21-1233

IN THE UNITED STATES COURT OF APPEALS
FOR THE SECOND CIRCUIT

AMERICAN CIVIL LIBERTIES UNION IMMIGRANTS' RIGHTS PROJECT,
Plaintiff-Appellant,

v.

UNITED STATES IMMIGRATION AND CUSTOMS ENFORCEMENT,
Defendant-Appellee.

---

*On Appeal from the United States District Court
for the Southern District of New York*

**BRIEF OF *AMICI CURIAE* THE CENTER FOR INVESTIGATIVE
REPORTING, THE MEDIA LAW RESOURCE CENTER, AND THE
MUCKROCK FOUNDATION IN SUPPORT OF APPELLANT AND
REVERSAL**

Dated: August 27, 2021

Mason A. Kortz
Cyberlaw Clinic, Harvard Law School
Wasserstein Hall, Suite WCC 5018
1585 Massachusetts Avenue
Cambridge, MA 02138
(617) 495-2845
mkortz@law.harvard.edu

Counsel for *amici curiae*

CORPORATE DISCLOSURE STATEMENT

Amici curiae the Center for Investigative Reporting, the Media Law Resource Center, and the MuckRock Foundation represent that they have no parent corporations and that they have no stock, and therefore, no publicly held company owns 10% or more of their stock.

Dated: August 27, 2021   /s/ Mason A. Kortz

                                         Mason A. Kortz
                                         Counsel for *amici curiae*

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Cases

Legislative Materials

Docket Materials

Other Authorities

STATEMENT OF INTEREST OF *AMICI CURIAE*

Amici curiae consist of three media and media-related organizations that regularly engage in data-focused journalism.[1] Collectively, *amici* have deep experience with the technical aspects of working with structured data, including databases, and with government transparency processes, including the Freedom of Information Act. As such, *amici* have a significant interest in a strong right of access to records held in government databases.

The Center for Investigative Reporting ("CIR"), founded in 1977, is the nation's oldest nonprofit investigative newsroom. CIR produces investigative journalism for its website, https://www.revealnews.org/, the Reveal national public radio show and podcast, and various documentary projects. CIR often works in collaboration with other newsrooms across the country.

The Media Law Resource Center, Inc. ("MLRC") is a non-profit professional association for content providers in all media, and for their defense lawyers, providing a wide range of resources on media law and policy issues. These include news and analysis of legal, legislative and regulatory developments;

---

[1] Pursuant to Federal Rule of Appellate Procedure 29(a)(4)(E), *amici* certify that no party's counsel authored this brief in whole or in part; no party or party's counsel contributed money that was intended to fund the preparation or submission of this brief; and no person — other than the *amici*, their members, or their counsel — contributed money that was intended to fund the preparation or submission of this brief. Pursuant to Federal Rule of Appellate Procedure 29(a)(2), *amici* represent that all parties have consented to the filing of this brief.

litigation resources and practice guides; and national and international media law conferences and meetings. The MLRC also works with its membership to respond to legislative and policy proposals and speaks to the press and public on media law and First Amendment issues. It counts as members over 130 media companies, including newspaper, magazine and book publishers, TV and radio broadcasters, and digital platforms, and over 200 law firms working in the media law field.

The MuckRock Foundation is a journalism and government transparency non-profit that has helped thousands of requesters around the United States better file, share, and understand Freedom of Information requests. This work has often involved obtaining and analyzing federal databases, including data on the government's 1033 program that led to reforms of this program. They often work with agency FOIA personnel and IT departments to help craft requests for data that protects privacy and reduces the burden on agency staff while providing key insights into government operations.

SUMMARY OF ARGUMENT

The Freedom of Information Act and the subsequent Electronic Freedom of Information Amendments make it clear that advances in record-keeping technology are to come with commensurate advances in public access to government records. However, some agencies seem to take a narrow view of this directive. In this case, U.S. Immigration and Customs Enforcement asserts that, because information in its electronic databases is organized around a personal identifier, it can lock up its records and throw away the key—or at least withhold it as exempt. This position is technically inaccurate, legally indefensible, and in direct contradiction to the presumption of transparency that underlies FOIA and E-FOIA.

Databases store both individual data points and information about what those data points mean. This combination of content and structure means that when records are extracted from a database, they can be quickly and efficiently filtered, sorted, or restructured into whatever arrangement is most suited for the task at hand. A database cannot, however, create information about of thin air. Any information in a database, from a single value to the entire database schema, is fixed and recorded as concretely as any paper record.

