UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

WILMER GARCIA RAMIREZ, SULMA HERNANDEZ ALFARO, ANA P., on behalf of themselves and others similarly situated,	)	
Plaintiffs,	)	Case No. 1:18-cv-00508
v.	)	Class Action
U.S. IMMIGRATION AND CUSTOMS ENFORCEMENT ("ICE"), et al.;	)	
Defendants.	)	

DEFENDANTS' PROPOSAL CONCERNING PRODUCTION OF ELECTRONICALLY STORED INFORMATION AND RESPONSE TO PLAINTIFFS' PROPOSAL

INTRODUCTION

The Court should reject Plaintiffs' proposal concerning production of electronically stored information ("ESI"), ECF No. 85, as both needlessly delaying production and unnecessarily embroiling the Court and Plaintiffs in Defendants' review process. Defendants, in contrast, propose a straightforward extension of five months, up to and including June 17, 2019, to complete production of ESI in this litigation. During the discovery period, Defendants propose rolling productions of responsive, non-privileged identified during the course of reviewing 5,000 documents per week[1] (with the exception of the weeks of December 24, 2018-January 11, 2019). Defendants' proposal results in orderly, ongoing, and regular provision of responsive ESI to Plaintiffs while protecting Defendants' ability to conduct a reasonable review

[1] To clarify, Defendants will review 5,000 documents per week and produce responsive, non-privileged documents identified on a weekly basis. While Defendants can control how many documents that are reviewed each week, Defendants are unable to predict how many of the documents will be suitable for production.

to protect applicable privileges of Defendants law enforcement documents. The parties have discharged their obligation to confer regarding ESI pursuant to Federal Rule of Civil Procedure 26(f)(3)(C) and this Court's Orders. Consequently, Plaintiffs' request for additional, gratuitous involvement in Defendants' review and production procedures is not warranted, not required by any authority, not consistent with best practices for the conduct of e-discovery, and is not unrealistic in the volume it contemplates.

I. DEFENDANTS' REQUEST FOR A FIVE-MONTH EXTENSION OF THE DISCOVERY PERIOD TO COMPLETE PRODUCTION OF ESI

On December 3, 2018, the parties finalized their agreed upon search terms for collection from the eighteen baseline custodians. *See* Defendants' Protocol for Use of Active Learning to Assist with Responsiveness Review ("Exhibit A"); *see also* Search Terms Report ("STR") of December 6, 2018 ("Exhibit B").[2] As of the time of this filing, Defendants are in the process of training Relativity Assisted Review Active Learning ("Active Learning"), an application that utilizes an iterative training regimen in which the training set is repeatedly augmented by additional documents chosen by the Machine Learning Algorithm and coded by subject matter experts for responsiveness. Exhibit A; *see also* *https://help.relativity.com/9.6/Content/Relativity/Active_Learning/Active_Learning.htm*. During the course of ICE's Active Learning responsiveness review of the 256,247 documents (including families) collected utilizing the parties' agreed upon search terms, Exhibit B, it is ICE's understanding that the current number of documents identified for privilege review is approximately 20,856 from the eighteen baseline custodians identified in Defendants' Active

[2] The STR of December 17, 2018, reflects an additional term eliminated by the parties and correction of a typographical error, as such, it is the report of the parties' finalized search terms. Exhibit B.

Learning Protocol. Exhibit A at Appendix A. This number may fluctuate as Defendants complete the Active Learning process. Defendants propose to produce responsive, non-privileged documents identified during the course of reviewing 5,000 documents per week for the eighteen baseline custodians during the five-month discovery period.

As noted in Defendants' protocol, not all documents collected from the eighteen baseline custodians utilizing the parties' agreed upon search terms can be submitted to Active Learning for responsiveness review. *See* Exhibit A at 3 section II.1. Certain file types such as media (.jpg, .jpeg, .bmp, .gif, .wmv, .wav, .mov, .avi), spreadsheets (.xls, .xlsx, .csv), system files, container files, and documents without text or documents with too much text and extensive markup language (such as .xml) are inappropriate for machine learning review using Active Learning. ICE proposes to manually review these excluded documents for responsiveness and privilege and produce responsive, non-privileged files during the five-month discovery period.

Given the current set of documents requiring review (including documents reviewed through Active Learning and excluded file types), ICE proposes to produce responsive, non-privileged documents identified during the course of reviewing 5,000 documents per week for a period of twenty (20) weeks, up to and including May 31, 2019. This period is adjusted for the weeks of December 24, 2018-January 11, 2019, when ICE anticipates limited review resources that will prevent productions. Additionally, Defendants' proposal provides for two extra weeks of discovery following Defendants' last rolling ESI production in order to accommodate any final depositions resulting from the last production. Accordingly, Defendants request a five-month extension of the discovery period up to and including June 17, 2019.

Given Defendants' reasonable, straightforward request for a five-month extension of the discovery period to produce ESI in this litigation, Defendants respectfully request the Court adopt Defendants' proposal.

II. DEFENDANTS MUST BE ALLOWED TO DETERMINE REVIEW AND PRODUCTION OF THEIR OWN ESI WITHOUT PREEMPTIVE RESTRAINT

Plaintiffs' proposal would needlessly delay discovery by entangling the Court in Defendants' review process merely to allay Plaintiffs' premature fears *before* rolling productions even begin and to allow Plaintiffs to circumvent their burden on a motion to compel under Federal Rule of Civil Procedure 37. The Court should reject this diversion. As many courts have noted, the producing party is in the best position to "'evaluate the procedures, methodologies, and technologies appropriate for preserving and producing their own electronically stored information.'" *Winfield v. City of New York*, 2017 WL 5664852, at *9 (S.D.N.Y. Nov. 27, 2017) (quoting *Hyles v. New York City*, 2016 WL 4077114, at *3 (S.D.N.Y. Aug. 1, 2016)); *see The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 Sedona Conf. J. 1 (2018) ("Sedona Conference") at Principle 6 ("Principle 6 recognizes that a responding party is best situated to preserve, search, and produce its own ESI . . . . without direction from the court or opposing counsel and eschewing 'discovery on discovery,' unless a specific deficiency is shown in a party's production.").[3]

[3] As Plaintiffs recognized during the December 13, 2018, conference with the Court, The Sedona Conference is the leading voice of the legal profession in addressing e-discovery concerns. The Sedona Conference publishes several frequently cited commentaries in this area, including the authoritative Sedona Principles, currently in its Third Edition and publically available at *https://thesedonaconference.org/publications*.

