

DOJ Sensitive Data Rule

Covered Data Handout

The DOJ Sensitive Data Rule restricts or prohibits both you and Georgia Tech from allowing access by countries of concern or covered persons to specifically defined categories of U.S. government-related data and U.S. sensitive personal data when the quantity of data reaches specific “bulk” amounts or thresholds.

When you evaluate the data sharing activities involved in your research projects, it is important to assess carefully whether the DOJ Sensitive Data Rule might be applicable. In considering whether your research involves generating, receiving, providing, or otherwise allowing access to data regulated by the Rule, please include in your consideration:

- **De-identified** and/or **anonymized data**, as the DOJ Sensitive Data Rule allows for no exceptions for de-identified or anonymized data.
- **Human biospecimens** that can be used to derive any category of U.S. sensitive personal data.
- The **total amount** of data provided by Georgia Tech **and any external partners** to a project, data set, or repository.

Note: This Informational Resource is intended to accompany the Georgia Tech DOJ Sensitive Data Rule FAQs website. For more information about the DOJ Sensitive Data Rule and compliance with the DOJ Sensitive Data Rule at Georgia Tech, please visit <https://generalcounsel.gatech.edu/doj-sensitive-data-rule>.

Bulk U.S. Sensitive Personal Data

CATEGORIES OF U.S. SENSITIVE PERSONAL DATA	BULK VOLUME THRESHOLD
Human Genomic Data Meaning, data representing the nucleic acid sequences that constitute the entire set or a subset of the genetic instructions found in a human cell, including the result or results of an individual's “genetic test” (as defined in 42 U.S.C. 300gg-91(d)(17)) and any related human genetic sequencing data. Note: Include biospecimens from which this data may be derived.	100 U.S. citizens, nationals, permanent residents, or individuals located in the U.S.
Human ‘omic Data Including: • Human Genomic Data, as defined above; • Human Epigenomic Data, meaning data derived from a systems-level analysis of human epigenetic modifications, which are changes in gene expression that do not involve alterations to the DNA sequence itself. These epigenetic modifications include modifications such as DNA methylation, histone modifications, and non-coding RNA regulation. Routine clinical measurements of epigenetic modifications for individualized patient care purposes would not be considered epigenomic data under the DOJ	1,000 U.S. citizens, nationals, permanent residents, or individuals located in the U.S.

CATEGORIES OF U.S. SENSITIVE PERSONAL DATA

BULK VOLUME THRESHOLD

Sensitive Data Rule because such measurements would not entail a systems-level analysis of the epigenetic modifications in a sample;

- Human Proteomic Data, meaning data derived from a systems-level analysis of proteins expressed by a human genome, cell, tissue, or organism. Routine clinical measurements of proteins for individualized patient care purposes would not be considered proteomic data under the DOJ Sensitive Data Rule because such measurements would not entail a systems-level analysis of the proteins found in such a sample; and
- Human Transcriptomic Data, meaning data derived from a systems-level analysis of RNA transcripts produced by the human genome under specific conditions or in a specific cell type. Routine clinical measurements of RNA transcripts for individualized patient care purposes would not be considered transcriptomic data under the DOJ Sensitive Data Rule because such measurements would not entail a systems-level analysis of the RNA transcripts in a sample.

Note: Include biospecimens from which this data may be derived.

Biometric Identifiers

Meaning, measurable physical characteristics or behaviors used to recognize or verify the identity of an individual, including facial images, voice prints and patterns, retina and iris scans, palm prints and fingerprints, gait, and keyboard usage patterns that are enrolled in a biometric system and the templates created by the system.

1,000 U.S. citizens, nationals, permanent residents, or individuals located in the U.S.

Precise Geolocation Data

Meaning, data, whether real-time or historical, that identifies the physical location of an individual or a device with a precision of within 1,000 meters.

1,000 U.S. citizens, nationals, permanent residents, or individuals located in the U.S. (or their devices)

Personal Financial Data

Meaning, data about an individual's credit, charge, or debit card, or bank account, including purchases and payment history; data in a bank, credit, or other financial statement, including assets, liabilities, debts, or trades in a securities portfolio; or data in a credit report or in a "consumer report" (as defined in [15 U.S.C. 1681a\(d\)](#)).

10,000 U.S. citizens, nationals, permanent residents, or individuals located in the U.S.

Personal Health Data

Meaning, health information that indicates, reveals, or describes the past, present, or future physical or mental health or condition of an individual; the provision of healthcare to an individual; or the past, present, or future payment for the provision of healthcare to an individual. This term includes basic physical measurements and health attributes (such as bodily functions, height and weight, vital signs, symptoms, and allergies); social, psychological, behavioral, and medical diagnostic, intervention, and treatment history; test results; logs of exercise habits; immunization data; data on reproductive and sexual health; and data on the use or purchase of prescribed medications.

10,000 U.S. citizens, nationals, permanent residents, or individuals located in the U.S.

Personal Identifiers

Including any of the following listed identifiers in combination with another listed identifier or in combination with other data that may link a listed

100,000 U.S. citizens, nationals, permanent residents, or

CATEGORIES OF U.S. SENSITIVE PERSONAL DATA	BULK VOLUME THRESHOLD
identifier to another listed identifier or to another category of U.S. Sensitive Personal Data: • <u>demographic or contact data</u> (e.g., first and last name, birth date, birthplace, ZIP code, residential street or postal address, phone number, email address, or similar public account identifiers); • <u>full or truncated government ID or account number</u> (e.g., SSN, driver's license or State ID #, passport #, or Alien Registration #); • <u>full financial account numbers or personal ID numbers</u> associated with a financial institution or financial services company; • <u>network-based identifier</u> (e.g., IP address or cookie data); • <u>device-based or hardware-based identifier</u> (e.g., IMEI, MAC address, SIM card number); • <u>advertising identifier</u> (e.g., Google Advertising ID or Apple ID for Advertisers); • <u>account-authentication data</u> (e.g., account username, account password, or answers to security questions); and • <u>call-detail data</u> (e.g., Customer Proprietary Network Information).	individuals located in the U.S. (or their devices)

U.S. Government-Related Data

The DOJ Sensitive Data Rule generally defines [U.S. government-related data](#) as:

1. Any precise geolocation data (as defined in the table above), of any volume, for any location within any area on the Government-Related Location Data [List](#) and
2. Sensitive personal data (from one or more categories in the table above) that is marketed as linkable to employees, contractors, or officials of the United States government.

Note: There is no “bulk” threshold for U.S. government-related data.