Immigration and Customs Enforcement nevertheless argues that the "new records" doctrine obviates its duty to produce records that it controls. This is inconsistent with the development of that doctrine, both before and after the

3

widespread government use of databases. From the start, and more emphatically since 1996, courts have held that searching and sorting information in a database does not constitute creation of a new record, even if the process requires some degree of programmatic data manipulation. That rule is no different when applied to relational information stored in databases, as is the case here.

Ultimately, Immigration and Customs Enforcement advances a rule that would allow agencies to hide public information about government activities behind technical distinctions. Rather than focus on *how* an agency organizes its data, the technically and legally correct approach asks *what* records an agency possesses. Accordingly, *amici* urge this Court to rule that querying a database does not constitute the creation of a new record, at least where the query does not require analysis or judgment beyond that inherent in the FOIA process.

4

ARGUMENT

Since its passage in 1967, the Freedom of Information Act ("FOIA") has been updated and interpreted to fulfill its core purpose: "to hold the governors accountable to the governed." *NLRB v. Robbins Tire & Rubber Co.*, 437 U.S. 214, 242 (1978). With the Electronic Freedom of Information Amendments of 1996 ("E-FOIA"), Congress made clear that "[t]he format in which data is maintained is not relevant under the FOIA." H.R. Rep. No. 104-795, at 19-20 (1996). Nevertheless, as federal agencies have shifted overwhelmingly to electronic records, courts have been tasked with deciding whether, and to what extent, agencies' data management practices impact the availability of agency records. It is imperative that courts make these decisions with a complete understanding of the underlying technologies, including electronic databases. Otherwise, agencies might use technology to withhold information from the public, rather than to "enhance the operation of the Act" as Congress intended. *Id.* at 11.

The present case exemplifies this concern. The information the American Civil Liberties Union Immigrants' Rights Project ("ACLU") seeks—records of apprehensions, detentions, bond determinations, and removals on a per-person basis—undoubtedly exists in U.S. Customs and Immigration Enforcement's ("ICE")

databases.[2] Decl. of Vassilio-Diaz at JA037-38 ¶¶ 6-9. Yet the district court, at ICE's

urging, held that producing this information would require creation of new records.

Opinion at JA182-87. In reaching this conclusion, the district court focused on the

fact that there is no single, non-exempt column in ICE's databases that connects the

records to distinct individuals. *Id.* at JA186. The result is that ACLU received a

jumbled set of records that seem designed to frustrate the purpose of their request.

Records in government databases have provided the foundation for reporting

on matters of significant public interest, from police violence to disaster buyouts to

nursing home abuse.[3] *Amici*, as organizations deeply invested in the dissemination

---

[2] Two of ICE's databases are at issue in this case: the Enforcement Integrated
Database ("EID") and the ICE Integrated Decision Support ("IIDS"). Decl. of
Vassilio-Diaz at JA037-40, ¶¶ 6-13. Both databases contain some or all of the
records ACLU seeks. *Id.* While ICE's search focused on IIDS, *id.* ¶ 17, the
apprehension, detention, removal, and bond data in IIDS is copied from EID. *See*
Ex. A to Decl. of Susan Long, ECF No. 116-2, *Long v. U.S. Immigration & Customs
Enforcement*, No. 5:17-CV-00506-BKS-TWD (N.D.N.Y. June 25, 2021) at 10-13
[hereinafter "ICE Data Management Plan"]. Accordingly, *amici* understand that
ACLU's request could be fulfilled from EID alone if necessary.

[3] *Get The Data: Explore Data on All Police Shootings From the Nation's 50 Largest
Local Police Departments*, VICE News (Dec. 10, 2017), https://news.vice.com/
en_us/article/a3jjpa/nonfatal-police-shootings-data [https://perma.cc/AZ2H-A6PJ];
Robert Benincasa, *Search The Thousands Of Disaster Buyouts FEMA Didn't Want
You To See*, NPR (Mar. 5, 2019), https://www.npr.org/2019/03/05/696995788/
search-the-thousands-of-disaster-buyouts-fema-didnt-want-you-to-see
[https://perma.cc/343U-DSTS]; Daniel Simmons-Ritchie, *Still Failing the Frail:
The Data and Records Behind Our Reporting and Our Database*, PennLive (Oct.
14, 2018), https://www.pennlive.com/midstate/2018/10/still_failing_the_frail_the_
da.html [https://perma.cc/5K7S-VC4T].

of public information, respectfully request that this Court correct the errors in the district court's analysis, reverse the judgment below, and articulate a rule that extracting existing information from an agency database does not constitute the creation of a new record, at least where this process requires no more independent analysis or judgment than that inherent in the FOIA process.

