Electronic Discovery in Surrogate's Court Litigation

Part I: An Introduction to Electronic Discovery Concepts

By Angelo M. Grasso

Long the bane of commercial litigators, electronic discovery, or e-discovery, has slowly but surely become a part of Surrogate's Court litigation. With its own language and issues that do not exist in conventional paper discovery, many practitioners have adopted the ostrich approach and tried to ignore the subject, or thrown up their arms in disgust after hearing a jargon-filled speech on technology and cost.

This is a mistake. Failing to understand and address e-discovery can lead to avoidable difficulties and unnecessary discovery disputes that only prolong litigation and add unnecessary cost. A solid command of the concepts behind e-discovery will aid not only the Surrogate's Court litigator, but also the planner and administrator who suspects she might need to produce documents in the future concerning her interactions with a testator or fiduciary.

This article, the first of two parts, is intended for Surrogate's Court practitioners who have had limited exposure to e-discovery, and provides an overview of some of the key terminology and processes that encompass e-discovery. The article will also discuss the preservation, collection and review of e-discovery, litigation holds and predictive coding, and the pitfalls for attorneys and clients who fail to comply with e-discovery rules and conventions.

Electronically Stored Information

The crux of e-discovery is the management of Electronically Stored Information (ESI). Rule 34 of the Federal Rules of Civil Procedure defines ESI as:

> Any designated documents or electronically stored information—including writings, drawings, graphs, charts, photographs, sound recordings, images, and other data or data compilations—stored in any medium from which information can be obtained either directly or, if necessary, after translation by the responding party into a reasonably usable form.[1]

This definition is both illuminating and vague, as it gives almost no specificity, perhaps in an effort to encompass technology that could not be anticipated, at the expense of precision. The comments to the 2006 Amendments to Rule 34 admit this, stating the rule "is intended to be broad enough to cover all current types of computer-based information and flexible enough to encompass future changes and developments."[2]

Fortunately, the Commercial Division of the Supreme Court of Nassau County has provided practitioners with more guidance. Long viewed as the New York State courts' leader in electronic discovery, the Court's rules[3] contain a list of what constitutes ESI:

- Native files
- Network access information
- Metadata
- Hard drives
- Internet usage files
- Offline storage
- Transaction logs
- Backup materials
- Spreadsheets
- Text files
- Emails
- Graphics
- Attachments
- Audio/visual files
- Voicemails
- Databases
- Instant messages
- Calendars
- Word processing documents
- Telephone logs
- Information stored on laptops, removable media, or "other portable devices"

While most of these terms are familiar, a few are worth exploring in detail. "Native Files" means ESI in the electronic format in which it was created, viewed, and/or modified. For example, an actual word processing file (with a .docx extension) would be a native file, while a printout of the document or a PDF of the document would not be. The same is true for spreadsheets: the .xls file is a native file; the hard copy spreadsheet is not. While those types of files are relatively easy to find and produce, much trickier are files that are not readily readable, such as databases.

"Static images" are representations of ESI made by converting a native file into a standard image format that can be viewed and printed. In other words, it is taking the file or data and putting it in a format that you can see, akin to preparing a trial exhibit. This would usually include PDFs of files, as PDF is not a format in which documents are ordinarily created. Similarly, it includes printouts of e-mails, rather than an entire Outlook data file.

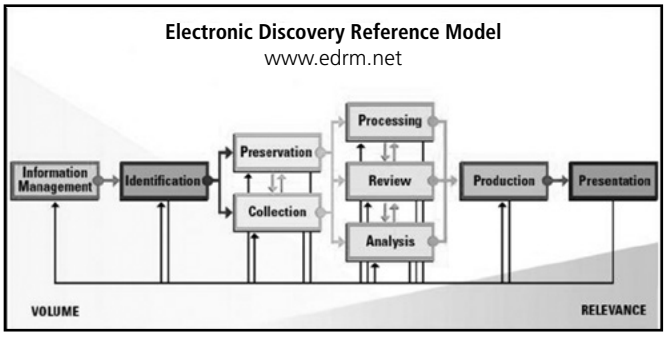
"Metadata" is information embedded in a Native File that is ordinarily neither viewable nor printable, but is generated when a file is created, modified, deleted, sent, received and/or manipulated. For example, a file's name, type, size, and location are all metadata. So

Angelo M. Grasso is a partner at Greenfield Stein & Senior, LLP in Manhattan and White Plains, who focuses on trusts and estates litigation. Angelo thanks his partner Gary B. Freidman and associate Tzipora Zelmanowitz for their assistance with this article.

are the dates the file was created and modified, and in each instance, who made the changes. A file's metadata is analogous to an internal diary that tracks and reads all modifications made to a file.

