

EXHIBIT A

**UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA**

_____)	
WILMER GARCIA RAMIREZ, <i>et al.</i> ,)	
)	
Plaintiffs,)	
)	Civil Action No. 1:18-CV-00508-RC
v.)	
)	Class Action
UNITED STATES IMMIGRATION AND)	
CUSTOMS ENFORCEMENT, <i>et al.</i> ,)	
)	
Defendants.)	
)	
_____)	

**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

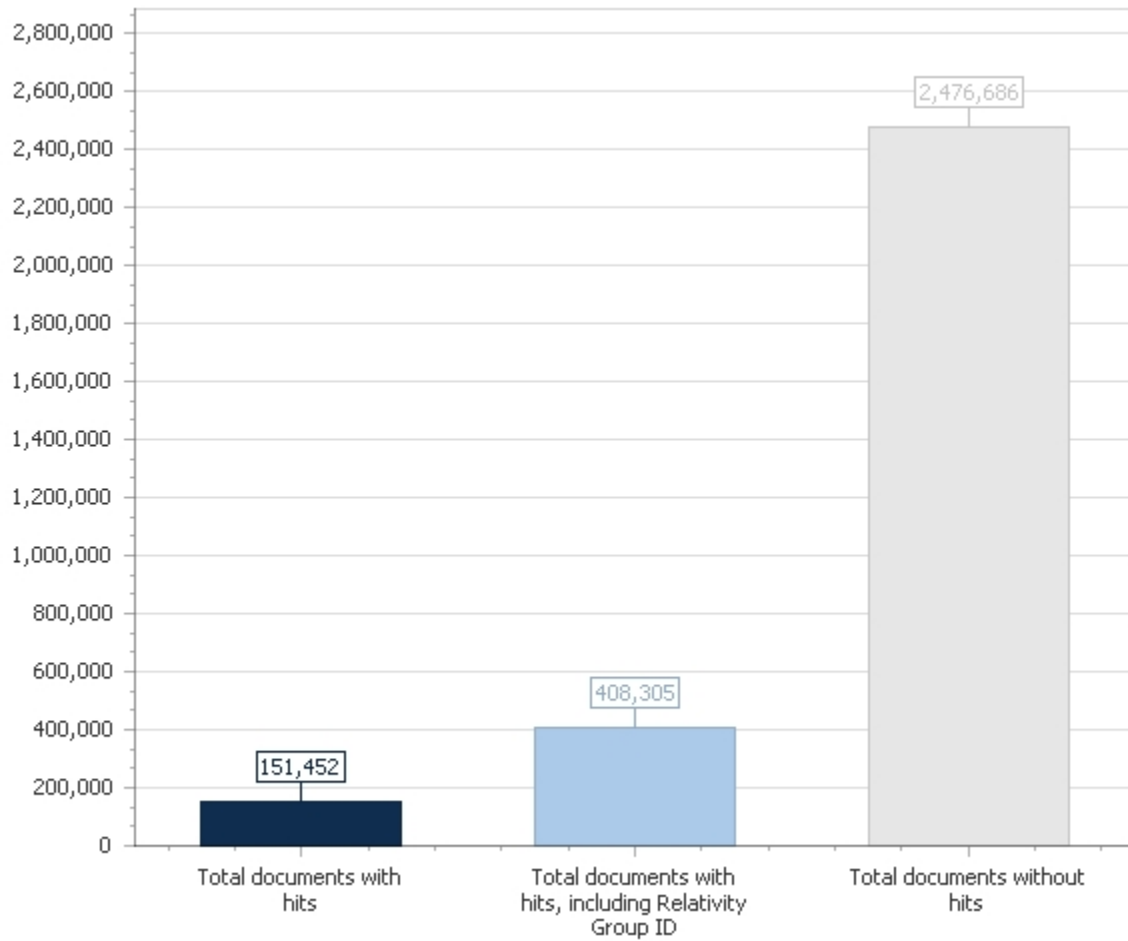
	NAME	TITLE	LOCATION	DATE RANGE
1.	Harper, Melissa	Chief, Juvenile and Family Residential Management Unit, Enforcement and Removal Operations (ERO)	ICE HQ	November 1, 2017 – October 26, 2018
2.	Johnson, Tae	Assistant Director of Custody Management, ERO	ICE HQ	January 1, 2016 – October 16, 2018
3.	Ravenell, 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

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

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 decimioctavo) 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

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