

IN THE UNITED STATES DISTRICT COURT  
FOR THE EASTERN DISTRICT OF TEXAS  
LUFKIN DIVISION

THE PROCTOR & GAMBLE CO.,	§	
	§	
<i>Plaintiff,</i>	§	Civil Action No. 9:08-CV-143
	§	
v.	§	JUDGE RON CLARK
	§	
S.C. JOHNSON & SON, INC.,	§	
	§	
<i>Defendant.</i>	§	

**ORDER**

At the January 23, 2009 Case Management Conference, the court ordered the parties to submit estimates of the cost involved in the production of documents in searchable Tagged Image File Format (“TIFF”) with Optical Character Recognition (“OCR”) before it made a ruling on the format of electronic discovery. The parties have now agreed that all documents are to be produced electronically in TIFF format; the only remaining issue is whether Defendant S.C. Johnson & Son, Inc.<sup>1</sup> should have to perform the OCR process on these documents prior to production. Defendant estimates that the total cost of the OCR is likely to exceed \$200,000, while Plaintiff The Proctor & Gamble Co. argues that this estimate is unsupported and is more likely around three cents per page for those documents maintained in hard copy form.

In OCR, static images of text are translated into a format, via a computer software program, that can be searched or read electronically. *See In re Classicstar Mare Lease Litigation*, 2009 WL 260954 at \*1, n.2 (E.D. Ky. Feb. 2, 2009); *In re Payment Card Interchange*

---

<sup>1</sup>Defendant states that it does not itself seek to use the OCR process, and any extra expense would be incurred on its behalf solely for Plaintiff’s convenience.

*Fee & Merchant Discount*, 2007 WL 121426 at \*1, n.2 (E.D.N.Y. Jan. 12, 2007). OCR does not have to be performed on files created electronically, like word processing documents, e-mail, or spreadsheets. Rather, it is used to render documents maintained in hard copy format and scanned into a computer searchable.

Defendant argues that the cost of converting documents to OCR “is likely to exceed \$200,000,” but fails to provide the court with a per page cost estimate or even the total number of pages on which it expects to perform the process. All Defendant supplies is an estimate of \$1,900 per custodian for document conversion, and does not even identify the total number of custodians from whom it anticipates documents will be obtained. Even assuming Plaintiff’s estimate of three cents per page is low<sup>2</sup>, the court finds it difficult to believe that in the age of electronically stored information so much of the requested data is only maintained in hard copy format which would require OCR to be searchable. In short, Defendant has failed to provide any support whatsoever for its estimate.

Additionally, Defendant does not contend that performing the OCR process would not make documents easier to examine and evaluate, thus reducing costs for attorney time. Rather, it simply asserts that the cost burden should be shifted. In a slightly different context, other courts have adopted multi-factor tests when considering whether cost shifting is appropriate. *See Zubulake v. UBS Warburg LLC*, 217 F.R.D. 309 (S.D.N.Y. 2003); *Rowe Entertainment, Inc. v.*

---

<sup>2</sup>Plaintiff does not provide a source for this estimate. However, a simple Internet search reveals multiple software programs available for OCR conversion at costs ranging from about \$200 to \$600. While this option may not be appropriate for some or all Defendant’s document custodians, and does not take into account the cost of the time needed to convert the documents, it does tend to make the \$200,000 figure provided by Defendant appear somewhat inflated.

*William Morris Agency, Inc.*, 205 F.R.D. 421 (S.D.N.Y. 2002), *aff'd* 2002 WL 975713 (S.D.N.Y. May 9, 2002).

In *Zubulake*, the court held that cost shifting could be considered when the data requested was relatively inaccessible and formulated a seven factor test: (1) the extent to which the request was specifically tailored to discover relevant information; (2) the availability of such information from other sources; (3) the total cost of production, when compared to the amount in controversy; (4) the total cost of production, when compared to the resources available to each party; (5) the relative ability of each party to control costs and the 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. *Zubulake*, 217 F.R.D. at 322. These seven factors should not be weighted equally; rather, they should be considered in numerical order of descending importance, where factors one and two are the most important and factor seven is the least important. *Id.* at 323.

Analysis of these seven factors does not favor cost-shifting. First, Defendant argues in passing that about \$137,000 of the estimated \$200,000 in OCR costs can be attributed to Plaintiff's request for documents and information relating to fabric freshening and/or malodor reducing products other than the accused Glade Fabric & Air Odor Eliminating product. However, other than a statement that this request is "overbroad," Defendant does not contend that the information requested is not relevant, nor unlikely to lead to the discovery of admissible information. Discovery pursuant to the Federal Rules is a "broad. . . regime," *O2 Micro Int'l v. Monolithic Power Sys.*, 467 F.3d 1355, 1366 (5th Cir. 2006), and the Local Rules for the Eastern District of Texas reflect this. Discovery need not even be admissible at trial to be obtained under

the Federal Rules, since its purpose is to “remove surprise from trial preparation so the parties can obtain evidence necessary to evaluate and resolve their dispute.” *Stone v. Unocal Termination Allowance Plan*, 2007 U.S. Dist. LEXIS 33082 at \*4 (S.D. Tex. May 7, 2007) (citation omitted). Absent a showing, or even a claim, that the evidence sought is neither relevant nor reasonably likely to lead to the discovery of admissible information, this factor does not favor Defendant’s position.

The remaining factors also do not favor cost-shifting. There has been no showing that the documents requested are obtainable from other sources. As discussed with the court at the Case Management Conference, the parties’ respective litigation budgets in this case are estimated to be several million dollars apiece, so the parties have adequate resources. Sales of the accused product have been between one and two million dollars per month. Defendant has budgeted a great deal of money in this case, in defense of a market which, according to Plaintiff, is worth many millions of dollars per year.

OCR, while perhaps not absolutely necessary to litigation, is a tool that greatly decreases the time and effort counsel must invest in searching and examining documents. Presumably, each party would perform the OCR process in a cost-effective manner to minimize their costs. Requiring the parties to incur this cost, when the OCR process is likely to streamline the discovery process and reduce the chance that either side will employ tactics designed to hide relevant information in a mountain of difficult-to-search documents is neither unreasonable nor burdensome. Therefore, after review of the *Zubulake* factors, the court concludes that cost shifting is not appropriate in this case.

IT IS THEREFORE ORDERED that documents will be produced in searchable TIFF. The parties will perform the OCR process on any scanned documents stored in hard copy format prior to production.

So **ORDERED** and **SIGNED** this **19** day of **February, 2009**.

A handwritten signature in black ink, appearing to read "Ron Clark". The signature is written in a cursive style with a large initial "R".

---

Ron Clark, United States District Judge