The Electronic Discovery Reference Model

The most common scheme under which e-discovery proceeds is the electronic discovery reference model (EDRM), which is illustrated by the flowchart below. The EDRM is an attempt to break down the e-discovery process from start to finish into nine stages. Much like the "inverted pyramid" used for depositions, the goal is to go from general to specific, starting with the most documents and ending up with the fewest, while relevancy increases.



The EDRM has nine stages:

- **Information management**: The client's internal procedures for maintaining ESI.
- **Identification**: Locating potential sources of ESI.
- **Preservation**: Preserving ESI.
- **Collection**: Gathering ESI for further use.
- **Processing**: Reducing and converting ESI to useable formats.
- **Review**: Evaluating ESI for relevance and privilege.
- **Analysis**: Evaluating ESI for content.
- **Production**: Delivering ESI to other parties in the correct form.
- **Presentation**: Using ESI, such as at depositions, trial, and hearings.

The EDRM is a broad template; frequently, entire stages will be skipped. As the chart indicates, some of these stages will occur concurrently. While all nine stages have their relative importance, the most common subject of discovery disputes concern the preservation, collection, and processing of ESI.

The Litigation Hold: *Zubulake v. UBS Warburg*

The most important e-discovery concept for a practitioner to understand are your and your client's duties and obligations concerning the collection and preservation of ESI. This concept is encompassed in the litigation hold, which was developed in the case *Zubulake v. UBS Warburg, LLC*.[4] *Zubulake* began as a conventional gender discrimination litigation; five decisions[5] and a jury trial later, it stands as the case outlining the contours of electronic discovery; it was the precursor to the sweeping 2006 amendments to the Federal Rules of Civil Procedure that govern e-discovery.

Ms. Zubulake was an equities trader at UBS who was allegedly denied a promotion. She brought an action in the Southern District of New York for gender discrimination, failure to promote, and retaliation.[6] Her case was assigned to Judge Shira Sheindlin, who had long expressed an interest in electronic discovery, and who used the case to write extensively on the subject.

From the beginning, Ms. Zubulake alleged that the evidence to prove her case was in emails on UBS' servers. Accordingly, she demanded production of e-mails sent among her former colleagues and superiors.[7] In response, UBS produced only 100 emails—a rather low number, considering that Ms. Zubulake had printed out and retained over 400 relevant emails before she left UBS.[8]

Judge Sheindlin first considered Ms. Zubulake's motion to compel, where she argued that the relevant e-mails were on UBS' backup media and should be produced.[9] After finding that the e-mails in question were relevant, the court turned to the issue of cost-shifting, as UBS claimed that it would cost over $300,000 to produce the requested e-mails because they were stored on backup tapes, and a timely and costly process was required to convert the data to readable emails.[10] The court determined that UBS was obligated to produce all emails that were readily accessible—*i.e.*, on optical disks or active servers—plus those from any five backup tapes that Ms. Zubulake selected.[11] In addition, UBS was directed to submit an affidavit outlining the results of its search and the time and cost expended so the court could make a proper cost-shifting analysis.[12]

In a subsequent opinion, the court addressed the issue of who would pay the $273,000 cost of restoring, searching and producing the emails from the 77 backup tapes.[13] Ultimately, the court assigned 25% of the cost of restoration to Ms. Zubulake, but held that UBS was to bear entire cost of producing the documents, noting that it was UBS' prerogative to have a senior associate at a large law firm to conduct the e-mail review, which Ms. Zubulake had no obligation to pay.[14]

By *Zubulake IV* it was clear that many e-mails were missing and could not be recovered.[15] Plaintiff moved for various sanctions, including an adverse inference

instruction against UBS with respect to the missing e-mails.[16] Noting that a spoliation sanction can only be levied if UBS destroyed evidence it had a duty to preserve, the court addressed three key questions: (i) When did UBS' duty to preserve arise?; (ii) What is the scope of UBS' duty to preserve?; and (iii) If UBS failed to preserve ESI, what is the appropriate remedy?