Traditionally, courts have not micro-managed parties' internal review processes for a number of reasons. First, attorneys, as officers of the court, are expected to comply with Rules 26 and 34 in connection with their search, collection, review and production of documents, including ESI. *Winfield*, 2017 WL 5664852, at *9. Second, internal attorney ESI work processes may reveal work product, litigation tactics, and trial strategy. *Id.* (citing generally *Disability Rights Council of Greater Wash. v. Wash. Metro. Transit Auth.*, 242 F.R.D. 139, 142-43 (D.D.C. 2007) (holding that a compilation of documents culled from a larger protection is protectable as attorney work product)); *see* Sedona Conference at Principle 6, Comment 6.c. ("Parties may voluntarily agree to produce or exchange documentation of their discovery processes, but should do so only after due consideration of privilege and work product issues." (citation omitted)).[4] Third, the producing party is better equipped than the court to identify and utilize the best process for producing their own ESI consistent with their obligations under the Federal Rules of Civil Procedure. *Hyles*, 2016 WL 4077114, at *3 (citing Principle 6 of the Sedona Conference). Fourth, perfection in ESI discovery is not required; rather, a producing party must take reasonable steps to identify and produce responsive documents. *Winfield*, 2017

[4] Materials may be subject to the work product immunity, even when created by others and not in anticipation of litigation, where (1) counsel has sifted through a large number of documents in order to locate and identify a smaller number of documents which are relevant to the case; and (2) although the documents themselves are not protected work product, requiring production of the culled documents would reveal to the other side the mental processes, impressions, and opinions of the attorneys who separated "the wheat from the chaff." *San Juan Dupont Plaza Hotel Fire Litigation*, 859 F.2d 1007, 1010 (1st Cir.1988); *accord Sporck v. Peil*, 759 F.2d 312, 316 (3d Cir.1985) (holding that the selection and compilation of documents by counsel in preparation for pretrial discovery may constitute work product). "The selection and compilation of otherwise unprivileged documents by counsel for litigation purposes should give rise to the work product immunity only where there is 'a real, rather than a speculative, concern' that the thought processes of the compiling lawyer will be exposed." *Wollam v. Wright Med. Grp., Inc.*, 2011 WL 4375016, at *1 (D. Colo. Sept. 20, 2011) (citing *Gould Inc. v. Mitsui Mining & Smelting Co. ., Ltd.*, 825 F.2d 676, 680 (2d Cir. 1987)).

WL 5664852, at *9 (citing *HM Elecs., Inc. v. R.F. Techs., Inc.*, 2015 WL 471498, at *12 (S.D. Cal. Aug. 7, 2015), *vacated in part on other grounds*, 171 F. Supp. 3d 1020 (S.D. Cal. 2016); *Pension Comm. of the Univ. of Montreal Pension Plan v. Banc of Am. Secs., LLC*, 685 F. Supp. 2d 456, 461 (S.D.N.Y. 2010). Finally, deference regarding the selection and implementation of the discovery process should be given to the party that will incur the costs. Sedona Conference at Principle 6, Comment 6.b. ("[the responding party] should be permitted to elect how best to allocate its resources and incur the costs required to comply with its obligations.") (citing *Kleen Prods. LLC et al., v. International Paper et al.*, Civ. No. 1:10-cv-05711, Doc. 319, Ex. A, at 6667-68 (N.D. Ill. May 17, 2012) (Tr. of Proceedings before the Hon. Nan Nolan)).

In keeping with these principles, the Court should reject Plaintiffs' attempt to inject themselves into Defendants' review process based on their opinion that Defendants' proposal for rolling productions of 5,000 documents per week is "unreasonably slow" and because they view Defendants' Active Learning protocol as "opaque." ECF No. 85 at 1. Plaintiffs, however, assert no prejudice which will result from Defendants' proposal. Nor can they, as they have asked the Court for additional depositions, which will require additional time. As a preliminary matter, Plaintiffs' assertion that it may take Defendants "as long as 51 weeks" to complete productions of ESI from the eighteen custodians is nonsensical. ECF No. 85 at 2. Plaintiffs' claim ignores Defendants' Active Learning Protocol (Exhibit A) and Draft Plan for Review of Electronically Stored Information ("Exhibit C"), stating responsiveness review would be conducted utilizing Active Learning by December 14, 2018, and only documents determined to be responsive (or excluded from Active Learning because of their file type) would undergo further review. Accordingly, Plaintiffs' estimate—inexplicably assuming all 256,247 documents hitting on search terms would be submitted to privilege review—is grossly over inclusive.

Regarding Plaintiffs' request that Defendants be required to share metrics and "meet and confer with Plaintiffs to reach a consensus on appropriate parameters" for use of Active Learning, the Court should reject this attempt to direct Defendants' review, particularly prior to rolling productions commencing and in the total absence of any discovery failure.[5] ECF No. 85 at 8. In *Winfield*, 2017 WL 5664852, a district court denied a similar request based on unspecified, generalized concerns regarding a party's technology assisted review. There, after productions of ESI, plaintiffs objected to the defendants' continued use of its Technology Assisted Review ("TAR") system. Plaintiffs contended that the system was improperly trained because the human document reviewers over-designated documents as non-responsive during both the linear review and during the TAR training stages. *Id.* at *9. As a result, plaintiffs claimed the TAR software was unable to recognize and properly categorize responsive documents. *Id.* The plaintiffs sought to check the defendants' document review by requesting an order providing information about the ranking system used (*i.e.*, what cut-off was used, and how many documents were deemed responsive and unresponsive at each ranking) and a random sample of various categories of documents that fell above and below defendants' selected cut-off rank. *Id.* at *6. In denying plaintiffs' request, the court cited case law endorsing the Sedona Conference's authoritative view captured in Principle 6 that "the producing party is in the best position to 'evaluate the procedure, methodologies, and technologies appropriate for preserving

[5] Although Plaintiffs inconsistently contend their request does not violate Sedona Conference Principle 6 because "Plaintiffs simply seek metrics to ensure the process is operating as it should." ECF No. 85 at 9-10. In the same filing, however, Plaintiffs specifically request that Defendants share these metrics and "meet and confer with Plaintiffs to reach a consensus on appropriate parameters" of these metrics for Defendants' use of Active Learning. Thus, Plaintiffs undeniably aim to direct the production of Defendants' ESI, in contravention of Principle 6. ECF No. 85 at 8.

and producing their own electronically stored information.'" *Id.* at *9 (citations omitted). The court explained:

> . . . there is nothing so exceptional about ESI production that should cause courts to insert themselves as super-managers of the parties' internal review processes, including training of TAR software, or to permit discovery about such process, in the absence of evidence of good cause such as a showing of gross negligence in the review and production process, the failure to produce relevant specific documents known to exist or that are likely to exist, or other malfeasance.

Id.

Here, the Court should deny Plaintiffs' proposal for the same reasons outlined in *Winfield*. Like the plaintiffs in *Winfield*, Plaintiffs' speculative concerns do not warrant intervention in Defendants' ESI review process. Indeed, Plaintiffs here have even less basis to request involvement in Defendants' review process because they have not yet received ESI productions subjected to any review. Plaintiffs' concerns are, thus, entirely premature. "A requesting party should not normally be able to restrain the responding party's discovery process to prevent an anticipated, but uncertain, future harm." Sedona Conference at Principle 6, Comment 6.b. (citing *In re Ford Motor Co.*, 345 F.3d 1315, 1317 (11th Cir. 2003) (vacating order for discovery of certain databases where no finding of "some non-compliance with discovery rules by Ford"); *Koninklijke Philips N.V. v. Hunt Control Sys., Inc.*, 2014 WL 1494517, at *4 (D.N.J. Apr. 16, 2014) (moving party failed to show a "material deficiency" in the responding party's electronic discovery process); *Freedman v. Weatherford Int'l*, 2014 WL 4547039 (S.D.N.Y. Sept. 12, 2014) (request for "discovery on discovery" denied for failure in absence of factual basis to find original production deficient)). Thus, it is patently insufficient to require Defendants to obtain Plaintiffs' approval of Defendants' review process, based completely on conjecture devoid of any factual basis.

Further, pursuant to Federal Rule of Civil Procedure 37(a), and as delineated by the Sedona Conference at Principle 7, "[t]he requesting party has the burden on a motion to compel to show that the responding party's steps to preserve and produce relevant electronically stored information were inadequate." Plaintiffs seek to avoid their burden by simply injecting themselves in Defendants' review process. According to Plaintiffs, metrics and inclusion in Defendants' review is necessary because they would otherwise not know if production is incomplete. ECF No. 85 at 9, n.4. Plaintiffs, however, cite no authority for the proposition that a responding party must grant access to their review process merely because the requesting party does not know how to determine when to file a motion to compel under Rule 37. Indeed, a court in this jurisdiction has held:

> Speculation that there is more will not suffice; if the theoretical possibility that more documents exist sufficed to justify additional discovery, discovery would never end. Instead of chasing the theoretical possibility that additional documents exist, courts have insisted that the documents that have been produced permit a reasonable deduction that other documents may exist or did exist and have been destroyed.

