C. What Is the Proper Cost-Shifting Analysis?

In the year since *Rowe* was decided, its eight factor test has unquestionably become the gold standard for courts resolving electronic discovery disputes.⁶² But there is little doubt that the *Rowe* factors will generally favor cost-shifting. Indeed, of the handful of reported opinions that apply *Rowe* or some modification thereof, *all of them* have ordered the cost of discovery to be shifted to the requesting party.⁶³

62 See *In re Livent, Inc. Noteholders Sec. Litig.*, 211 F.R.D. 219, 2003 WL 23254, at *3 (S.D.N.Y. 2003) ("the attorneys should read Magistrate Judge Francis's opinion in [*Rowe*]. Then Deloitte and plaintiffs should confer, in person or by telephone, and discuss the eight factors listed in that opinion."); *Bristol-Myers Squibb*, 205 F.R.D.

at 443 ("For a more comprehensive analysis of cost allocation and cost shifting regarding production of electronic information in a different factual context, counsel are directed to the recent opinion in [*Rowe*].").

[**37]

63 See *Murphy Oil*, 2002 U.S. Dist. LEXIS 3196, 2002 WL 246439; *Bristol-Myers Squibb*, 205 F.R.D. 437; *Byers*, 2002 U.S. Dist. LEXIS 9861, 2002 WL 1264004.

In order to maintain the presumption that the responding party pays, the cost-shifting analysis must be neutral; close calls should be resolved in favor of the presumption. The *Rowe* factors, as applied, undercut that presumption for three reasons. *First*, the *Rowe* test is incomplete. *Second*, courts have given equal weight to all of the factors, when certain factors should predominate. *Third*, courts applying the *Rowe* test have not always developed a full factual record.

[*321] 1. The *Rowe* Test Is Incomplete

a. A Modification of *Rowe*: Additional Factors

Certain factors specifically identified in the Rules are omitted from *Rowe*'s eight factors. In particular, *Rule 26* requires consideration of "the amount in controversy, the parties' resources, the importance of the issues at stake in the litigation, and the importance of the proposed discovery in resolving the issues."⁶⁴ Yet *Rowe* makes no mention of either [**38] the amount in controversy or the importance of the issues at stake in the litigation. These factors should be added. Doing so would balance the *Rowe* factor that typically weighs most heavily in favor of cost-shifting, "the total cost associated with production." The cost of production is almost always an objectively large number in cases where litigating cost-shifting is worthwhile. But the cost of production when compared to "the amount in controversy" may tell a different story. A response to a discovery request costing \$ 100,000 sounds (and is) costly, but in a case potentially worth millions of dollars, the cost of responding may not be unduly burdensome.⁶⁵

64 *Fed. R. Civ. P. 26(b)(2)(iii)*.

65 A word of caution, however: in evaluating this factor courts must look beyond the (often inflated) value stated in the *ad damnum* clause of the complaint.

Rowe also contemplates "the resources available to each party." But here too -- although this consideration may be implicit in the *Rowe* test [**39] -- the absolute wealth of the parties is not the relevant factor. More important than comparing the relative ability of a party to pay for discovery, the focus should be on the total cost of production as compared to the resources available to each party. Thus, discovery that would be too expensive for one defendant to bear would be a drop in the bucket for another.⁶⁶

66 UBS, for example, reported net profits after tax of 942 million Swiss Francs (approximately \$ 716 million) for the third quarter of 2002 alone. See 11/12/02 UBS Press Release, available at http://www.ubswarburg.com/e/port_genint/index_genint.html.