As to when the duty to preserve arose, the court held that the duty attached when litigation was reasonably anticipated.[17] UBS argued this occurred when Ms. Zubulake filed her human resources complaint; the court disagreed, holding the duty to preserve attached when the "key players" at UBS believed that litigation was possible.[18] About four months before Ms. Zubulake filed her human resources complaint, several key players began marking their internal e-mails as "privileged and confidential."[19] Additionally, a supervisor testified that the possibility of litigation "was in the back of his mind" as early as four months before she filed the complaint with human resources.[20] The court held UBS' duty to preserve arose at that point in time, as this was clear evidence that UBS knew that litigation was possible.[21]

As to the scope, the court stopped short of saying that once the duty to preserve arose, all ESI had to be preserved, noting this would paralyze a large institution like UBS, which is often involved in litigations.[22] Instead, the court held that it is the party's obligation to set aside the "unique" ESI:

> While a litigant is under no duty to keep or retain every document in its possession, it is under a duty to preserve what it knows, or reasonably should know is relevant in the action, is reasonably calculated to lead to the discovery of admissible evidence, is reasonably likely to be requested during discovery, and/or is the subject of a pending discovery request.[23]

To this end, the court created the now-famous "litigation hold":

> Once a party reasonably anticipates litigation, it must suspend its routine document retention/destruction policy and put in place a "litigation hold" to ensure the preservation of relevant documents.[24]

The court focused on the ESI from the people it dubbed the "key players," and held that the litigation hold must encompass all documents from key players in existence at the time litigation was reasonably anticipated or created thereafter.[25]

After concluding that UBS failed to preserve e-mails, the court turned to ascertaining the appropriate remedy. It began by noting its reluctance to impose an adverse inference sanction, as it is often too difficult to overcome.[26] The court instead employed a three-part test, holding that the party seeking an adverse inference instruction must show that:

1. A duty to preserve existed when the materials were destroyed;
2. The materials were destroyed with a "culpable state of mind"; and
3. The destroyed evidence was "relevant" to the party's claim or defense such that a reasonable trier of fact could find it would support the claim or defense.[27]

The second prong is the fulcrum in determining the severity of the sanction, as the court held that if the destruction of evidence was done in bad faith—*i.e.*, willfully or intentionally—this demonstrates the evidence destroyed was relevant, satisfying the third prong.[28] Hence, the key question becomes whether the destruction of ESI was merely negligent, or whether the offending party engaged in a willful pattern of destruction.[29] Here, the court held that the destruction of the e-mails appeared to have been mere negligence, not willful, and declined to apply an adverse inference instruction, but did order UBS to sit for additional depositions (at its own expense) concerning the missing e-mails.[30]

Sanctions: *Zubulake V* and *Montreal Pension*

The issue of sanctions rose to the forefront in *Zubulake V*. After the *Zubulake IV* depositions were completed, it was clear that many critical e-mails were deleted, never produced, and lost forever. The court held that both UBS and its counsel failed to hold and produce relevant evidence, which prejudiced Ms. Zubulake, noting the interplay between the client and counsel:

> A lawyer cannot be obliged to monitor her client like a parent watching a child. At some point, the client must bear responsibility for a failure to preserve. At the same time, counsel is more conscious of the contours of the preservation obligation; a party cannot reasonably be trusted to receive the "litigation hold" instruction once and to fully comply with it without the active supervision of counsel.[31]

The court suggested three steps that counsel "should take to ensure compliance with the preservation obligation": issuing a timely litigation hold, communicate directly with the "key players," and instruct all employees to produce electronic copies of the relevant, active files.[32] Practitioners who find themselves in actual or potential litigation where ESI is a factor are

well-advised to take these three steps and "get ahead" of e-discovery before running the risk of sanctions.

The court held that "the duty to preserve and produce documents rests on the party," and since UBS had continued to delete relevant e-mails beyond when the duty to preserve attached, they willfully destroyed potentially relevant information, warranting an adverse inference instruction.[33] Not surprisingly, in large part because of the sanction, Ms. Zubulake prevailed at trial, and UBS was directed to pay her $29.3 million—a hefty price to pay for deleting e-mails.