I.   The data and structural information stored in ICE's relational databases are records under FOIA.

The district court erroneously held that the information ACLU seeks—specifically, the correlation of immigration enforcement events[4] to individuals—only exists in the abstract. Opinion at JA186. This analysis rests on an incorrect understanding of how information is stored in databases generally and relational databases in particular. A relational database is not just a collection of unrelated data points; it also contains structural information about how the data is organized. This structural information is as concrete as any other information held by an agency, and thus equally amenable to disclosure under FOIA.

A.   Relational databases contain structured data that is optimized for retrieval in multiple different arrangements.

Databases are powerful tools for modern record-keeping. Unlike paper files or discrete electronic documents, which can only exist in one configuration at a time,

---

[4] *Amici* use the term "enforcement event" to refer collectively to apprehensions, detentions, bond determinations, and removals.

7

databases are designed to provide access to information in multiple different arrangements with minimal effort. Databases have this functionality because they store information in a highly structured, easily searchable form. As the District Court of D.C. has explained,

> A traditional FOIA search could include physically retrieving 1000 different paper forms or "records," each of which has been placed in a separate file, and each of which contains only one relevant paragraph. This would be analogous to (and much more difficult than) a computer "query search" for those paragraphs. If defendant has categorized its data by size of employer, for example, and the plaintiff's request can be retrieved by a single search or a simple series of searches, then the information exists in the form of parts of multiple "records."

Thompson Pub. Grp., Inc. v. Health Care Fin. Admin., No. 1:92-CV-02431-LFO, 1994 WL 116141, at *2 (D.D.C. Mar. 15, 1994). In order to provide this flexibility, databases contain two types of information. The first is the *data*—discrete values such as text, numbers, and so on. The second is the *schema*—information that explains what the data values mean, how they are grouped together, and how these grouping relate to each other. The schema describes the data in such a way that it can be easily searched, sorted, and otherwise manipulated.

Relational databases[5] are databases that use a particular model for optimizing data storage and retrieval. In a typical relational database, the schema describes multiple *tables*, each of which is composed of rows and columns. Generally, each

---

[5] To the best of *amici*'s knowledge, EID and IIDS are both relational databases. *See* Decl. of Wu at JA096 ¶ 11.

8

table represents a conceptual unit (like a person or a case) or a relationship between conceptual units (like which people were involved with which cases and how). The schema also describes the relationships between tables, usually by designating *keys*—columns that are the same in multiple tables. *See Long v. U.S. Dep't of Justice*, 450 F. Supp. 2d 42, 48 (D.D.C. 2006), *order amended on reconsideration*, 457 F. Supp. 2d 30 (D.D.C. 2006), *amended*, 479 F. Supp. 2d 23 (D.D.C. 2007) ("In a relational database, information is subdivided into tables, and connections or linkages can be programmed between the tables, allowing the data maintained in one table to be related to the data stored in another."). The process of separating data into multiple related tables is called *normalization* and is an important practice that helps prevent errors from being introduced into the database.[6] *Description of the Database Normalization Basics*, Microsoft Docs, https://docs.microsoft.com/en-us/office/ troubleshoot/access/database-normalization-description [https://perma.cc/54QU-EPGU].

---

[6] ICE avers that extracting the requested records from IIDS would be difficult in part because "IIDS was not designed as a well-formed and normalized database," Supp. Decl. of Vassilio-Diaz at JA104-05 ¶¶ 8-10. However, public records show that IIDS is at least partially normalized. *See* ICE Data Management Plan at 21 (noting that tables in IIDS are normalized to the Second or Third normal form). In any event, ICE has not averred that EID is not normalized and, to the best of *amici*'s understanding, the requested records exist in that database as well. *See* Department of Homeland Security, *Privacy Impact Assessment for the Enforcement Integrated Database*, at 2 (Jan. 14, 2010) (explaining that EID "provides users the capability to access a person-centric view of [enforcement event] data").

However, while normalized data is optimal for database operations, it is often not user-friendly. Users can extract a particular arrangement of data from a database by running a *query*. A query is an instruction that tells a database management system to select a specific subset of information from a database and return it the requested arrangement.[7] In a relational database, a query specifies which rows and columns to extract from which tables, how they should line up with each other, and how they should be sorted, filtered, or otherwise manipulated. Thus, users rarely interact directly with data as stored in a database; rather, they interact with it as conceptually organized by a query. This allows agencies to store data in a structure that is optimized for data management but interact with that data in whatever arrangement is most appropriate for the task at hand.