Hubbard v. Potter, 247 F.R.D. 27, 29 (D.D.C. 2008) (citations excluded). This rule is no less true for e-discovery and Plaintiffs may not circumvent their burden. *Id.* at 31 (declining to order discovery on discovery based on "nothing more than [plaintiffs'] own speculation that other electronic documents exist."). *See Harris v. Koenig*, 271 F.R.D. 356, 370 (D.D.C. 2010) ("If plaintiffs are speculating that documents responsive to these requests do exist, there must be a reasonable deduction that that is true, and not a mere hunch.") (citations omitted).

Moreover, as recognized by the Sedona Conference in Principle 6, the responding party must make numerous determinations to identify, preserve, collect, process, analyze, review, and produce relevant responsive and non-privileged and discoverable ESI for each case. Determining what is relevant and discoverable under the circumstances for each matter often

requires a highly fact-specific inquiry. "Thus, the responding party—not the court or requesting party—is both tasked with making those determinations and generally in a better position to make those decisions. Because of the dynamic nature of litigation, the analysis cannot be reduced to a generalized checklist of reasonable steps for every party to take in every action." Sedona Conference at Principle 6, Comment 6.a. Thus, there should be no preemptive restraint placed on a responding party that chooses to proceed on its own with determining how best to fulfill its preservation and discovery obligations. *Id.* Specifically, Sedona Conference, Principle 6 at Comment 6.a. explains:

> . . . neither a requesting party nor the court should prescribe or detail the steps that a responding party must take to meet its discovery obligations, and there should be no discovery on discovery, absent an agreement between the parties, or specific, tangible, evidence-based indicia (versus general allegations of deficiencies or mere "speculation") of a material failure by the responding party to meet its obligations.

Accordingly, adoption of Plaintiffs' proposal—requiring Defendants to obtain Plaintiffs' approval for responsiveness review criteria in the absence of any deficiency—functions as a preemptive restraint and is highly disfavored.

In support of their novel proposal, Plaintiffs cite two cases that are manifestly distinguishable. ECF No. 85 at 3-5. The first is *Progressive Casualty Insurance Co. v. Delaney*, 2014 WL 3563467 (D. Nev. July 18, 2014). In that case, "[t]he parties submitted a joint proposed ESI protocol," "and agreed to search terms to run across ESI which [the plaintiff] represented was in its possession . . . . " *Id.* at *6. After agreeing to that ESI protocol, the plaintiff unilaterally "began utilizing predictive coding techniques to review ESI without the defendant's agreement *to amend the parties' stipulated ESI protocol Order* . . . and without seeking leave of the court to amend the ESI Order." *Id.* at *2 (emphasis added). The predictive coding the plaintiff sought "would [have] relieve[d] it of the burden of manual review of ESI

according to the ESI protocol it [had originally] stipulated to . . . . " *Id.* at *10. Ultimately, the court did not allow the plaintiff to use predictive coding. *See id.* at *11. *Progressive* is distinguishable from this case because in *Progressive* the parties "agreed to search terms to run across ESI," which defendants then attempted to unilaterally modify. *Id.* at *6. Here, the parties never reached any agreement regarding review methodology. Accordingly, there is no breach of any such agreement requiring the Court's intervention. *Id.* *Progressive* is inapposite.

Stambler v. Amazon.com, Inc., 2011 WL 10538668 (E.D. Tex. May 23, 2011), is likewise distinguishable. In *Stambler*, defendants opposed production of documents after having agreed to search terms and custodians for e-mail production. *Id.* at *9. Specifically, two months before the close of discovery defendants asserted the agreed upon terms were overly burdensome and unilaterally determined to produce only a subset of the e-mails agreed upon. *Id.* As to the purported burden, plaintiff argued that quick peek agreements or claw back agreements could mollify the harm. *Id.* Hence, in order to uphold the parties' prior agreement, the court granted plaintiffs' motion to compel and ordered productions subject to the parties' quick peek and claw back agreements. *Id.* at *11. In so doing, the court specifically noted "[t]his case is unusual—*and therefore of limited precedential value*—because the parties reached an agreement but then [d]efendants argued pursuant to Rule 26(b)(2)(C)(iii) that the resulting review and production would be overly burdensome." *Id.* (emphasis added). *Stambler* is not instructive here. Not only does *Stambler* lack precedential value in the very court that issued it, but it is also plainly distinguishable. Here, the parties never came to any agreement regarding productions based on

key words, nor are Defendants refusing to make productions or unilaterally modifying any prior agreement with Plaintiffs. Accordingly, the relief ordered in *Stambler* is wholly irrelevant here.[6]

In their filing, Plaintiffs make much of the Sedona Conference Cooperation Proclamation (July 2008) in support of their proposition that Defendants are required to share metrics and seek Plaintiffs' approval concerning the application of Active Learning. ECF No. 85 at 7, n.3. Plaintiffs misread the Cooperation Proclamation. As an initial matter, the Cooperation Proclamation is a publication seeking to institute a paradigm shift in the parties' approach to e-discovery. *See The Sedona Conference*, *Cooperation Proclamation*, 10 SEDONA CONF. J. 331 (2009 Supp.). Consistent with the authoritative Sedona Principles, the Cooperation Proclamation is not independent authority, but is intended to track the Federal Rules of Civil Procedure and recommended best practices. Accordingly, the parties' obligations begin and end with the Federal Rules. Critically, Rule 1 requires parties to employ the Rules "to secure the just, speedy, and inexpensive determination of every action and proceeding." Additionally, requesting and responding parties share the mutual obligation to meet and confer in good faith to discuss the preservation and production of ESI, as required by Rule 26(f)(3)(C) directing the parties to confer on and prepare a discovery plan that addresses "any issues about disclosure, discovery, or preservation of electronically stored information, including the form or forms in which it should be produced." The Rules also provide that requests for ESI proceed under Rule 34 and action to compel production proceed under Rule 37(a). In addition to what is required by those Rules, Sedona Conference Principle 6 generally recommends the responding party engage in

[6] *Progressive* and *Stambler* are also distinguishable in that they did not involve discovery on a federal law enforcement agency as is the case here. Production of ICE documents necessitates a reasonable review to preserve applicable governmental privileges that were not at stake in *Progressive* or *Stambler*.

meaningful cooperation with opposing parties to attempt to reduce the costs and risk associated with the preservation and production of ESI. Thus, Comment 6.b. states that "[i]f both requesting and responding parties voluntarily elect to cooperatively evaluate and agree upon the appropriate procedures, methodologies, and technologies to be employed in the case, both may potentially achieve significant monetary savings and non-monetary efficiencies." However, consistent with the Federal Rules, the Sedona Conference emphasizes that such agreements are "voluntary." Sedona Conference at Principle 6, Comment 6.b. In the absence of any voluntary agreement, Principle 6 specifies that a responding party is best situated to preserve, search, and produce its own ESI "without direction from the court or opposing counsel and eschewing 'discovery on discovery,' unless a specific deficiency is shown in a party's production."[7]