Sanctions once again became a key issue before the same court when Judge Scheindlin revisited the state of e-discovery in *Univ. of Montreal Pension Plan v. Bank of America Securities, LLC*,[34] which she titled "*Zubulake* Revisited: Six Years Later." *Montreal Pension* was a fairly complicated case with 96 plaintiffs suing concerning the loss of over half a billion dollars from the liquidation of two funds in the British Virgin Islands. During discovery, the defendants claimed they found substantial gaps in some of the plaintiffs' document productions. Depositions were held, after which defendants moved for sanctions, alleging that 13 plaintiffs failed to preserve and produce documents and submitted false declarations concerning their collection and preservation efforts.[35]

In *Montreal Pension*, the court found that plaintiffs "failed to timely institute written litigation holds and engaged in careless and indifferent collection efforts after their duty to preserve arose."[36] This led to the court to consider two key issues:

1. **Level of Negligence**. The court noted that ordinary negligence is a "failure to conform" to the standards "set by years of judicial decisions analyzing allegations of misconduct and reaching a determination as to what a party must do to meet its obligation to participate meaningfully and fairly in the discovery phase of a judicial proceeding."[37] (Artfully, Judge Scheindlin noted that someone is negligent even if the conduct "results from a pure heart and an empty head.")[38] By contrast, gross negligence is failing "to exercise even that care which a careless person would use." Willful negligence goes a step further, and is "an act of unreasonable character in disregard of a known of obvious risk that was so great as to make it highly probable that harm would follow, and which thus is usually accompanied by a conscious indifference to the consequences."[39]

Helpfully, the court provided examples of each form of negligence. Actively destroying documents is willful negligence.[40] Failing to issue a written litigation hold is gross negligence, as it is "likely to result in destruction of relevant information."[41] Failing to obtain ESI from all employees, take all measures on ESI, or use proper search terms is ordinary negligence.[42] The tricky line is the destruction of backup tapes or the failure to collect ESI from key players: this could be either willful or gross, depending on the circumstances.[43]

2. **Burden of Proof and Sanctions**. As every litigator knows, a case's success can turn on who bears the burden of proof. Here, the court creates multiple standards for the burden of proof on sanctions motions, holding that it "differs depending on the severity of the sanction."[44] For less severe sanctions (such as money), the court will focus on the conduct of the spoliator and whether there was any relevance to the documents that were destroyed.[45] However, if the party is seeking a severe sanction such as dismissal or striking a pleading, then the innocent party has the burden of showing that the spoliator had control of the evidence, acted with a culpable state of mind, and that the missing evidence was relevant to the claim or defense.[46]

As to the relevance prong, the court looked back to *Zubulake* and noted that relevance *and* prejudice may be presumed if the spoliator acted in bad faith or was grossly negligent.[47] If it was just "regular" negligence, then the innocent party must go the extra mile and show via extrinsic evidence that the destroyed evidence would have been *favorable* (not just relevant) to the claim or defense. Moreover, even if the negligence were gross, the presumption can be rebutted by showing that there was no prejudice.[48] The court's goal is clear:

> The party seeking relief has some obligation to make a showing of relevance and eventually prejudice, lest litigation become a "gotcha" game rather than a full and fair opportunity to air the merits of a dispute.[49]

Ultimately, the court noted its reluctance to hand out sanctions, because it "divert[s] court time from other important duties—namely, deciding cases on the merits."[50] Indeed, the *Montreal Pension* court estimated that it spent 300 hours on this motion alone.[51] Hence, the court noted that the goal is to reach a balance between keeping parties in line but keeping sanctions applications from becoming common, which it concluded, "is not a good thing."[52]

Predictive Coding: *Da Silva Moore* and *Delaney*

Preserving and collecting ESI is only one step in the e-discovery process. Equally critical is reviewing the ESI for responsiveness and producing it. No attorney has ever reveled in document review, but e-discovery adds the complicating factor of volume. Clients will often provide their attorneys with millions of ESI documents, which must then be sifted for responsiveness and privilege. Even putting aside the attendant boredom and indifference to reviewing over a million documents manually, doing so is woefully

ineffective. A 1985 study showed that attorneys wildly overestimate their ability to find responsive documents on manual review.[53] Later studies have shown that the level of agreement for manual review is approximately 70-75%, disproving the notion that human review of documents is the "gold standard."[54]