B. Structured data created and retained by an agency is a record subject to FOIA.

FOIA obligates an agency "to provide access to those [records] which it in fact has created and retained." *Kissinger v. Reporters Comm. for Freedom of the Press*, 445 U.S. 136, 152 (1980). As explained in the prior section, relational

---

[7] There are other types of queries that instruct a database management system to insert, update, or delete data from a database, or to alter the structure of the tables, rows, or columns that make up the database. Such queries are beyond the scope of FOIA, as they would require agencies to create, alter, or destroy records. For the purpose of this brief, *amici* use the term "query" to refer exclusively to queries that select information without changing the structure or contents of the database.

10

databases work by spreading data across multiple tables when it is stored and recombining it when it is accessed. However, no database, however well designed, can create information out of thin air. When an agency chooses to store data in a database, the agency is creating a record. *See Yeager v. Drug Enf't Admin.*, 678 F.2d 315, 321 (D.C. Cir. 1982) ("The type of storage system in which the agency has chosen to maintain its records cannot diminish the duties imposed by the FOIA.").

By way of analogy, imagine a collection of government contracts—classic FOIA records. If an agent labels a file folder "Contracts FY 2020," they are creating and recording a piece of structural information: records placed in this folder are relevant to the 2020 fiscal year. If an agent then places a contract in that folder, they are recording another piece of information: that this contract is relevant to the 2020 fiscal year. If the agency later received a FOIA request for all contracts entered into in fiscal year 2020, the contracts in the folder would be responsive records. This is true even if the contracts themselves are not dated; the combination of the schema (i.e., the label on the folder) and the data (i.e., the contracts in the folder) is sufficient to identify the records as responsive. *See People for the Am. Way Found. v. U.S. Dep't of Just.*, 451 F. Supp. 2d 6, 15 (D.D.C. 2006) (explaining that an agency must employ available information to "identify[] . . . pre-existing agency records that are indisputably within [the agency's] control").

Relational databases are similar to this imaginary filing system—except that every contract would be parsed, annotated, and cross-referenced with other records before being filed. While this information is more complex than a label on a folder, it is no more abstract. *Cf. Forsham v. Harris*, 445 U.S. 169, 185 (1980) (finding information existed "in the abstract" where it had never been in the agency's possession). Moreover, and perhaps most relevant for FOIA purposes, this structural information would be created as each contract is added to the database and, as a result, would constitute a record under FOIA. *See Am. Immigr. Laws. Ass'n v. Exec. Off. for Immigr. Rev.*, 830 F.3d 667, 678 (D.C. Cir. 2016) (explaining that once an agency defines a collection of related material, that collection is a record).

C. The district court erred in holding that the information ACLU seeks does not exist in ICE's databases.

The district court's assessment of what information does (and does not) exist in ICE's databases, which formed the basis of its new records doctrine analysis, was flawed in two ways. First, it found that the relational information ACLU seeks would have to be "extrapolated" from the A-Numbers in the databases. Opinion at JA185. Second, it suggested that "information about [the] contents" of a database cannot be a record. *Id*. Neither of these findings reflect a correct understanding of the organization and use of relational databases.

The district court held that the relationships between enforcement events in ICE's database were "conceptual abstractions," *id.* at JA186, and that extracting

12

them from existing information would require extrapolation, *id.* at JA185. The term "extrapolation" suggests that some analytic process would be required. This is simply not the case. The information that an event is related to a specific person, and thus to other events involving that person, is created when the event is added to the database. For example, in IIDS, each enforcement event contains an EID_PERSON_ID column,[8] which is a "foreign key" (as designated by the "FK" after it in the schema). ICE Data Management Plan at 43-93. The foreign key designation shows that EID_PERSON_ID is used to relate enforcement events to another table—in this case a table containing each individual's personal profile, including their A-Number. *Id.* When an enforcement event is added to the database, the EID_PERSON_ID column would be populated with the same value in both the relevant enforcement event table and the personal information table, recording the link between the event and the individual. When retrieving information about the event via a query, that linkage would be explicit in the database—no judgment, analysis, or extrapolation would be required.