Last, "the importance of the issues at stake in the litigation" is a critical consideration, even if it is one that will rarely be invoked. For example, if a case has the potential for broad public impact, then public policy weighs heavily in favor of permitting extensive discovery. Cases of this ilk might include toxic tort class actions, environmental actions, so-called "impact" or social [**40] reform litigation, cases involving criminal conduct, or cases implicating important legal or constitutional questions.

b. A Modification of *Rowe*: Eliminating Two Factors

Two of the *Rowe* factors should be eliminated:

First, the *Rowe* test includes "the specificity of the discovery request." Specificity is surely the touchstone of any good discovery request,⁶⁷ requiring a party to frame a request broadly enough to obtain relevant evidence, yet narrowly enough to control costs. But relevance and cost are already two of the *Rowe* factors (the second and sixth). Because the first and second factors are duplicative, they can be combined. Thus, the first factor should be: the extent to which the request is specifically tailored to discover relevant information.

67 See *Sedona Principles* (Principle 4: "Discovery requests should make as clear as possible what electronic documents and data are being asked for, while responses and objections to discovery should disclose the scope and limits of what is being produced.").

[**41] *Second*, the fourth factor, "the purposes for which the responding party maintains the requested data" is typically unimportant. Whether the data is kept for a business purpose or for disaster recovery does not affect its *accessibility*, which is the practical basis for calculating the cost of production.⁶⁸ Although a business purpose will often coincide with accessibility -- data that is inaccessible [**322] is unlikely to be used or needed in the ordinary course of business -- the concepts are not coterminous. In particular, a good deal of accessible data may be retained, though not in the ordinary course of business. For example, data that should rightly have been erased pursuant to a document retention/destruction policy may be inadvertently retained. If so, the fact that it should have been erased in no way shields that data from discovery. As long as the data is accessible, it must be produced.

68 Indeed, although Judge Francis weighed the purpose for which data is retained, his analysis also focused on accessibility:

If a party maintains electronic data for the purpose of utilizing it in connection with *current activities*, it may be expected to respond to discovery requests at its own expense. . . . Conversely, however, a party that happens to retain *vestigal data* for no current business purpose, but only in case of an emergency or simply because it has neglected to discard it, should not be put to the expense of producing it.

205 *F.R.D. at 431* (emphasis added). It is certainly true that data kept solely for disaster recovery is often relatively inaccessible because it is stored on backup tapes. But it is important not to conflate the purpose of retention with accessibility. A good deal of accessible, easily produced material may be kept for no apparent business purpose. Such evidence is no less discoverable than paper documents that serve no current purpose and exist only because a party failed to discard them. See, e.g., *Fidelity Nat. Title Ins. Co. of New York v. Intercounty Nat. Title Ins. Co.*, 2002 U.S. Dist. LEXIS 11916, No. 00 C. 5658, 2002 WL 1433584, at *6 (N.D. Ill. July 2,

2002) (requiring production of documents kept for no purpose, maintained "chaotically" and "cluttered in unorganized stacks" in an off-site warehouse); *Dangler v. New York City Off Track Betting Corp.*, 2000 U.S. Dist. LEXIS 14938, No. 95 Civ. 8495, 2000 WL 1510090, at *1 (S.D.N.Y. Oct. 11, 2000) (requiring production of documents kept "disorganized" in "dozens of boxes").

[**42] Of course, there will be certain limited instances where the very purpose of maintaining the data will be to produce it to the opposing party. That would be the case, for example, where the SEC requested "communications sent by [a] broker or dealer (including inter-office memoranda and communications) relating to his business as such." Such communications must be maintained pursuant to SEC *Rule 17a-4*.⁶⁹ But in such cases, cost-shifting would not be applicable in the first place; the relevant statute or rule would dictate the extent of discovery and the associated costs.⁷⁰ Cost-shifting would also be inappropriate for another reason -- namely, that the regulation itself requires that the data be kept "in an accessible place."

⁶⁹ See *supra*, note 20.

⁷⁰ However, while Zubulake is not the stated beneficiary of SEC *Rule 17a-4*, see *Touche Ross & Co. v. Redington*, 442 U.S. 560, 569-70, 61 L. Ed. 2d 82, 99 S. Ct. 2479 (1979), to the extent that the e-mails are accessible because of it, it inures to her benefit.

