

MAY 1, 2012

## The BUZZ on Predictive Coding — Global Aerospace Inc., et al. v. Lindow Aviation LP, et al.<sup>1</sup>

On April 23, 2012, Judge James H. Chamblin issued an order in *Global Aerospace Inc., et al. v. Lindow Aviation LP, et al.* approving “the use of predictive coding for purposes of the processing and production of electronically stored information.”<sup>2</sup> The Order expressly reserves the right of any receiving party to object to the “completeness of the contents of the production or the ongoing use of predictive coding.”<sup>3</sup>

The two Lindow defendants moved for a protective order because plaintiffs were objecting, without explanation, to the defendants’ proposed use of predictive coding in discovery.<sup>4</sup> The Memorandum in Support requested the court approve the use of predictive coding to avoid the undue burden and expense associated with alternative means of culling the ESI. The defendants alternatively requested that the court order a shifting of the incremental costs to plaintiffs because the plaintiffs objected to the cost-saving measures afforded by predictive coding.

In their Memorandum in Support, the defendants asserted that the use of predictive coding is a reasonable means of locating and retrieving documents that may be responsive to requests for production. They argued that, as a result, predictive coding satisfies the Civil Rules regarding discovery. In showing its reasonableness, the defendants compared computerized predictive coding to human linear review and keyword searching. They maintained that a traditional first-pass review of their data would yield only 60% of the potentially relevant documents, while keyword searching would yield only 20%; these two methods would be costly and time consuming. Predictive coding, on the other hand, would locate “upwards of seventy-five percent of the potentially relevant documents and can be implemented at a fraction of the cost and in a fraction of the time of linear review and keyword searching.”<sup>5</sup>

The defendants’ discovery plan is to use OraTec’s predictive coding software to retrieve potentially relevant documents from a “massive collection of electronically stored information” (~250 GB) that would then be reviewed and produced if responsive and not privilege. The defendants will include a

---

<sup>1</sup> Global Aerospace Inc., et al. v. Lindow Aviation LP /dba/ Dulles Jet Center, et al., Consolidated Case No. CL 00061040-00, (Circuit Court Loudoun Cty., Apr. 23, 2012) (J. Chamblin).

<sup>2</sup> Order Approving the Use of Predictive Coding for Discovery (“Order”).

<sup>3</sup> Id.

<sup>4</sup> Memorandum in Support of Motion for Protective Order Approving the Use of Predictive Coding (“Memorandum in Support”).

<sup>5</sup> Memorandum in Support, p. 2.



“statistically sound validation protocol” that should discharge the “reasonable inquiry” obligations under the court’s discovery rules.<sup>6</sup>

Citing an article by Maura Grossman and Gordon Cormack evaluating technology-assisted review,<sup>7</sup> defendants argued that linear human review is neither consistent nor particularly effective at identifying and segregating relevant documents from non-relevant ones. The defendants included in their argument a discussion of two studies that evaluated the consistency among human reviewers and revealed that agreement among lawyers on relevancy occurred between 28 and 50% of the time.<sup>8</sup> The defendants state that, according to Grossman and Cormack’s article, linear review misses on average 40% of the relevant documents and pulls in irrelevant documents close to 70% of the time.<sup>9</sup>

When discussing keyword searching, defendants call on the study by Blair & Maron<sup>10</sup> that revealed that keyword searching missed 80% of the relevant documents and pulled in irrelevant documents only 21% of the time; however, that number has been as high as 93%.<sup>11</sup> The defendants’ Cleveland-based vendor, JurInnov Ltd., evaluated the data set against a set of keywords proposed by the plaintiffs and found that a significant number of irrelevant documents would be included in the set and it would be difficult to determine why certain keywords were contained in either the relevant or irrelevant sets such that no modification of the searching could be tailored to produce better results.

The defendants explained how predictive coding works. While different technologies exist, they asserted that “predictive coding will be more effective at retrieving relevant documents and eliminating irrelevant documents than either linear review or keyword searching.” The defendants outlined their protocol for using predictive coding, which included providing the plaintiffs with the full set of training documents minus privileged and irrelevant sensitive documents that will be sufficiently logged.

Defendants intend to provide opposing counsel the opportunity to seek a modification of the coding on the training material if objections to the coding arise. Defendants also intend to implement a statistically valid sampling program to establish whether an acceptable level of recall (75%, according to defendants) has been achieved by the completed predictive coding.

---

<sup>6</sup> Id. pp. 1-2.