Finally, the Court should deny Plaintiffs' proposal because reasonableness and proportionality, not perfection and scorched-earth, must be the guiding principles illuminating e-discovery. *See* Fed. R. Civ. Pro. 1 (emphasizing the parties' and courts' obligation to control costs); Fed. R. Civ. Pro. 26(b) (restoring proportionality to the definitional scope of discovery). Thus, the Court should reject Plaintiffs' attempt to hold Defendants' use of Active Learning to a higher standard than keywords or manual review. As noted by the court in *Rio Tinto PLC v. Vale S.A.*, 306 F.R.D. 125, 129 (S.D.N.Y. 2015), holding technology assisted review tools to a higher standard "discourages parties from using TAR for fear of spending more in motion practice than the savings from using TAR for review." Thus, because it is Defendants' prerogative to design a review process that leverages technological resources, and Plaintiffs

[7] Moreover, Defendants note they have cooperated with Plaintiffs to finalize the parties' agreed upon collection search terms, discuss appropriate custodians, and provide a detailed protocol outlining Defendants' application of Active Learning. *See* Exhibit A; B.

cannot prematurely establish any production failure will occur (unless the sheer volume that the Court orders proves impossible for ICE to meet), the Court should decline Plaintiffs' invitation to ignore the weight of authority and enmesh Plaintiffs in Defendants' review process.[8] *See* Declaration of Michael P. Davis, ICE Executive Deputy Principal Legal Advisor ("Exhibit D") (outlining ICE resource limitations).

III. PLAINTIFFS' PROPOSAL FUNDAMENTALLY MISUNDERSTANDS ACTIVE LEARNING

In addition to incorrectly assuming Active Learning will have no impact on the number of documents submitted for privilege review, Plaintiffs' proposal also contains material misunderstandings regarding Active Learning. First, Plaintiffs' filing misstates Defendants' definition of "elusion rate" defined in Defendants' protocol. *Compare* ECF No. 85 at 8; Exhibit A at I. Specifically, Defendants' Active Learning Protocol defines "elusion rate" as "[t]he percentage of documents coded as responsive after review of the Elusion Test sample." Exhibit A at I. Inexplicably, Plaintiffs' filing states the definition is "the percentage of documents coded as responsive after initial training of the computer." ECF No. 85 at 8. Whether or not inadvertent, Plaintiffs' definition misinterprets Defendants' protocol. Under Plaintiffs' rephrased definition, the elusion rate would be all documents coded responsive by the reviewer(s). The elusion rate, however, is the percentage of responsive documents in a sample of "non-coded, predicted non-relevant documents (not reviewed, skipped, suppressed duplicates)." *See*

[8] For the same reasons, the Court should deny Plaintiffs' request to require Defendants to provide them with periodic status reports concerning the progress of Defendants' review and production of ESI. ECF No. 85 at 10-11. If Plaintiffs elect to challenge Defendants' productions—once made—their recourse is under Federal Rule of Civil Procedure 37(a). The Court should not prematurely require Defendants to provide information to Plaintiffs on matters that are not the subject of any claim or defense relevant to the litigation as required for discovery under Rule 26.

https://help.relativity.com/9.6/Content/Relativity/Active_Learning/Elusion_Test.htm. at 20. Further, as Defendants have repeatedly explained to Plaintiffs, Active Learning is an iterative process and the elusion rate cannot be arbitrarily set before initialization. Notably, the Cormack-Grossman Glossary defines "elusion" as "[t]he fraction of documents identified as Non-Relevant by a search or review effort, that are in fact Relevant." Maura R. Grossman and Gordon V. Cormack, *The Grossman-Cormack Glossary of Technology-Assisted Review*, 2013 Fed. Cts. L. Rev. 7 (January 2013). Elusion is computed by taking a random sample from the null set and determining how many (or what proportion of) documents are actually relevant. *Id.* A low elusion value has commonly been advanced as evidence of an effective search or review effort. *Id.* Accordingly, in Relativity, the elusion rate cannot be set prior to the project start and Plaintiffs cannot expect this metric before training is complete.

Further, the Court should reject Plaintiffs' unfounded request that Defendants provide "back end" quality-control metrics used to evaluate a "technology-assisted review process," namely, recall and precision. ECF No. 85 at 8-9. Plaintiffs claim they require these metrics to determine "whether the computer did a good job." *Id.* at 8. Even assuming Plaintiffs may challenge Defendants' review procedure before any production is made, without specifying any failure, and outside of a Rule 37 motion to compel—which they may not—Relativity does not automatically generate recall and precision in Active Learning. Potentially, Plaintiffs are confusing continuous Active Learning with Sample-Based Learning. The central difference between Active Learning and Sample-Based Learning, however, is that Active Learning uses a type of technology called Support Vector Machine Learning (SVM). As the review progresses, the SVM algorithm continues to learn, taking advantage of the additional judgments made by the reviewers. Because Active Learning does not utilize a control set to train the index, recall and

precision are simply not features of a standard Active Learning project. *See* https://help.relativity.com/9.6/Content/Relativity/Active_Learning/Active_Learning.htm. Accordingly, Plaintiffs' contention that these metrics "can be generated with a keystroke" is simply wrong. Consequently, Plaintiffs' request for precision and recall metrics should be rejected. The Court should allow Defendants' election to review their own ESI utilizing Active Learning, not Sample-Based Learning, and reject Plaintiffs' impermissible preemptive restriction. *Winfield*, 2017 WL 5664852, at *9; *Hyles*, 2016 WL 4077114, at *3; Sedona Conference at Principle 6.

CONCLUSION

Because it is both unreasonable and untenable, the Court should deny Plaintiffs' Proposed Plan for Production of Electronically Stored Information (ECF No. 85). The Court should and grant Defendants' Proposed Plan for Production of Electronically Stored Information because it is both orderly and reasonable.

DATE: December 17, 2018

Respectfully submitted,

JOSEPH H. HUNT
Assistant Attorney General
Civil Division

WILLIAM C. PEACHEY
Director

COLIN A. KISOR
Deputy Director

CHRISTINA PARASCANDOLA
Senior Litigation Counsel

CARA E. ALSTERBERG
Trial Attorney

KEVIN HIRST
Trial Attorney

By: */s/ Yamileth G. Davila*
YAMILETH G. DAVILA
Assistant Director
U.S. Department of Justice, Civil Division
Office of Immigration Litigation –
District Court Section
P.O. Box 868, Washington, DC 20044
Telephone: 202-305-0137
Facsimile: 202-305-7000
Yamileth.g.davila@usdoj.gov

ATTORNEYS FOR DEFENDANTS

Exhibit A

UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

WILMER GARCIA RAMIREZ, *et al.*, Plaintiffs, v. UNITED STATES IMMIGRATION AND CUSTOMS ENFORCEMENT, *et al.*, Defendants.	Civil Action No. 1:18-CV-00508-RC Class Action

PROTOCOL FOR USE OF ACTIVE LEARNING TO ASSIST WITH RESPONSIVENESS REVIEW

Defendant United States Immigration and Customs Enforcement ("ICE") hereby proposes to use the following protocol for the use of Active Learning for purposes of assisting with the responsiveness review in this civil action.

I. DEFINITIONS

1. **Active Learning**: An Iterative Training regimen in which the Training Set is repeatedly augmented by additional documents chosen by the Machine Learning Algorithm, and coded by one or more Subject Matter Expert(s).

2. **Confidence Level**: In this protocol, Confidence Level is one of the variables used to create a statistically valid sample. The other variables are the size of the total document collection to be sampled from and the tolerable margin of error. Confidence Level is a measurement of statistical reliability, and, in this context, refers to the likelihood that a measurement reached through the sample is accurate.