[**43] c. A New Seven-Factor Test

Set forth below is a new seven-factor test based on the modifications to *Rowe* discussed in the preceding sections.

1. The extent to which the request is specifically tailored to discover relevant information;

2. The availability of such information from other sources;

3. The total cost of production, compared to the amount in controversy;

4. The total cost of production, compared to the resources available to

each party;

5. The relative ability of each party to control costs and its incentive to do so;

6. The importance of the issues at stake in the litigation; and

7. The relative benefits to the parties of obtaining the information.

2. The Seven Factors Should Not Be Weighted Equally

Whenever a court applies a multi-factor test, there is a temptation to treat the factors as a check-list, resolving the issue in favor of whichever column has the most checks.⁷¹ But "we do not just add up the factors."⁷² When evaluating cost-shifting, the central question must be, does the request impose an "undue burden or expense" on the responding party?⁷³ Put another way, "how important is the sought-after [**44] evidence [**323] in comparison to the cost of production?" The seven-factor test articulated above provide some guidance in answering this question, but the test cannot be mechanically applied at the risk of losing sight of its purpose.

⁷¹ See, e.g., *Big O Tires, Inc. v. Bigfoot 4X4*, , 167 F. Supp. 2d 1216, 1227 (D. Colo. 2001) ("A majority of factors in the likelihood of confusion test weigh in favor of Big O. I therefore conclude that Big O has shown a likelihood of success on the merits.").

⁷² *Noble v. United States*, 231 F.3d 352, 359 (7th Cir. 2000).

⁷³ *Fed. R. Civ. P. 26(b)(iii)*.

Weighting the factors in descending order of importance may solve the problem and avoid a mechanistic application of the test. The first two factors -- comprising the marginal utility test -- are the most important. These factors include: (1) The extent to which the request is specifically tailored to discover relevant information and (2) the availability of such information from other [**45] sources. The substance of the marginal utility test was well described in *McPeck v. Ashcroft*:

The more likely it is that the backup tape contains information that is relevant to a

claim or defense, the fairer it is that the [responding party] search at its own expense. The less likely it is, the more unjust it would be to make the [responding party] search at its own expense. The difference is "at the margin."⁷⁴

74 202 F.R.D. at 34.

The second group of factors addresses cost issues: "How expensive will this production be?" and, "Who can handle that expense?" These factors include: (3) the total cost of production compared to the amount in controversy, (4) the total cost of production compared to the resources available to each party and (5) the relative ability of each party to control costs and its incentive to do so. The third "group" -- (6) the importance of the litigation itself -- stands alone, and as noted earlier will only rarely come into play. But where it does, this [**46] factor has the potential to predominate over the others. Collectively, the first three groups correspond to the three explicit considerations of *Rule 26(b)(2)(iii)*. Finally, the last factor -- (7) the relative benefits of production as between the requesting and producing parties -- is the least important because it is fair to presume that the response to a discovery request generally benefits the requesting party. But in the unusual case where production will also provide a tangible or strategic benefit to the responding party, that fact may weigh *against* shifting costs.

D. A Factual Basis Is Required to Support the Analysis

Courts applying *Rowe* have uniformly favored cost-shifting largely because of assumptions made concerning the likelihood that relevant information will be found. This is illustrated in *Rowe* itself:

Here, there is a high enough probability that a broad search of the defendants' e-mails will elicit some relevant information that the search should not be precluded altogether. However, there has certainly been no showing that the e-mails are likely to be a gold mine. No witness has testified, for example, about any e-mail communications [**47] that allegedly reflect discriminatory or

anti-competitive practices. Thus, the marginal value of searching the e-mails is modest at best, and this factor, too, militates in favor of imposing the costs of discovery on the plaintiffs.⁷⁵

But such proof will rarely exist in advance of obtaining the requested discovery. The suggestion that a plaintiff must not only demonstrate that probative evidence exists, but also prove that electronic discovery will yield a "gold mine," is contrary to the plain language of *Rule 26(b)(1)*, which permits discovery of "any matter" that is "relevant to [a] claim or defense."