<sup>7</sup> M. R. Grossman & G. V. Cormack, Technology-Assisted Review in E-Discovery Can Be More Effective and More Efficient than Exhaustive Manual Review, 42 Rich. J.L. & Tech. 11 (2011).

<sup>8</sup> The two studies are: Ellen M. Voorhees, Variations in Relevance Judgments and the Measurement of Retrieval Effectiveness, 36 Info. Processing & Mgmt. 697 (2000), available at [http://www.cs.cornell.edu/courses/cs430/2006fa/cache/Trec\\_8.pdf](http://www.cs.cornell.edu/courses/cs430/2006fa/cache/Trec_8.pdf), and Herbert L. Roitblat et al., Document Categorization in Legal Electronic Discovery: Computer Classification vs. Manual Review, 61 J. Am. Soc’y for Info. Sci. and Tech. 70 (2010), available at <http://www.clearwellsystems.com/e-discovery-blog/wp-content/uploads/2010/12/man-v-comp-doc-review.pdf>.

<sup>9</sup> Memorandum in Support, p. 7.

<sup>10</sup> D. C. Blair & M. E. Maron, An Evaluation of Retrieval Effectiveness for a Full Text Document-Retrieval System, 28 Commc’ns ACM 289 (1985).

<sup>11</sup> Defendants also find support in and cite to The Sedona Conference Best Practices Commentary on the Use of Search & Information Retrieval Methods in E-Discovery, 8 The Sedona Conf. Journal. Fall 2007, at p. 189.



Defendants ultimately argued that predictive coding “does not truly present a question of the application or extension of legal precedent,” because “[t]here is no judicial or other legal mandate requiring, or even advocating, the use of one method of document retrieval over another.” Notably, the defendants cite to Judge Peck’s *Da Silva Moore* decision in support of their position.<sup>12</sup>

Plaintiffs maintain that document production must involve these steps: 1) talk to the client; 2) locate files that might reasonably contain responsive documents; 3) look at the documents; 4) select the ones that are responsive and not privileged; and 5) produce them.<sup>13</sup> Plaintiffs complain that the predictive coding measures intended by defendants would only yield 75% of the relevant documents, not **all** responsive documents as obligated under the Civil Rules. Plaintiffs characterize predictive coding as a “radical departure from the standard practice of human review of documents as a necessary step in responding to a request for production.”<sup>14</sup>

Perhaps important to the Judge’s decision to permit predictive coding as planned by the defendants was this statement made by the plaintiffs:

This is not to say that the Landow Defendants cannot use computer technologies to make their review process more efficient. To ensure that relevant evidence is not left behind, Defendants may (and indeed should) use keyword searching to spot-check the files deemed nonresponsive. ... These tools can and should be used as well to ensure a complete response.

In making this statement in their Opposition, the plaintiffs, perhaps inadvertently, concede to the defendants’ proposal to use predictive coding to bring forth all (or at least 75%) of the potentially relevant documents so that human reviewers can then evaluate them for responsiveness and privilege. Judge Chamblin issued his Order over the plaintiffs’ objections.

---

<sup>12</sup> *Da Silva Moore v. Publicis Groupe*, Case No. 11- cv-01279, 2012 LEXIS 23350 (S. D. NY Feb. 24, 2012). Magistrate Judge Peck’s controversial order was upheld by Judge Andrew Carter in an opinion issued on April 26, 2012, stating “there is insufficient evidence to conclude that the use of the predictive coding software will deny Plaintiffs access to liberal discovery ... [and] is difficult to ascertain that the predictive software is less reliable than the traditional keyword search.... To call the method unreliable at this stage is speculative.” *Da Silva Moore*, Case No. 11- cv-01279 (Doc. 175, Apr. 26, 2012), at p. 4.

<sup>13</sup> Opposition of Plaintiffs M.I.C. Industries, Inc., Factory Mutual Insurance Co., Global Aerospace, Inc., and BAE systems Survivability Systems, LLC to the Landow Defendants’ Motion for Protective Order Regarding Electronic Documents and “Predictive Coding” (“Opposition”), p. 1.

<sup>14</sup> *Id.* at 2.



---

*This newsletter is an information source for clients and friends of Redgrave LLP. The content should not be construed as legal advice, and readers should not act upon information in this publication without professional counsel. This material may be considered advertising under certain rules of Professional Conduct. ©2012 Redgrave LLP. All Rights Reserved.*

*Contact Us: For further information or if you have any questions regarding this Alert, please contact your Redgrave LLP attorney or Managing Partner Victoria Redgrave at (202) 681-2599 or [vredgrave@redgravellp.com](mailto:vredgrave@redgravellp.com).*