3. **Cutoff Rank:** This value determines the minimum rank needed for a document to receive a responsive categorization. Documents ranking above the Cutoff Rank shall consist of documents most likely to be responsive. Documents ranking below the Cutoff Rank will be deemed non-responsive and the results will be validated by the Elusion Test.

4. **Deduplication**: A method of replacing multiple identical copies of a document by a single instance of that document. Deduplication can occur within the data of a single custodian (also referred to as Vertical Deduplication), or across all custodians (also referred to as Horizontal or Global Deduplication).

5. **Elusion Rate:** The percentage of documents coded as responsive after review of the Elusion Test sample.

6. **Elusion Test:** A process by which Subject Matter Experts review a random sample of non-coded, predicted non-responsive documents (not reviewed, skipped, suppressed duplicates) below a certain Cutoff Rank to determine the Elusion Rate in order to validate the Active Learning process.

7. **Iterative Training**: The process of repeatedly augmenting the Training Set with additional examples of documents coded by Subject Matter Expert(s) as Responsive or Non-Responsive until the effectiveness of the Machine Learning Algorithm reaches an acceptable level.

8. **Random Sample / Random Sampling**: Selection of a subset of the total document collection, using a method that is equally likely to select any document from the total document collection for inclusion in the Random Sample.

9. Subject Matter Expert(s) ("SMEs"): One or more individuals who are familiar with what is Responsive and can render an authoritative determination as to whether a document is Responsive or not.

10. Training Set: A sample of documents coded by one or more Subject Matter Expert(s) as Responsive or Non-Responsive, from which a Machine Learning Algorithm then infers how to distinguish between Responsive or Non-Responsive documents beyond those in the Training Set.

II. SCOPE OF THIS PROTOCOL

1. The procedures described in this protocol apply only to text-rich, electronically stored information ("ESI") collected by ICE from the custodians in Appendix A. *See* ECF No. 69-3. The custodians listed in Appendix A are the following: ICE Headquarters personnel Tae Johnson and Melissa Harper; two National Juvenile Coordinators; and fourteen Field Office Juvenile Coordinators ("FOJC") from ten field offices. Certain file types shall be excluded from the Active Learning process described in this protocol, including, but not limited to, media (for example, .jpg, .jpeg, .bmp, .gif, .wmv, .wav, .mov, .avi), spreadsheets (for example, .xls, .xlsx, .csv), system files, container files, documents containing little or no text, documents which contain too much text and extensible markup language (for example, .xml) files (collectively, "Excluded File Types"). ICE will deduplicate the ESI collection across custodians (Global Deduplication) before beginning the Active Learning process. ICE will also apply the search terms listed in Appendix B before beginning the Active Learning process and documents that do not hit on any of the terms (unless part of a family with a document that has a search term) will be omitted from the Active Learning process described in this protocol. Absent other agreement,

the remaining files will be subject to the Active Learning process described in this protocol ("Documents").

2. Nothing in this protocol shall prevent the United States from using other search, review, or coding methodologies in addition to, or in place of, Active Learning to help identify Documents that are responsive to the document requests from the Plaintiff or to review Documents for potentially privileged or confidential information.

III. PROTOCOL

The United States has determined that it will use Relativity's Active Learning to assist with a responsiveness review. The Active Learning process is described in the sections below.

1. Step One – Training Phase

After Documents are loaded into the Active Learning platform, the SMEs will begin coding documents for responsiveness. Once the reviewers have coded at least five documents as responsive and at least five documents as non-responsive, the Active Learning model will be built. Once the model is built, the prioritized review will commence. During the prioritized review process, the Active Learning system presents the SMEs with additional documents to code for responsiveness. This process continues until a particular Elusion Rate is met.

2. Step Two – Validation of Active Learning Through the Elusion Test

ICE will perform an Elusion Test to validate the Active Learning process. The Elusion Test is conducted by creating a statistically valid Random Sample, using a 95% Confidence Level and a 5% margin of error, of documents which fall below a particular Cutoff Rank (likely non-responsive documents). The SMEs will manually code that sample of documents for responsiveness. ICE will then determine how many documents were coded as responsive and

divide that number by the total number of documents in the sample. The resulting percentage is the Elusion Rate. If the target Elusion Rate is not met, ICE will continue training the system until the desired Elusion Rate is met. Thus, the validation process may occur several times. After training has concluded, ICE will determine the appropriate Cutoff Rank and Elusion Rate.

3. Step Three –Review

Documents above the Cutoff Rank will then be reviewed for confidentiality and privilege. During the course of confidentiality and privilege review, should non-responsive documents be identified, they will be coded as such

4. Step Four – Production

Responsive, non-privileged Documents will be produced pursuant to the 502(d) Order and Protective Order.

DATED: December 3, 2018

Respectfully submitted,

COLIN A. KISOR
Deputy Director

_/s/Yamileth G. Davila__
YAMILETH DAVILA
Assistant Director
U.S. Department of Justice
Civil Division
Office of Immigration Litigation
District Court Section
P.O. Box 868
Ben Franklin Station
Washington, DC 20044
Tel: (202) 305-0137
Fax: (202) 305-7000
Yamileth.g.davila@usdoj.gov

Counsel for Defendant

APPENDIX A

	NAME	TITLE	LOCATION	DATE RANGE
1.	Harper, Melissa	Chief, Juvenile and Family Residential Management Unit, Enforcement and Removal Operations (ERO)	ICE HQ	November 1, 2017 – October 26, 2018
2.	Johnson, Tae	Assistant Director of Custody Management, ERO	ICE HQ	January 1, 2016 – October 16, 2018
3.	Ravenell, Eric	National Juvenile Coordinator	ICE HQ	January 1, 2016 – October 26, 2018
4.	Helland, Dawn	National Juvenile Coordinator	ICE HQ	June 1, 2018 – October 26, 2018
5.	Munguia, Jose	FOJC	San Antonio Field Office	January 1, 2016 - October 26, 2018
6.	Sullivan, Nancy Z.	FOJC	Phoenix Field Office	October 1, 2016 – October 26, 2018
7.	Love, Joe	FOJC	San Francisco Field Office	October 1, 2016 – September 30, 2018
8.	Black, Lika	FOJC	Seattle Field Office	March 12, 2018 – October 26, 2018
9.	Galvez, Victor	FOJC	El Paso Field Office	October 1, 2016 – October 26, 2018
10.	Barnes, Edwin	FOJC	Houston Field Office	January 1, 2016 – October 26, 2018
11.	Hyde, Linda	FOJC	New York Field Office	January 1, 2016 – October 26, 2018
12.	Jones, Rebecca	FOJC	Washington Field Office	June 10, 2018 – October 26, 2018
13.	Johnson, Kareem	FOJC	New York Field Office	August 25, 2016 – October 26, 2018
14.	DeJesus, Carlos	FOJC	New York Field Office	August 25, 2016 – October 26, 2018
15.	Garcia, Alexis	FOJC	Phoenix Field Office	January 1, 2017 – October 26, 2018

16.	Pepple, Geoffrey	FOJC	Chicago Field Office	January 1, 2016 – October 26, 2018
17.	Kaskanlian, Andrew	FOJC	San Francisco Field Office	October 1, 2018 – October 26, 2018
18.	Reardon, Sean	FOJC	Miami Field Office	January 1, 2016 – October 26, 2018

