



**Federal Trade Commission
Privacy Impact Assessment
for the: Spam Database**

April 2009

1 **Introduction**

The Federal Trade Commission's (FTC) Bureau of Consumer Protection (BCP) protects consumers from a variety of fraudulent, deceptive, and unfair practices in the marketplace, including identity theft, telemarketing fraud, Internet fraud, and consumer credit issues. To further its consumer protection mission, the FTC brings civil and administrative law enforcement actions to enforce its laws and provides consumer and business education to enable the public to avoid common harms. The FTC works to ensure that consumers have accurate information for purchasing decisions, and confidence in the traditional and electronic marketplaces.

BCP's consumer protection-related activities include consumer complaint collection and analysis, individual company and industry-wide investigations, administrative and federal court litigation, rulemaking proceedings, consumer and business education, and the operation of consumer protection programs.

BCP's consumer protection-related activities also include enforcement of the Controlling the Assault of Non-Solicited Pornography and Marketing Act (the CAN-SPAM Act of 2003, 15 U.S.C. § 7704), which establishes requirements for those who send commercial email, spells out penalties for spammers and companies whose products are advertised in spam if they violate the law, and gives consumers the right to ask emailers to stop spamming them.

To support BCP's investigations and consumer protection-related activities, BCP has created a "Spam Database" (SpamDB). The SpamDB typically receives hundreds of thousands of messages each day, and approximately one-hundred million messages each year.

The SpamDB is managed by BCP's Division of Planning and Information (DPI). DPI and the FTC's Office of Information Technology Management (OITM) staff work together to maintain the SpamDB.

The SpamDB is primarily used by law enforcers (e.g. attorneys, investigators, paralegals) in the FTC's Bureau of Consumer Protection, and by technologists in DPI. Information from the SpamDB may also be shared with and then used by authorized law enforcement partners (e.g. the Department of Justice), and by authorized private sector and academic partners. In addition, the FTC may retain experts or contractors who may be given access to the SpamDB. These various groups are referred to in this PIA as "users."

2 **System Overview**

The SpamDB provides the public with an email address (spam@uce.gov) to which they can forward those email messages they believe to be spam (also known as unsolicited commercial email or "UCE"). Participation is voluntary, and the Commission's website

notifies consumers that the FTC maintains this information for use in law enforcement investigations.¹

The SpamDB is physically located in BCP's Washington D.C. Internet Lab (Lab) facility² and is isolated from the FTC production (computer) network that the rest of the agency uses by a firewall and router. This separation protects the FTC production network from malicious attachments frequently associated with spam while maintaining the integrity of emails collected in the database. Messages received by the FTC mail servers are routed to separate BCP mail servers, at which time an automated "script" (set of computer instructions) saves the files to the SpamDB using filenames that reflect the time and date they are received.

The SpamDB organizes and indexes the messages it receives, and makes those messages available to law enforcers via a web-based search engine. The database is accessible to local FTC staff through the workstations in the Lab. The SpamDB is also accessible to authorized remote users through the FTC's secure Virtual Private Network (VPN).³

Information users may obtain from the SpamDB is not saved or stored on Lab workstations. Users can download and save selected spam messages identified in their search and print or save them to digital media (e.g. CD/DVD) for use in their investigations. Lab policies require that all data collected in the Lab be removed by users once it is saved or printed. Information that is removed from the SpamDB (typically, to be included as part of a larger investigation file), is subject to internal FTC data protection and privacy policies, including those pertaining to the safeguarding of sensitive personally identifiable information and sensitive health information.

The SpamDB includes information in its "raw" / "unstructured" form – i.e. as plain text email messages and attachments. The SpamDB also includes index information, which facilitates the search and retrieval of messages, as well as various statistics that are derived from the information that has been submitted (e.g. the number of messages received in a given period).

3 Information Collected and Stored within the System

.1 What information is to be collected, used, disseminated, or maintained by the system?

The SpamDB collects and maintains information that is voluntarily submitted via email to spam@uce.gov. Any information included in an email message submitted to the SpamDB is included at the submitter's discretion.

¹ See <http://www.ftc.gov/ftc/contact.shtm>.

² The SpamDB is not part of the BCP Internet Laboratory (Lab); rather the SpamDB is housed within the Lab's server room. A discussion of the Internet Lab is available in that system's Privacy Impact Assessment, which is available at the following location - <http://www.ftc.gov/os/2009/04/internetlabpia.pdf>

³ Currently, only staff in the FTC regional offices have been provided with the ability to access the SpamDB remotely. However, such access may be provided to other law enforcers in the future.