One of the most innovative developments in e-discovery is predictive coding, the most common method of automating ESI review. To make predictive coding work, the attorneys will review and code an initial group of documents (the "seed set") to "train" the computer by telling the computer which of the documents in the seed set is and is not responsive. The computer "learns" from the seed set, and applies this "knowledge" to the remaining documents to determine what is and is not relevant. Once this has been done, the attorneys will manually review sample responsive and non-responsive results to determine whether the computer review reached a predetermined "confidence level." If it has not, then the seed set and algorithm will be refined to produce a response with an increased confidence level.

Predictive coding sounds a bit like hocus-pocus and a recipe for mistakes. However, it is less foreign than it sounds: anyone who uses e-mail has inadvertently bumped up against predictive coding, which is how your spam filter works. While there will be a larger upfront expense to code and teach the computer, on a large enough set of documents, it is almost certainly more cost-effective than having attorneys and paralegals bill hourly at an inferior success rate.

Predictive coding first gained acceptance in *Da Silva Moore v. Publicis Groupe*,[55] where Magistrate Judge Andrew Peck encouraged the parties to use predictive coding. *Da Silva Moore* contains a fairly extensive review of the process used to try to narrow the universe of responsive documents from three million. Judge Peck noted that predictive coding has two enormous benefits:

1. It greatly reduces the amount of manual review that needs to be done by attorneys, as "technology-assisted review requires, on average, human review of only 1.9% of the documents."[56]

2. Keyword searches, the default for document review, are of limited utility because they are often over-inclusive and yield too many non-responsive documents. Equally problematic, "the way lawyers choose keywords is the equivalent of the child's game Go Fish," as the requesting party selects words "without having much, if any, knowledge of the responding party's 'cards.'"[57]

Ultimately, the court determined that predictive coding was the proper approach for sifting through the vast quantity of documents, holding:

> 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. Courts and litigants must be cognizant of the aim of Rule 1, to "secure the just, speedy, and inexpensive determination" of lawsuits.[58]

However, a key to predictive coding is transparency and cooperation between the parties. The pitfalls of conducting predictive coding without transparency was made apparent in the Nevada case *Progressive Casualty Ins. Co. v. Delaney*.[59] In a declaratory judgment action concerning failed banks in multiple jurisdictions, the parties submitted a Joint ESI Protocol that the court approved and so-ordered. In the first search, Progressive collected approximately 1.8 million ESI documents. Using search terms set forth in the Joint ESI Protocol reduced it to "merely" 565,000 documents, which, as Thomas Jefferson said, "is too many damn pages for any man to understand."[60] Progressive's attorneys attempted a manual review, but determined after 125,000 documents that it was simply too voluminous.

Progressive then decided to unilaterally ignore the Joint ESI Protocol and turn to predictive coding, which narrowed the field from 565,000 to 90,575 "potentially relevant" documents.[61] Progressive further noted that adding a "privilege" filter identified approximately 27,000 documents as "more likely privileged," and proposed manually reviewing these documents while producing the remaining 63,000 documents without manual review, subject to a clawback agreement.[62]

The issue raised in opposition was the lack of transparency behind Progressive's predictive coding. While the search terms that reduced the universe of documents from 1.8 million to 565,000 was established in the Joint ESI Protocol, the defendants had no way of knowing what method was used to "seed" and "teach" the predictive coding system, nor would plaintiffs give this information, claiming it was discovery about discovery.[63] The court took a dim view of this reluctance, noting that "courts which have allowed predictive coding...have emphasized the need for cooperation and transparency in adopting predictive coding processes and methods."[64] The court explained transparency was necessary because predictive coding was vulnerable to the "garbage in, garbage out" phenomenon:

> Predictive coding, or technology assisted review, uses software that can

> be trained by a human being to distinguish between relevant and non-relevant documents. However, the quality of its product depends on the quality of the information used to "train" the software.[65]

As Progressive's lack of transparency failed to "comply with [its own expert's] recommended best practices," the court determined that allowing secretive predictive coding "will only result in more disputes."[66] Because Progressive ignored the discovery Order, the court directed them to run a filter for privilege, but otherwise, to produce all of the 565,000 emails on the basis that transparency trumps work product.[67] While this appears to have been a victory for the defendants, it is arguable that they were punished, as it now became their burden to review over half a million ESI documents.