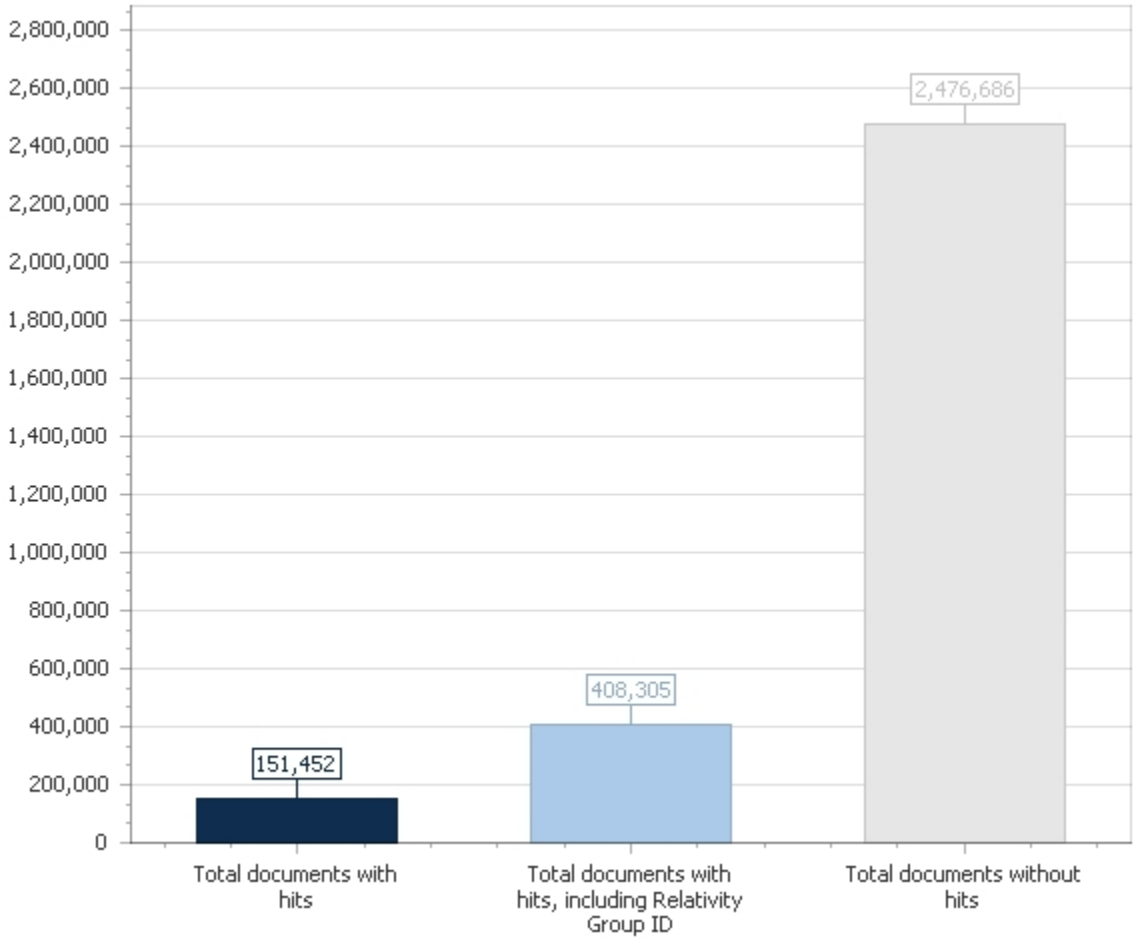
APPENDIX B

OPLA - DCLD - GARCIA-RAMIREZ eMAILS

Search Terms Report

Report Name: Baseline Custodians - STR

Searchable Set: Global Search - baseline custodians



Results Summary

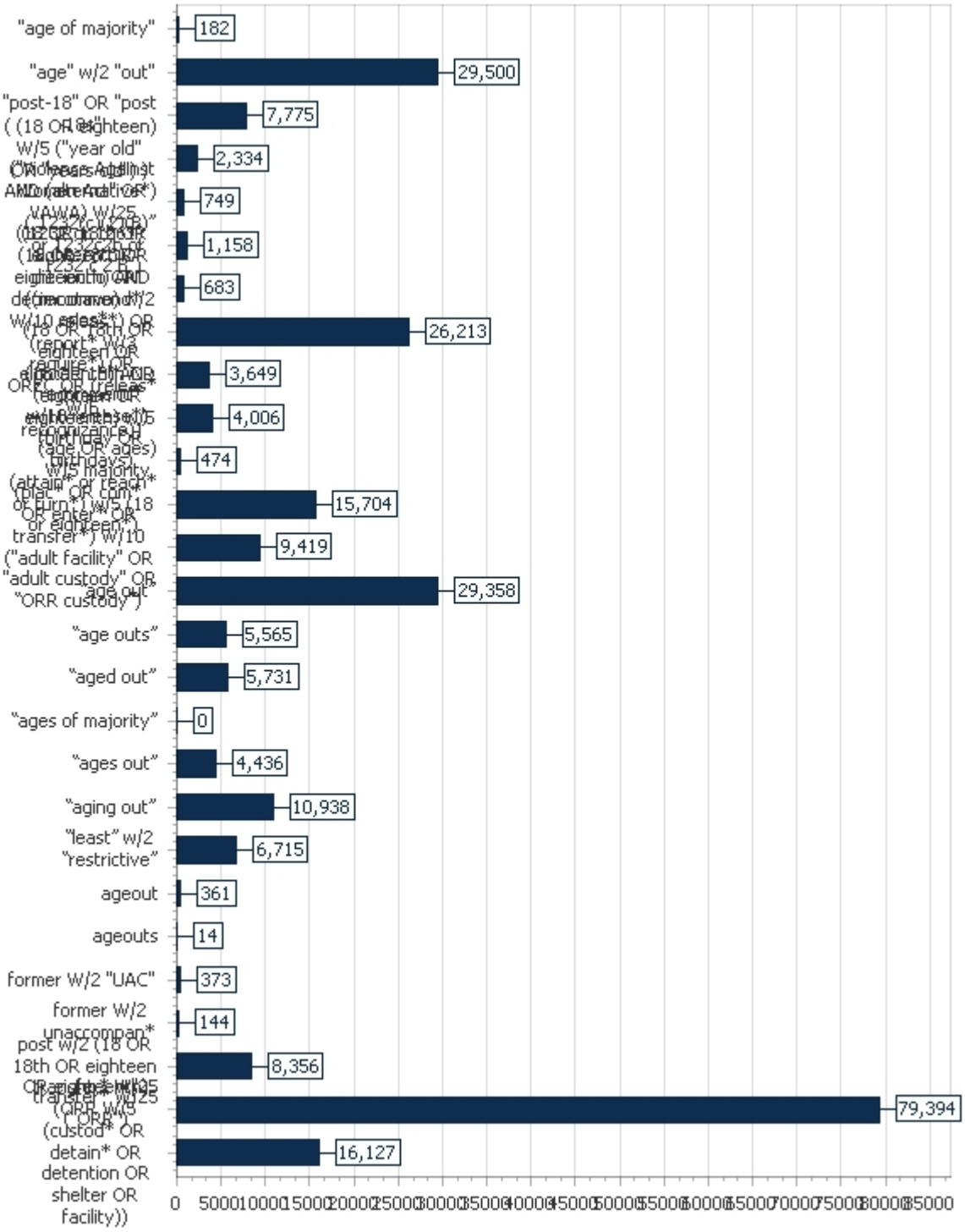
Documents in searchable set	Total documents with hits	Total documents with hits, including Relativity Group ID	Total documents without hits
2,884,991	151,452	408,305	2,476,686