The district court also held that "ACLU's request . . . does not seek the contents of the IIDS database . . . but instead seemingly seeks information about

---

[8] It is not clear from the schema whether the EID_PERSON_ID column itself contains an A-Number, a name, or some other personally identifying information. If it is not, *amici* are at a loss to explain why ICE could not simply disclose the EID_PERSON_ID as an anonymous unique identifier.

those contents." Opinion at JA185. In a sense, this is correct; a database schema describes the organization of the contents of the database, including the relationships between data points. However, there is no rule that agency records cannot describe other records. Email headers, *Ctr. for Pub. Integrity v. United States Dep't of Com.*, 401 F. Supp. 3d 108, 120-21 (D.D.C. 2019), cover sheets, *Jud. Watch, Inc. v. U.S. Dep't of Treasury*, 796 F. Supp. 2d 13, 29 (D.D.C. 2011), and database schemas, *Long v. Immigr. & Customs Enf't*, 149 F. Supp. 3d 39, 48 (D.D.C. 2015), are all examples of records "about the content" of other records. The correct question, then, is not whether the requested information describes the organization of other records, but whether the requested information was created and maintained by the agency prior to the request. In this case, it is clear that ICE was recording the relationships between immigration enforcement events long before ACLU filed its request. *See* Department of Homeland Security, *Privacy Impact Assessment for the Enforcement Integrated Database*, at 2 (Jan. 14, 2010) (describing operation of EID in 2010).

II. Querying structured data from ICE's relational databases does not entail creating a new record.

For the reasons explained above, the records ACLU seeks exist, concretely, in ICE's databases. While this addresses the district court's primary contention with the request, the court also cited two cases about the process of extracting data from databases: *National Security Counselors v. C.I.A.*, 898 F. Supp. 2d 233 (D.D.C. 2012), *aff'd sub nom. Nat'l Sec. Couns. v. Cent. Intel. Agency*, 969 F.3d 406 (D.C.

Cir. 2020), and *People for the American Way Foundation v. U.S. Dep't of Justice*, 451 F. Supp. 2d 6 (D.D.C 2006). Accordingly, *amici* address the reasoning in these two cases and urge the Court to hold that fulfilling ACLU's request would not amount to creation of a new record.

> **A.   The "new records" doctrine does not apply to databases searches that merely restructure or reorganize records without the application of human judgment or analysis.**

The principle that "[FOIA] only requires disclosure of certain documents which the law requires the agency to prepare or which the agency has decided for its own reasons to create," *NLRB v. Sears, Roebuck & Co.*, 421 U.S. 132, 162 (1975), predates the widespread use of databases by federal agencies. As agency databases have become more ubiquitous and more complex, courts have sometimes struggled to determine whether database queries produce existing records or create new ones. *See, e.g.*, *Nat'l Sec. Couns.*, 898 F. Supp. 2d at 270 (describing the application of the new records doctrine to databases as "somewhat muddled"); *Am. Civil Liberties Union v. Arizona Dep't of Child Safety*, 240 Ariz. 142, 150 (Ct. App. 2016) ("We acknowledge that distinguishing between searching an electronic database and creating a new record . . . may be a difficult task."). Nevertheless, three guiding principles can be discerned.

The first, longstanding principle is that merely reorganizing or recompiling information does not constitute creation of a new record, even if the result is a

15

document that did not previously exist. *Schladetsch v. U.S. Dept of H.U.D.*, No. 1:99-CV-00175-ESH, 2000 WL 33372125, at *3 (D.D.C. Apr. 4, 2000) ("The fact that . . . the net result of complying with [a] request will be a document the agency did not previously possess is not unusual in FOIA cases, nor does this preclude the applicability of the Act."). This rule predates the widespread use of databases and the passage of E-FOIA. *Everytown for Gun Safety Support Fund v. Bureau of Alcohol, Tobacco, Firearms & Explosives*, 403 F. Supp. 3d 343, 357 (S.D.N.Y. 2019), *rev'd and remanded on other grounds*, 984 F.3d 30 (2d Cir. 2020) (citing *Disabled Officer's Ass'n v. Rumsfeld*, 428 F. Supp. 454, 456-57 (D.D.C. 1977)).

The second principle stems from the first: searching or sorting a database does not constitute creation of a new record. Even courts that have applied the new records doctrine broadly agree that the E-FOIA is unambiguous on this point. *See Nat'l Sec. Couns.*, 898 F. Supp. 2d at 270 (citing Pub. L. No. 104–231, 110 Stat. 3048 (1996)); *People for the Am. Way*, 451 F. Supp. 2d at 14 (same). This is true even if the search requires some data manipulation, *Everytown*, 403 F. Supp. 3d at 360 (citing H.R. Rep. No. 104-795, at 22 (1996)), or if additional programming is required to effect the search, *Schladetsch*, 2000 WL 33372125 at *3.