Conclusion

E-discovery presents issues that require careful attention by all practitioners to the management, collection and dissemination of their ESI. A future article will discuss how the Surrogate's Courts have treated ESI and e-discovery, and explore the contours of the practitioner's obligations to preserve and produce ESI.

Endnotes

1. Fed. R. Civ. P. 34(a)(1)(A).
2. *Id.*
3. https://www.nycourts.gov/courts/comdiv/PDFs/Nassau-E-Filing_Guidelines.pdf.
4. 217 F.R.D. 309 (S.D.N.Y. 2003).
5. Specifically: 217 F.R.D. 309 (S.D.N.Y. 2003) ("*Zubulake I*"); 230 F.R.D. 290 (S.D.N.Y. 2003) ("*Zubulake II*"); 216 F.R.D. 280 (S.D.N.Y. 2003) ("*Zubulake III*"); 220 F.R.D. 212 (S.D.N.Y. 2003) ("*Zubulake IV*"); 229 F.R.D. 422 (S.D.N.Y. 2004) ("*Zubulake V*").
6. *Zululake I,* 217 F.R.D. at 312.
7. *Id.* at 312-13.
8. *Id.* at 313.
9. *Id.*
10. *Id.*
11. *Zubulake I,* 217 F.R.D. at 324.
12. *Id.*
13. *Zubulake III,* 216 F.R.D. at 283.
14. *Id.* at 290.
15. *Zubulake IV,* 220 F.R.D. at 215.
16. *Id.*
17. *Zubulake IV,* 220 F.R.D. at 217.
18. *Id.*
19. *Id.*
20. *Id.*
21. *Id.*
22. *Id.*
23. *Id.*
24. *Zubulake IV,* 220 F.R.D. at 218.
25. *Id.*
26. *Zubulake IV,* 220 F.R.D. at 219-220.
27. *Id.* at 220.
28. *Id.* at 221.
29. *Id.* at 220.
30. *Id.* at 222.
31. *Zubulake V*, 229 F.R.D at 433.
32. *Id.* at 433-34.
33. *Id.* at 436.
34. 685 F. Supp. 2d 456 (S.D.N.Y. 2010).
35. *Montreal Pension,* 685 F. Supp. 2d 462-63.
36. *Id.* at 463.
37. *Id.* at 464.
38. *Id.*
39. *Id.*
40. *Montreal Pension,* 685 F. Supp. 2d at 465.
41. *Id.*
42. *Id.*
43. *Id.*
44. *Montreal Pension,* 685 F. Supp. 2d at 467.
45. *Id.*
46. *Id.*
47. *Id.*
48. 685 F. Supp. 2d at 468-69.
49. *Id.* at 468.
50. *Id*. at 471.
51. *Id.*
52. The least interesting part of *Montreal Pension* is how it was resolved, as the court meted out particularized penalties and judgments for each of the thirteen plaintiffs over the course of sixteen pages, which mattered to the litigants, but had little precedential value.
53. David E. Blair & M.E. Maron. *An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System*, Communications of the ACM, Vol. 28, Issue 3 (1985).
54. Maura R. Grossman & Gordon V. Cormack, *Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient Than Exhaustive Manual Review*, XVII Rich. J.L. Tech. 11, (2011); 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 (2010).
55. 287 F.R.D. 182 (S.D.N.Y. 2012).
56. *Da Silva Moore,* 287 F.R.D. at 190.
57. *Id.* at 191.
58. *Id*. at 191 (internal citations omitted).
59. 2014 WL 3563467 (D. Nevada 2014).
60. Lin-Manuel Miranda, *Hamilton*, "Cabinet Battle #1."
61. *Progressive* at *2.
62. *Id.* at *3.
63. *Id.* at *9.
64. *Id.* at *4.
65. *Id.* at *8.
66. *Id.* at *11.
67. *Id*.