OPLA - DCLD - GARCIA-RAMIREZ eMAILS

Search Terms Report

Report Name: Baseline Custodians - STR

Searchable Set: Global Search - baseline custodians



OPLA - DCLD - GARCIA-RAMIREZ eMAILS

Search Terms Report

Report Name: Baseline Custodians - STR

Searchable Set: Global Search - baseline custodians

Terms Summary

Term	Documents with hits	Documents with hits, including Relativity Group ID	Unique hits
"age of majority"	182	642	0
"age" w/2 "out"	29,500	62,211	39
"post-18" OR "post-18s"	7,775	19,923	0
((18 OR eighteen) W/5 ("year old" OR "years old")) AND (alternative*)	2,334	10,656	969
("Violence Against Women Act" OR VAWA) W/25 (1232 or 1261)	749	3,201	0
("1232(c)(2)(B)" or 1232c2b or "1232 c 2 B")	1,158	3,175	29
(18 OR 18th OR dieciocho OR decimoctavo) w/2 anos*	683	4,339	606
(18 OR 18th OR eighteen OR eighteenth) AND ((recommend* W/10 releas*) OR (report* W/3 require*) OR OREC OR (releas* W/5 recognizance))	26,213	109,846	14,693
(18 OR 18th OR eighteen OR eighteenth) AND (recommend* w/10 release*)	3,649	20,455	0
(18 OR 18th OR eighteen OR eighteenth) w/5 (birthday OR birthdays)	4,006	15,564	713
(age OR ages) W/5 majority	474	1,684	33
(attain* or reach* or turn*) w/5 (18 or eighteen*)	15,704	50,753	4,987

OPLA - DCLD - GARCIA-RAMIREZ eMAILS

Search Terms Report

Report Name: Baseline Custodians - STR

Searchable Set: Global Search - baseline custodians

Term	Documents with hits	Documents with hits, including Relativity Group ID	Unique hits
(plac* OR com* OR enter* OR transfer*) W/10 ("adult facility" OR "adult custody" OR “ORR custody”)	9,419	32,480	2,276
“age out”	29,358	61,913	0
“age outs”	5,565	15,746	1,296
“aged out”	5,731	18,776	1,661
“ages of majority”	0	0	0
“ages out”	4,436	11,822	722
“aging out”	10,938	27,820	4,738
“least” w/2 “restrictive”	6,715	33,314	1,433
ageout	361	726	29
ageouts	14	50	8
former W/2 "UAC"	373	966	111
former W/2 unaccompan*	144	349	2
post w/2 (18 OR 18th OR eighteen OR eighteenth)	8,356	23,586	384
transfer* W/25 ("ORR")	79,394	220,681	59,256
transfer* W/25 (ORR W/5 (custod* OR detain* OR detention OR shelter OR facility))	16,127	52,218	0

