



Privacy Impact Assessment of the Consumer Complaints Topic Modeling Process

For Questions or Comments, please email: privacy.comments@frb.gov

Description of the IT system:

In order to identify trends and risks within financial services, the Board of Governors of the Federal Reserve System's (Board) Division of Consumer and Community Affairs (DCCA), Risk and Surveillance unit, built the Consumer Complaints Topic Modeling Process. This process classifies consumer complaints submitted to the Consumer Financial Protection Bureau (CFPB) into topics. Consumer complaints are accessible through CFPB's Consumer Complaint Database.¹ These consumer complaints are made available to government agencies through CFPB's secure non-public Government Portal (Government Portal).

A machine learning natural language technique called "topic modeling," uses narrative data to help DCCA analysts understand trends in the types of consumer complaints received by CFPB. Data from the Government Portal (e.g., the consumer financial product associated with a complaint, the non-public complaint narrative, and public and non-public complaint IDs) are downloaded as part of the Consumer Complaints Topic Modeling Process. Complaint narratives are entered into the machine learning model, which generates a topic number and topic label. None of these data are made available to the public. The output of the Consumer Complaints Topic Modeling Process is for internal use only and does not contain personally identifiable information (PII).²

1. The information concerning individuals that is being collected and/or maintained:

The Board collects unredacted complaints from the Government Portal. CFPB advises consumers not to include PII in complaint narratives, though consumers may still choose to do so. To mitigate this risk, DCCA employs a masking algorithm to obscure social security numbers and dates (including dates of birth) provided by consumers in open-text fields. Structured fields containing PII are not downloaded or retained, other than the public and non-public complaint identifiers provided by CFPB, none of which are made available to the public through the Consumer Complaints Topic Modeling Process.

¹ CFPB Consumer Complaint Database <https://www.consumerfinance.gov/data-research/consumer-complaints/>.

² PII is information that can be used to distinguish or trace an individual's identity, either alone or when combined with other information that is linked or linkable to a specific individual. The Office of Management and Budget (OMB), [OMB Circular A-130, Managing Information as a Strategic Resource](#) (July 28, 2016).

2. Source(s) of each category of information listed in item 1:

CFPB's Government Portal.

3. Purposes for which the information is collected:

These data help DCCA analysts understand trends in consumer complaints and consumer risks related to problems people are experiencing in the marketplace, while also helping the Board conduct research to increase its understanding of the impacts of financial services policies and practices on consumers and communities.

4. Who will have access to the information:

Access to the Consumer Complaints Topic Modeling Process is restricted to authorized DCCA staff who require access for official business purposes. An individual's job responsibilities determine the type of access they are granted. Periodic reviews are conducted to determine whether an individual still requires access, and has the appropriate role to access these data. Disclosures may also be made under the Freedom of Information Act (5 USC §552).

5. Whether the individuals to whom the information pertains have an opportunity to decline to provide the information or to consent to particular uses of the information (other than required or authorized uses):

A complaint submitted to the CFPB is entirely voluntary, so individuals may choose not to file a complaint. Individuals do not have an opportunity to consent to the Board's particular uses of their information.

6. Procedure(s) for ensuring that the information maintained is accurate, complete and up-to-date:

Individuals submitting complaints to the CFPB are responsible for the accuracy, completeness, and timeliness of their information.

7. The length of time the data will be retained:

The data used by the Board for Consumer Complaints Topic Modeling is retained for less than 10 years.

8. The administrative and technological procedures used to secure the information against unauthorized access:

The Board applies applicable privacy and security controls in the National Institute of Standards and Technology's (NIST) Special Publication 800-53 (Rev. 5), *Security and Privacy Controls for Information Systems and Organizations*, at the moderate impact level.