The information collected in the SpamDB varies depending on what the submitter has chosen to forward to the FTC. Most often, emails submitted to the SpamDB include the body of the original email received by a consumer, along with standard email header information, which includes the email address of the consumer or entity that forwarded the email to the FTC. In addition, email header information includes sender and recipient email addresses, timestamps for each transmission between the sender and recipient, and a subject line.

Occasionally, messages forwarded to the SpamDB contain additional information, including personally identifiable information (PII) such as name, address, telephone numbers, and email addresses. Typically, such information is provided by those who forward messages to the SpamDB that include their “signature line” / contact information. In addition, because information is submitted to the SpamDB via email, messages can include more sensitive information (e.g. social security or tax ID numbers, credit card numbers, and bank account numbers). However, the submission of sensitive information does not occur frequently.⁴

.2 What are the sources of the information in the system?

Information is submitted by individuals (e.g. consumers) who voluntarily forward email messages to spam@uce.gov. Information also may be voluntarily submitted by other entities, including Internet Service Providers (ISPs), which may receive spam complaints from customers, or which may use other mechanisms (e.g. spam filters) to identify messages to forward to spam@uce.gov.

.3 Why is the information being collected, used, disseminated, or maintained?

Information is collected to support the FTC’s law enforcement mission and to enforce the CAN-SPAM Act (see section 1), including rulemaking under the Act. Information also may be collected to support research and studies, and to help develop consumer and business education materials and announcements.⁵

.4 How is the information collected?

Information is submitted by individuals and other entities, who voluntarily forward emails to spam@uce.gov (see section 2).

.5 How will the information be checked for accuracy and timeliness (currency)?

⁴ Based on a random sample of 300 messages contained within the SpamDB, approximately three percent (3%) of all submissions contain some signature line information, and approximately three tenths of one percent (0.3%) of all submissions may contain more sensitive information.

⁵ See <http://www.ftc.gov/spam/>, <http://www.ftc.gov/bcp/menus/consumer/tech/spam.shtml>, and <http://www.onguardonline.gov/> for more information.

The information forwarded to the SpamDB is not systematically checked for accuracy and timeliness. Messages collected from consumers are considered an accurate representation of the email as of the point-in-time it was received by the FTC. This information is provided voluntarily.

.6 Is the system using technologies in ways that the FTC has not previously employed (e.g., monitoring software, Smart Cards, etc.)? If so, how does the use of this technology affect individuals' privacy?

The SpamDB does not employ any new technologies, although the establishment of the SpamDB centralizes collection and preservation functions for efficiency. The SpamDB does not use technologies in ways that raise privacy concerns not otherwise discussed in this document. The FTC is not engaged in data mining and the data is not used for this purpose.⁶

.7 What law or regulation permits the collection of this information?

Information is collected pursuant to the FTC's general law enforcement and investigatory authority, which is primarily set forth in the Federal Trade Commission Act, 15 U.S.C. §§ 41-58. In addition, information is collected pursuant to the FTC's law enforcement and investigatory authority under the CAN-SPAM Act of 2003, 15 U.S.C. § 7704.

.8 Considering the type of information collected and sources of collection, what privacy risks were identified and how were these risks mitigated?

The SpamDB collects email messages that are voluntarily provided by individuals and other entities. This information primarily consists of the contents of the unsolicited commercial email messages (i.e. spam) sent to consumers, and any information that may be included by consumers when they forward these messages to the FTC.

As discussed above (see section 3.1), the vast majority of the information submitted to the SpamDB is non-sensitive commercial information, although each submission does include the submitter's email address information, and from time-to-time may contain more sensitive information.

Nonetheless, any privacy risks associated with that information are significantly mitigated by the way the FTC stores it. Most information in the SpamDB is stored in an "unstructured" format (see section 2.1).⁷ The SpamDB does not store the messages that consumers forward to the FTC in "rows" and "columns" of data that are easy to search,

⁶ See the Federal Agency Data Mining Reporting Act of 2007, Pub. L. No. 110-53, 121 Stat. 266, §804(b)(1) for a definition/description of the term "data mining."

⁷ As discussed in section 2.1, the SpamDB includes "raw" / "unstructured" information, as well as index information, and various statistics that are derived from the information that has been submitted. Information contained within the indexes is stored in a "structured" format that does not provide for easy access to sensitive information.

as is the case with a traditional database. Rather, most information in the SpamDB is stored as “raw” text email messages and attachments. Storing data in this way dramatically increases the computing power that is required to find meaningful information. This “natural” barrier is further enhanced by the sheer volume of information stored within the SpamDB – hundreds of millions of messages, stored on multiple terabytes of storage devices. This technical challenge, when coupled with the low volume of messages that are received that contain sensitive information, should serve to make the SpamDB a low value “target,” and a low risk.