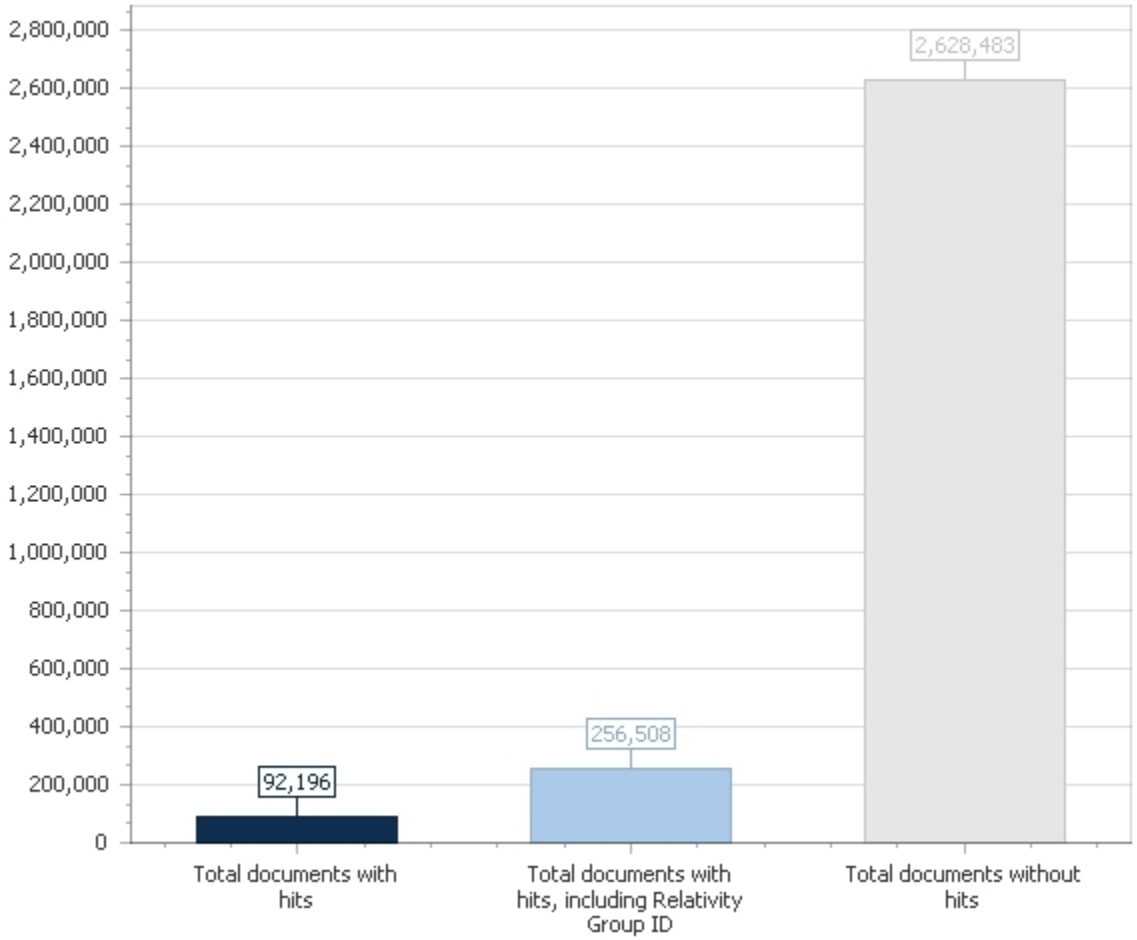
Exhibit B

OPLA - DCLD - GARCIA-RAMIREZ eMAILS

Search Terms Report

Report Name: Baseline Custodians - STR - Revised

Searchable Set: Global Search - baseline custodians



Results Summary

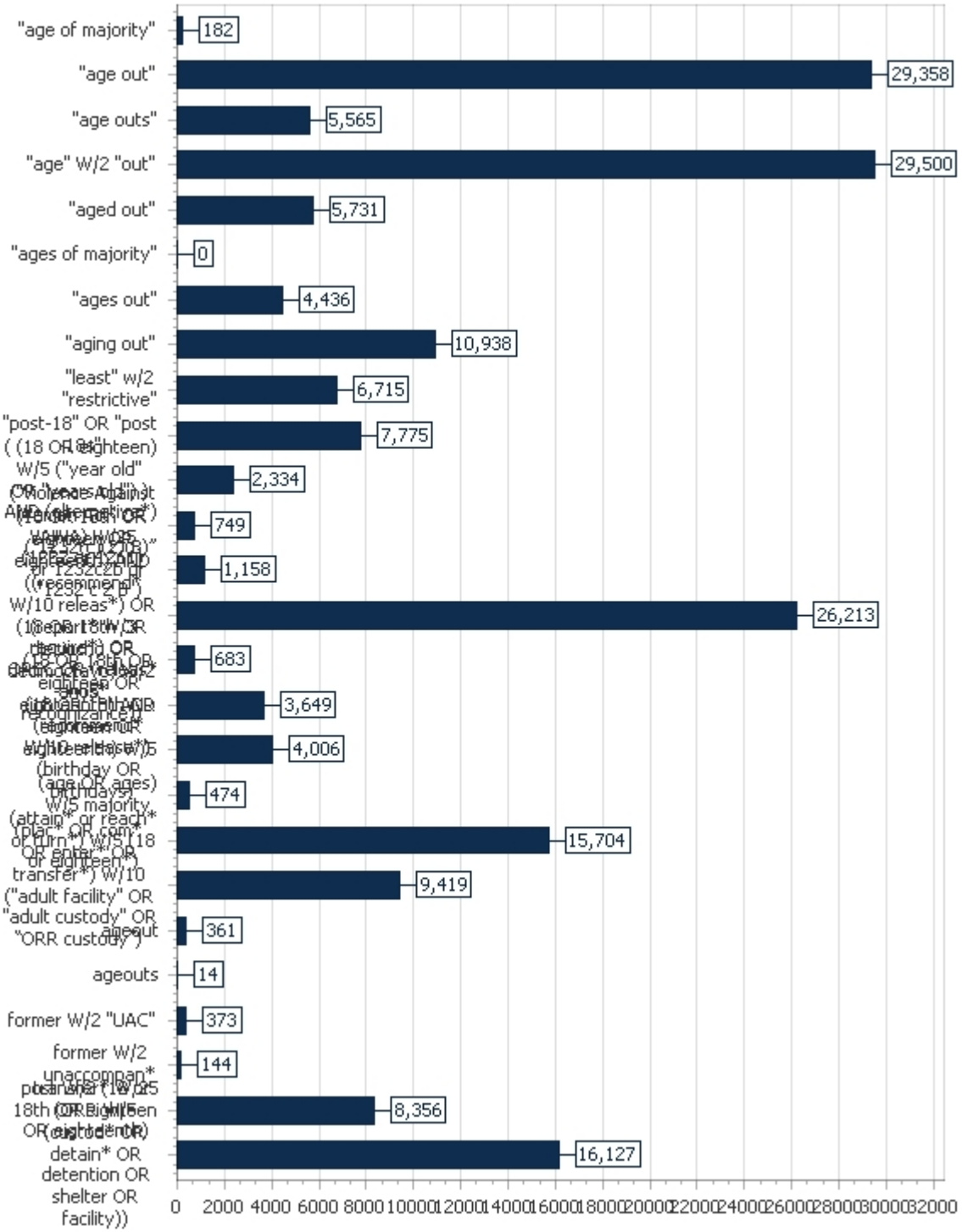
Documents in searchable set	Total documents with hits	Total documents with hits, including Relativity Group ID	Total documents without hits
2,884,991	92,196	256,508	2,628,483

OPLA - DCLD - GARCIA-RAMIREZ eMAILS

Search Terms Report

Report Name: Baseline Custodians - STR - Revised

Searchable Set: Global Search - baseline custodians



OPLA - DCLD - GARCIA-RAMIREZ eMAILS
Search Terms Report

Report Name: Baseline Custodians - STR - Revised

Searchable Set: Global Search - baseline custodians

Terms Summary

Term	Documents with hits	Documents with hits, including Relativity Group ID	Unique hits
"age of majority"	182	642	0
"age out"	29,358	61,913	0
"age outs"	5,565	15,746	1,316
"age" W/2 "out"	29,500	62,211	39
"aged out"	5,731	18,776	1,707
"ages of majority"	0	0	0
"ages out"	4,436	11,822	725
"aging out"	10,938	27,820	4,773
"least" w/2 "restrictive"	6,715	33,314	1,751
"post-18" OR "post-18s"	7,775	19,923	0
((18 OR eighteen) W/5 ("year old" OR "years old")) AND (alternative*)	2,334	10,656	993
("Violence Against Women Act" OR VAWA) W/25 (1232 or 1261)	749	3,201	0
(“1232(c)(2)(B)” or 1232c2b or "1232 c 2 B")	1,158	3,175	29
(18 OR 18th OR eighteen OR eighteenth) AND ((recommend* W/10 releas*) OR (report* W/3 require*) OR OREC OR (releas* W/5 recognizance))	26,213	109,846	15,011
(18 OR 18th OR dieciocho OR decimoctavo) W/2 anos*	683	4,339	611
(18 OR 18th OR eighteen OR eighteenth) AND (recommend* W/10 release*)	3,649	20,455	0

OPLA - DCLD - GARCIA-RAMIREZ eMAILS

Search Terms Report

Report Name: Baseline Custodians - STR - Revised

Searchable Set: Global Search - baseline custodians

Term	Documents with hits	Documents with hits, including Relativity Group ID	Unique hits
(18 OR 18th OR eighteen OR eighteenth) W/5 (birthday OR birthdays)	4,006	15,564	739
(age OR ages) W/5 majority	474	1,684	33
(attain* or reach* or turn*) W/5 (18 or eighteen*)	15,704	50,753	5,301
(plac* OR com* OR enter* OR transfer*) W/10 ("adult facility" OR "adult custody" OR “ORR custody”)	9,419	32,480	3,108
ageout	361	726	29
ageouts	14	50	8
former W/2 "UAC"	373	966	119
former W/2 unaccompan*	144	349	2
post W/2 (18 or 18th OR eighteen OR eighteenth)	8,356	23,586	384
transfer* W/25 (ORR W/5 (custod* OR detain* OR detention OR shelter OR facility))	16,127	52,218	10,092

Exhibit C

DRAFT PLAN FOR REVIEW OF ELECTRONICALLY STORED INFORMATION

Garcia-Ramirez v. ICE, No. 18-508 (D.D.C. filed Mar. 5, 2018)

December 4, 2018

SCOPE OF REVIEW: ICE collected 2.8 million documents (not pages) from 18 custodians.

PROPOSED SCHEDULE:

December 3, 2018	ICE ran a set of 26 agreed-upon search terms against the 2.8 million documents, resulting in: - Over 151,000 documents with hits - Over 408,000 documents in a family of documents with hits (for example, if an attachment to an email has a hit, the email and any other attachments are part of the family of documents).
December 5, 2018	ICE begin relevancy review in Relativity using Active Learning.
December 10, 2018	ICE begin privilege review of relevant documents. ICE has found that initial training of reviewers is crucial for an accurate production and much of the first week is consumed by training and responding to questions. ICE estimates that it is able to dedicate approximately 10 attorneys to the review process for 4 hours per day, 5 days a week, for first-line review. ICE is able to dedicate approximately 5 attorneys for second-line review. In addition, once documents are reviewed and ready for production, ICE has found that production of large volume of documents takes an average of 2 days.
December 14, 2018	ICE complete relevancy review.
December 21, 2018	Rolling production begins. ICE to make good faith efforts to produce 5,000 documents. Due to holidays, ICE anticipates a limited number of reviewers will be available from December 24 – January 11, 2018 and therefore cannot commit to document production during these weeks.
January 18, 2019	Rolling production continues. ICE to makegood faith efforts to produce an additional 5,000 documents. Rolling production will continue every week with ICE making its best efforts to produce 5,000 documents every week until all relevant documents are produced.

BASIS OF CALCULATIONS:

Based on prior document reviews, ICE has found that, on average, a reviewer can review and redact approximately 25 documents per hour for privilege, including describing the privilege for a privilege log. ICE has also found that second-line review is critical during production. Second-line reviewers are anticipated to review approximately 50 documents per hour.

Exhibit D