75 205 F.R.D. at 430. See also *Murphy Oil*, 2002 U.S. Dist. LEXIS 3196, 2002 WL 246439, [WL] at *5 (determining that "the marginal value of searching the e-mail is modest at best" and weighs in favor of cost-shifting because "Murphy has not pointed to any evidence that shows that 'the e-mails are likely to be a gold mine'").

The best solution to this problem is found in *McPeck*:

Given the complicated [**48] questions presented [and] the clash of policies . . . I have decided to take small steps and perform, as it were, a test run. Accordingly, I will order DOJ to perform a backup restoration of the e-mails attributable to Diegelman's computer during the period of July 1, 1998 to July 1, 1999. . . . The DOJ will have to carefully document the time and money [**324] spent in doing the search. It will then have to search in the restored e-mails for any document responsive to any of the plaintiff's requests for production of documents. Upon the completion of this search, the DOJ will then file a comprehensive, sworn certification of the time and money spent and the results of the search. Once it does, I will permit the parties an opportunity to argue why the results and the expense do or do not justify any further search.⁷⁶

Requiring the responding party to restore and produce responsive documents from a small sample of backup tapes will inform the cost-shifting analysis laid out above.

When based on an actual sample, the marginal utility test will not be an exercise in speculation -- there will be tangible evidence of what the backup tapes may have to offer. There will also be tangible [**49] evidence of the time and cost required to restore the backup tapes, which in turn will inform the second group of cost-shifting factors. Thus, by requiring a sample restoration of backup tapes, the entire cost-shifting analysis can be grounded in fact rather than guesswork.⁷⁷

⁷⁶ 202 F.R.D. at 34-35.

⁷⁷ Of course, where the cost of a sample restoration is significant compared to the value of the suit, or where the suit itself is patently frivolous, even this minor effort *may* be inappropriate.

IV. CONCLUSION AND ORDER

In summary, deciding disputes regarding the scope and cost of discovery of electronic data requires a three-step analysis:

First, it is necessary to thoroughly understand the responding party's computer system, both with respect to active and stored data. For data that is kept in an accessible format, the usual rules of discovery apply: the responding party should pay the costs of producing responsive data. A court should consider cost-shifting *only* when [**50] electronic data is relatively inaccessible, such as in backup tapes.

Second, because the cost-shifting analysis is so fact-intensive, it is necessary to determine what data may be found on the inaccessible media. Requiring the responding party to restore and produce responsive documents from a small sample of the requested backup tapes is a sensible approach in most cases.

Third, and finally, in conducting the cost-shifting analysis, the following factors should be considered, weighted more-or-less in the following order:

1. The extent to which the request is specifically tailored to discover relevant

information;

2. The availability of such information from other sources;

3. The total cost of production, compared to the amount in controversy;

4. The total cost of production, compared to the resources available to each party;

5. The relative ability of each party to control costs and its incentive to do so;

6. The importance of the issues at stake in the litigation; and

7. The relative benefits to the parties of obtaining the information.

Accordingly, UBS is ordered to produce *all* responsive e-mails that exist on its optical disks [**51] or on its active servers (*i.e.*, in HP OpenMail files) at its own expense. UBS is also ordered to produce, at its expense, responsive e-mails from any *five* backup tapes selected *by Zubulake*. UBS should then prepare an affidavit detailing the results of its search, as well as the time and money spent. After reviewing the contents of the backup tapes and UBS's certification, the Court will conduct the appropriate cost-shifting analysis.

A conference is scheduled in Courtroom 12C at 4:30 p.m. on June 17, 2003.

SO ORDERED:

Shira A. Scheindlin

U.S.D.J.

Dated: New York, New York

May 13, 2003