

EXHIBIT 17

September 15, 2015

VIA ELECTRONIC MAIL

Michael Sobol, Esq.
David Rudolph, Esq.
Melissa Gardner, Esq.
Lieff Cabraser Heimann & Bernstein, LLP
275 Battery Street, 29th Floor
San Francisco, CA 94111-3339

Hank Bates, Esq.
Allen Carney, Esq.
David Slade, Esq.
Carney Bates & Pulliam, PLLC
11311 Arcade Drive
Little Rock, AR 72212

Re: *Campbell et al. v. Facebook, Inc.*, N.D. Cal. Case No. 13-cv-05996-PJH

Dear David:

I write in response to your further follow-up letter dated September 1, 2015, regarding predictive coding.

As an initial matter, we would remind you that we affirmatively and proactively engaged you to discuss this widely-accepted process in an effort to expedite Facebook's document collection and review. In fact, before the predictive coding process even began, in the interest of reviewing and producing responsive materials to you in a timely manner, we manually reviewed more than a thousand documents and produced the documents that are the most relevant to your case.

Thereafter, and as explained in our initial letter discussing predictive coding (dated June 19, 2015), it became clear that the initial set of materials collected and processed constituted an extremely large volume of material – over two million documents. Thus, it was obvious that additional tools were necessary to identify responsive documents in a timely and cost-efficient manner, and predictive coding is “widely accepted for limiting e-discovery to relevant documents and effecting discovery of ESI without an undue burden.” *Dynamo Holdings Ltd. P'ship v. C.I.R.*, 143 T.C. 183, 192 (2014); *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 127 (S.D.N.Y. 2015) (“[C]ase law has developed to the point that it is now black letter law that where the producing party wants to utilize TAR [technology assisted review] for document review, courts will permit it.”). For that reason, we reached out to you and your colleagues to engage in a constructive dialogue regarding these methods and your thoughts on a fair, reasonable, and proportionate review process. In the meantime, we undertook a carefully crafted, systematic effort to develop a dynamic predictive algorithm based on nearly 5,000 training- and assessment-set documents that we manually reviewed and classified in order to make the model as accurate and robust as possible. Further, each time that we identified additional potential sources of relevant information (either internal repositories or individual email archives), we searched for and collected all relevant information from those sources.

Since our initial June 19th letter, and since our in-person discussion on June 30, Plaintiffs have sent multiple letters critiquing this process and asking numerous follow-up questions, but not offering a single constructive suggestion about how to complete or improve this process. At this

point, it is apparent that Plaintiffs' strategy is to attempt to manufacture a dispute and/or reserve their right to object to the process after the fact, even though we have attempted to engage you from the beginning. This lack of cooperation has forced us to work through the process ourselves, prioritizing the production of responsive materials while following best practices.

Turning to the further questions in your follow-up letter:

First, your assertion that Plaintiffs did not agree to the use of keyword searching is demonstrably false. Facebook has involved Plaintiffs in its efforts to assure a reasonable and proportionate discovery process, as contemplated by this Court's ESI Guidelines 1.02 and 1.03. The parties expressly agreed to a set of search terms in May, and since that time Plaintiffs have not suggested (or attempted to justify) any additional terms. As discussed above, these terms proved to be extremely broad, and yielded a significant number of non-responsive documents. The custodians likely to have discoverable and responsive information have sent or received millions of emails during the agreed upon time period for discovery, the overwhelming majority of which have nothing to do with the challenged practice. Although the agreed search terms narrowed this universe, the process still yielded an overbroad number of documents (around 800,000) for the agreed time period from the agreed list of custodians. When Facebook proposed a predictive coding process to narrow the universe, it was Plaintiffs who were hesitant to deviate from the more conventional process of using the agreed search terms and conducting a manual review, as you indicated during our in-person discussions on June 30. Accordingly, it is inaccurate to suggest at this late date that Plaintiffs never agreed to search terms.

Facebook's use of broad search terms to identify a universe of potentially relevant documents, and use of predictive coding to identify documents within that universe that are most likely to be relevant, is an effective way to reduce costs, increase efficiency, and expedite the process of identifying the relevant information, as contemplated by this Court's ESI Guideline 2.02(f).¹ Plaintiffs have not proposed any alternative process for identifying relevant material without undue burden, cost, and delay. To the extent that Plaintiffs' concerns about search term filtering are based on the ever-present risk of excluding a number of responsive documents, there are standard ways to assess such risks, including through sampling methods designed to validate the search (an exercise that Facebook has already advised you it intends to employ). Indeed, the United States District Court for the Northern District of California contemplates the use of such sampling for that very purpose in ESI Guideline 2.02(f). Plaintiffs cannot sit back, agree to the process, and then at the very last minute attempt to object to Facebook's months-long document collection and production process.

Moreover, no search terms were applied in collecting and processing the emails from the two key custodians (Alex Himel and Ray He) for the critical time period of October 2012, when the decision was made to end the challenged process. To be clear, *all* October 2012 emails for

¹ As for one of your other questions, as Facebook has explained on multiple previous occasions, the agreed search terms were used to cull the document population from the email repositories of the agreed custodians within the agreed date range. Facebook is unable to provide a count of total emails possessed by custodians, and instead only provides a count of exported files (i.e., search results exported).

custodians Ray He and Alex Himel were subjected to predictive coding and subsequent manual classification.