The third principle is that the new records doctrine exists to prevent agencies from being required to conduct analysis, answer questions, or make changes that alter the substance of a record. *See, e.g., Sears*, 421 U.S. at 161–62 (agency not

16

required to draft explanatory material); *DiViaio v. Kelley*, 571 F.2d 538, 542 (10th

Cir. 1978) (agency not required to respond to "interrogatories relating [to records]");

Tereshchuk v. Bureau of Prisons, 67 F. Supp. 3d 441, 451-52 (D.D.C. 2014) (agency

not required to create detailed summary of existing records); *Hudgins v. I.R.S.*, 620

F. Supp. 19, 21 (D.D.C. 1985), *aff'd*, 808 F.2d 137 (D.C. Cir. 1987) (agency not

required to conduct legal research on requester's behalf).

Taking these three principles together, courts have held that "using a query to

search for and extract a particular arrangement or subset of data already maintained

in an agency's database does not amount to the creation of a new record." *Ctr. for

Investigative Reporting v. United States Dep't of Just.*, 982 F.3d 668, 691 (9th Cir.

2020); *Everytown*, 403 F. Supp. 3d at 357. This is also precisely the point of E-FOIA.

As Senator Orrin Hatch explained,

> A specific "record" may not be created until a query is formed and the
> software associated with the database manipulates the information,
> which in turn compiles the record formulated by the query. Because the
> database itself is a public record, then any record created from
> information stored in that database is also a public record.

S. Rep. No. 104-272, at 29 (1996). Extracting data in multiple arrangements is the

entire point of keeping information in a database: the agency does not need to

commit to a particular arrangement of information. Rather, it has a multitude of

different arrangements at its fingertips, each of which is in the agency's "possession

or control." *Kissinger*, 445 U.S. at 152. To say that some of these available

17

arrangements "have been *in fact* obtained" by the agency, *Forsham*, 445 U.S. at 186,

but others have not, would amount to arbitrary line-drawing.

> **B.   The district court erred in holding that substituting unique identifiers for A-Numbers would entail creation of new records.**

As explained in Section I, *supra*, all the information needed to fulfill ACLU's

request is already stored in ICE's databases. Moreover, as explained in the preceding

section, extracting this data using a query does not entail creation of a new record,

even if the precise arrangement of data is novel. The sole remaining issue, then, is

whether ICE is obligated to transform A-Numbers into anonymous but unique

identifiers as part of that query. It is, because doing so is functionally equivalent to

other database operations that are required under FOIA.

Courts have long been in agreement that reorganizing information in records

does not constitute record creation. *See Disabled Officer's Ass'n v. Rumsfeld*, 428 F.

Supp. 454, 456 (D.D.C. 1977) (denying summary judgment where request required

agency to compile names and addresses from disparate records into single, novel

document). Fewer courts have addressed whether this extends to reorganizing

information in individual database columns. However, in *Center for Investigative

Reporting v. Department of Justice*, the Ninth Circuit faced just this question. There,

the plaintiff had requested data that was indisputably in the defendant agency's

database, namely, information about crimes involving firearms that had once

belonged to law enforcement officers. 982 F.3d at 674-75. However, due to the

18

operation of the law funding the database, the plaintiff opted to request annual totals rather than individual traces. *Id*. The defendant argued that because it did not have a spreadsheet or other document with those totals pre-calculated, the plaintiff was requesting a new record. *Id*.

The Ninth Circuit disagreed, holding that "using a query to search for and extract a particular arrangement or subset of data already maintained in an agency's database does not amount to the creation of a new record." *Id*. at 691. Noting that the database in question already contained "close-out codes" identifying the source of each firearm, the court concluded that

> ATF could produce the precise statistical aggregate data that CIR seeks, with no further counting or analysis required, if, for example, a query or queries for the relevant close-out codes produces a "hit count" reflecting the number of records involving a firearm traced to law enforcement, the number of matching records is contained in [the database's] metadata, or if the database produces an otherwise responsive result separate from the trace data itself.

Id. at 693; *see also Everytown*, 403 F. Supp. 3d at 359 (holding that "producing aggregate data or other descriptive information about an agency database" does not "necessarily constitute[] record creation").