Second, our August 20, 2015 correspondence explained the nature of the documents included in the training (which you refer to as “seeding”) sets used to develop the predictive coding model. By way of further explanation, the training set used in this case is comprised of two types of document sets: (1) documents used in previous assessments, which were randomly selected from and representative of the document population that existed at the time and which Gibson Dunn attorneys had manually classified as either responsive or non-responsive;² and (2) documents that were confirmed to be responsive and have been produced in this case. These two document sets together provide an effective sample of the documents from the population to train the model.

Regardless of the makeup of the training set, however, the results of the final assessment analysis demonstrate the performance of the model on the overall document population. *The most critical focus of the modeling process has been to develop a model that can achieve an acceptable recall rate.*

Because the final assessment is a random and representative sample of the overall population for review, applying the model to the overall review population should achieve the indicated recall rate from our final assessment. Our August 20 letter provided the information for this final assessment (or “control set”), which included the model cutoff score, the number of responsive documents found, the number of non-responsive documents found, the overturn rate, and the true positive and false positive counts above and below the cutoff score. Further, as explained in our earlier letters, we also plan to make use of statistical sampling to ensure the robustness of our review based on predictive coding. *See, e.g., Rio Tinto PLC*, 306 F.R.D. at 128 (“[R]equesting parties can insure that training and review was done appropriately by other means, such as statistical estimation of recall at the conclusion of the review as well as by whether there are gaps in the production, and quality control review of samples from the documents categorized as non-responsive.”).

Third, those documents confirmed to be responsive have been (and will continue to be) produced, but Plaintiffs are not entitled to non-responsive documents. You assert that Plaintiffs cannot “judge the effectiveness of Facebook’s TAR implementation” without that information. That is incorrect. While Plaintiffs cannot critique Gibson Dunn’s review, they can indeed assess the “TAR” implementation, which is merely a tool to consistently and efficiently apply the same criteria in Gibson Dunn’s review to the document population. Likewise, Plaintiffs have no good faith basis on which to demand production of the “randomly selected control set documents” and, with respect to each document in the set, the responsiveness classifications made by the reviewing attorneys and the Equivio software. If Facebook were conducting a purely “traditional” manual review and production of documents without the use of predictive coding, its attorneys’ responsiveness classifications would not be subject to second-guessing by Plaintiffs

² These assessments were generated both while performing iterations of the modeling process, as described in greater detail in our previous letter, and as needed to account for newly processed data (such as when new custodians emails were collected). These assessments include responsive and not responsive documents.

through the production of non-responsive documents. Likewise, there is no justification for Plaintiffs to receive productions of non-responsive documents to evaluate predictive coding. Courts have rejected similar attempts to subject predictive coding to a different (or higher) standard than manual review. *See, e.g., Rio Tinto PLC*, 306 F.R.D. at 129 (“One point must be stressed—it is inappropriate to hold TAR to a higher standard than keywords or manual review. Doing so discourages parties from using TAR for fear of spending more in motion practice than the savings from using TAR for review.”). The validity of the predictive coding process is measured by the recall rate. Plaintiffs are not entitled to non-responsive material.

Fourth, your claim that “Facebook’s production thus far—a significant portion of which are either publicly-available or highly duplicative (i.e., individual responses to email chains spread across many documents)—appears to be inadequate” is belied by even a cursory review of the multiple productions to date. Facebook has provided thousands of pages of confidential email exchanges discussing aspects of and decisions concerning the functionality at issue in this case. Facebook has also produced internal presentations, Wiki pages, Tasks, Differentials (or “Diffs” – documents showing changes in the source code), Salesforce documents related to Facebook advertising practices, documents from the Facebook Help Center, Statements of Rights and Responsibilities and Data Use Policies that discuss the relevant functionality during the relevant time period, and other materials.

Finally, as Facebook has reminded Plaintiffs on multiple occasions (and Plaintiffs have acknowledged on multiple occasions), this case is fundamentally about the functionality underlying the Facebook Messages product. Accordingly, the critical facts relating to Plaintiffs’ claims were contained in the technical documentation disclosed to Plaintiffs on June 1, 2015, and further verified through the source code repositories to which Plaintiffs have now had access for nearly two months. Accordingly, Rule 26’s proportionality requirement for discovery is especially instructive here: “the court must limit the frequency or extent of discovery otherwise allowed ... if it determines that ... *the burden or expense of the proposed discovery outweighs its likely benefit, considering the needs of the case, the amount in controversy, the parties’ resources, the importance of the issues at stake in the action, and the importance of the discovery in resolving the issues.*” Fed. R. Civ. P. 26(b)(2)(C) (emphasis added).

In sum, Facebook has now explained—on three separate occasions—the process by which it is identifying the documents most likely to be relevant from an enormous volume of materials in the most expedient and efficient manner possible. If Plaintiffs have a specific, concrete proposal regarding how to conduct a proportional and time-effective review, Facebook is willing to consider it, subject to any necessary cost-sharing to offset the considerable expense of reworking processes implemented months ago. *See, e.g., ESI Guideline 2.02(f)* (“Opportunities to reduce costs and increase efficiency and speed, such as by conferring about the methods and technology used for searching ESI to help identify the relevant information ... or by sharing expenses like those related to litigation document repositories.”). But Facebook will not continue to defend against an avalanche of open-ended inquiries and demands concerning its use of respected and highly defensible methods to produce responsive materials in this case, especially when it is

September 15, 2015

Page 5

apparent that Plaintiffs are attempting to manufacture a dispute as a means of increasing the cost and duration of this meritless litigation.

Sincerely,

/s/ Priyanka Rajagopalan
Priyanka Rajagopalan

cc: All counsel of record