The same logic applies to ACLU's request here. All that is needed to fulfill the request is for ICE to extract existing data grouped by the given criteria: individual identities. *See* Office of Information Policy, *Defining a "Record" under the FOIA*, U.S. Department of Justice (July 23, 2021), https://www.justice.gov/oip/oip-

19

guidance/defining_a_record_under_the_foia (recommending that agencies treat collections or grouping of information as discrete records). There are multiple ways ICE could achieve this without disclosing exempt records. One is to produce the records and substitute random unique values for the A-Numbers after the fact. *See* Appellant's Brief at 29. Another is to do the grouping as part of the search, using JOINs to correlate enforcement events by the A-Number column without disclosing the contents of that column for any given record. *See SQL Joins*, W3Schools, https://www.w3schools.com/sql/sql_join.asp [https://perma.cc/VA78-LNBT] ("A JOIN clause is used to combine rows from two or more tables, based on a related column between them."). A third option—and perhaps the most elegant—would be to use a one-way hashing function to transform A-Numbers into encrypted strings as part of each query.[9] This is analogous to the aggregation functions at issue in *Center for Investigative Reporting* and *Everytown for Gun Safety*; the only

---

[9] ICE avers that in order to accomplish this transformation, it would have to "develop a standardized formula." Supp. Decl. of Vassilio-Diaz at JA104-05 ¶ 10. *Amici* are unaware of the exact capabilities of ICE's databases. However, all major database management software with which *amici* are familiar support one-way hashing as a built-in function, obviating the need for ICE to develop its own formula. *See, e.g.*, *HASHBYTES (Transact-SQL)*, Microsoft Docs, https://docs.microsoft.com/en-us/sql/t-sql/functions/hashbytes-transact-sql?view=sql-server-ver15 [https://perma.cc/D64D-76PL]; *STANDARD_HASH*, Oracle Help Center, https://docs.oracle.com/database/121/SQLRF/functions183.htm#SQLRF55647 [https://perma.cc/CKQ4-L4BS].

difference is that instead of aggregating individual data points into totals, ICE would transform individual data points into anonymized, purely relational information.

Ultimately, the exact approach used by ICE should not be the deciding factor. Whether a query returns a single row of data, aggregated data, or relational information spread across multiple tables, the fact that a particular result has not previously been extracted does not mean that it is a "new" record for the purposes of FOIA. Rather, any query that "extract[s] and compile[s] existing data" should be treated as the retrieval of existing records. *Long v. Immigr. & Customs Enf't*, No. 1:17-CV-01097-APM, 2018 WL 4680278, at *4 (D.D.C. Sept. 28, 2018).

III. **This Court should rule that querying data from a database does not require creation of a new record so long as the query does not require significant judgment or analysis.**

For the reasons stated above, the district court's ruling does not align with either the technical realities of database management or the legal precedent of the new records doctrine. While this is a sufficient basis for reversal, *amici* respectfully advance one more argument: ICE's approach to record management represents a threat to the public interests at the heart of FOIA. *See Ctr. for Investigative Reporting*, 982 F.3d at 690 ("[I]f running a search across these databases necessarily amounts to the creation of a new record, much government information will become forever inaccessible under FOIA, a result plainly contrary to Congress's purpose in enacting FOIA.")

21

Under the expansive view of "new records" advanced by ICE and adopted by the district court, agencies could shield such records from disclosure through cheap record management practices. All an agency would need to do is break records into pieces and label them with some exempt information (say, social security numbers or agents' names) and voila, the agency has now "encrypted" its records in a way that stymies public accountability. The practice is a bit like locking paper records in a cabinet and claiming that the key is exempt—a result that no agency could seriously argue is permissible.

ICE contends that ACLU's request is beyond the scope of FOIA in part because its databases are "case-centric" rather than "person-centric."[10] Decl. of Vassilio-Diaz at JA038 ¶ 7. This highlights the arbitrary nature of the technical distinctions ICE raises. Return for a moment to the hypothetical file of government contracts organized by fiscal year in Section I.B. If a requester sought all contracts with Acme Corporation, the agency could not avoid disclosure by saying its filing

---

[10] ICE also asserts that the software it uses to access its databases only allow for per-person queries on an *ad hoc* basis. Decl. of Vassilio-Diaz at JA038-39 ¶¶ 7, 12. However, if ICE's databases contain information sufficient to collate records by person one-by-one, it necessarily contains information sufficient to do the same *en masse*. The fact that ICE's interface does not have this capability is not material; an agency is not excused from searching records by the requested criteria simply because it interacts with its data another way. *See Prop. of the People, Inc. v. United States Dep't of Just.*, --- F. Supp. 3d ---, No. 1:18-CV-01202-CJN, 2021 WL 1209280, at *4 (D.D.C. Mar. 31, 2021) (holding agency search inadequate where agency searched only indexed columns, rather than content of records).

cabinets were "date-centric" instead of "company-centric." The operative question would be whether the agency had records on Acme, not how it stored them. This principle applies with equal force to records in databases. *See Ctr. for Investigative Reporting*, 982 F.3d at 692 ("So long as the relevant information and data fields already exist in the database maintained by the agency, the result produced by a query is an existing record, regardless of the form it takes.").

This is not to say courts cannot set limits on database operations under FOIA. In keeping with the purpose of the new records doctrine—to avoid tasking agencies with creating information beyond what they have chosen to maintain—judicial consensus is emerging that whether a database query constitutes creation of a new record "hinges not on whether the information is housed in the form requested, but whether generating the information requires the agency to engage in additional research or conduct additional analyses above and beyond the contents of its database." *Everytown*, 403 F. Supp. 3d at 359; *Ctr. for Investigative Reporting*, 982 F.3d at 693 (holding that application of the new records doctrine turns on "whether the information the requester] seeks could be produced by a reasonable search . . . or would require more significant human analysis"). Courts have properly placed the burden of showing the need for judgment or analysis on the agency. *See Long*, 2018 WL 4680278, at *7 (requiring agency to provide "sufficient detail to allow the court

23

to conclude that responding to the requests would require the creation of new data points, as opposed to the extraction and compilation of existing ones").

Of course, any search for records involves some degree of judgment or analysis. For example, if a requester seeks records by name, an agency is obligated to search for reasonable permutations of that name. *See Negley v. F.B.I.*, 658 F. Supp. 2d 50, 60 (D.D.C. 2009). If an agency comes across information that indicates the existence of additional responsive records, it must follow up accordingly. *Campbell v. U.S. Dep't of Just.*, 164 F.3d 20, 28 (D.C. Cir. 1998), *as amended* (Mar. 3, 1999). Searches of databases must allow at least this much leeway as well. *See Yeager*, 678 F.2d at 321 ("Although accessing information from computers may involve a somewhat different process than locating and retrieving manually-stored records, these differences may not be used to circumvent the full disclosure policies of the FOIA.").

In accordance with these principles and policy concerns, *amici* urge the court to adopt the following rule: querying an agency database does not constitute the creation of a new record, at least where the agency cannot show that the query would require analysis or judgment beyond what is ordinarily contemplated as part of the FOIA process. A more restrictive rule would foil the very purpose of FOIA.

CONCLUSION

For the reasons stated above, *amici* respectfully urge this Court to hold that ACLU's request would not require the creation of new records, REVERSE the decision below, and REMAND for further proceedings consistent with this holding.

Dated: August 27, 2021

/s/ Mason A. Kortz

Mason A. Kortz
Cyberlaw Clinic, Harvard Law School
Wasserstein Hall, Suite WCC 5018
1585 Massachusetts Avenue
Cambridge, MA 02138
(617) 495-2845
mkortz@law.harvard.edu

Counsel for *amici curiae*

CERTIFICATE OF COMPLIANCE

Pursuant to the Fed. R. App. P. 32(g)(1), I hereby certify that:

1.   This document complies with the type volume limitations of Second Circuit Local Rule 29.1(c) because, excluding the parts of the document exempted by Fed. R. App. P. 32(f), this document contains 5,958 words as calculated by the word count feature of Microsoft Word 365.

2.   This document complies with the typeface requirements of Fed. R. App. P. 32(a)(5) and the type-style requirements of Fed. R. App. P. 32(a)(6) because this document has been prepared in a proportionally spaced typeface using Microsoft Word 365 in 14-point Times New Roman.

Dated: August 27, 2021   /s/ Mason A. Kortz

                                                          Mason A. Kortz
                                                          Counsel for *amici curiae*

CERTIFICATE OF SERVICE

I hereby certify that I electronically filed the foregoing Brief of *Amici Curiae* the Center for Investigative Reporting, the Media Law Resource Center, and the MuckRock Foundation in Support of Appellant and Reversal with the Clerk of the Court for the United States Court of Appeals for the Second Circuit by using the appellate CM/ECF system on August 27, 2021. I certify that all participants in this case are registered CM/ECF users and that service will be accomplished by the appellate CM/ECC system.

Dated: August 27, 2021   /s/ Mason A. Kortz

                                          Mason A. Kortz
                                          Counsel for *amici curiae*