In addition, several safeguards have been implemented to mitigate any residual risks that might be present, and to prevent disclosure of any sensitive information that might be stored in the database. The SpamDB is housed within BCP’s secure Internet Lab facility. To access the SpamDB, users must either be physically located in the Lab, or must have access to the SpamDB’s secure VPN. The SpamDB’s VPN complies with OMB and NIST guidelines regarding remote access to federal information systems, and includes multi-factor authentication controls. Only authorized users are given access to the SpamDB.

The FTC also follows applicable Federal Information Security Management Act (FISMA) requirements to ensure that information collected in the Spam Database is appropriately secured. For additional discussion of security, see section 7 below.

4 **Use and Access to Data in the System**

.1 Describe how information in the system will or may be used.

Information is collected to support the FTC’s law enforcement mission, including enforcement of the CAN-SPAM Act (see section 1), including sharing with other law enforcement partners.

.2 Which internal entities will have access to the information?

BCP investigators and case teams will have access to the information collected in the SpamDB, as will other authorized FTC and law enforcement personnel (see section 7.5).

.3 Which external entities will have access to the information?

On occasion, the FTC may share information with law enforcement and other government agencies, courts, and defendants, or as otherwise authorized by the law.

5 **Notice and Access for Individuals**

.1 How will individuals be informed about what information is collected, and how this information is used and disclosed?

Notice is provided to individuals on those FTC web pages that contain information about the SpamDB (i.e., pages that give the email address for consumers to forward copies of spam emails they have received to the database). Additionally the FTC Privacy Policy on the FTC's main web site, www.ftc.gov, provides consumers with notification about how the FTC collects, uses, shares, and protects personal information.⁸

.2 Do individuals have the opportunity and/or right to decline to provide information?

Yes. Individuals are not required to provide any information to the SpamDB. Submissions to the SpamDB are voluntary.

.3 Do individuals have the right to consent to particular uses of the information? If so, how would an individual exercise this right?

Consumers are notified of how the FTC intends to store and use the information provided (see section 4.1). Once the spam email is forwarded to the FTC, consumers do not have the right to consent to particular uses of the information.

.4 What are the procedures that allow individuals to gain access to their own information?

Individuals may seek access to information, if any, about themselves in the SpamDB by filing a written access request with the FTC's Office of General Counsel. The FTC's Privacy Act rules and procedures for making such a request are published in the Code of Federal Regulations at 16 C.F.R. 4.13.⁹ Because SpamDB records are maintained and used for law enforcement purposes, some or all portions of requested records may be legally exempt from disclosure and withheld from requesting individuals (e.g., investigatory targets).

.5 Discuss the privacy risks associated with the process of providing individuals access to their own records and how those risks are mitigated.

No privacy risks were identified, because the FTC does not give individuals access to their own records through the SpamDB. Rather, as explained in 5.4 above, individuals must file a written request with the FTC's Office of General Counsel to seek access.

6 Web Site Privacy Issues

⁸ The FTC's Privacy Policy is available at the following URL - <http://www.ftc.gov/ftc/privacy.shtm>.

⁹ See <http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&sid=e24bb12fb0671c830c6372c90efd2b5a&rgn=div8&view=text&node=16:1.0.1.1.5.0.5.13&idno=16>.

No website privacy issues were identified.

7 Security of Information in the System

- .1 Are all IT security requirements and procedures required by federal law being followed to ensure that information is appropriately secured?**

The FTC follows applicable Federal Information Security Management Act (FISMA) requirements to ensure that information in the Spam Database is appropriately secured.

- .2 Does the project employ technology that may raise privacy concerns? If so, please discuss its implementation.**

The Spam Database does not employ technologies that raise privacy concerns not already addressed in this document.

- .3 What procedures are in place to determine which users may access the system and are they documented?**

To access the SpamDB users must either have access to BCP's secure Internet Lab facility, or must have been given remote access via the SpamDB's secure VPN. Lab access is based on organization assignment. All BCP staff are granted access to the Lab as part of the FTC employee Check-In process. Other law enforcement users may request access to the SpamDB by contacting the Division of Planning and Information's (DPI) Assistant Director.

- .4 Describe what privacy training is provided to users either generally or specifically relevant to the program or system.**

All FTC employees are required to complete computer security training and privacy awareness training annually. Interactive online training covers topics such as how to properly handle sensitive PII and other data, online threats, social engineering, and the physical security of documents. Persons at the FTC with significant security responsibilities are required to undergo additional, specialized training, tailored to their respective responsibilities.

- .5 What auditing measures and technical safeguards are in place to prevent the misuse of data?**

An electronic keycard system restricts physical access to the Lab facilities. Once data is removed from the Lab, FTC policies for handling and safeguarding sensitive personally identifiable information and sensitive health information apply.

For remote access, users may only access the SpamDB via the system's secure VPN. Users who have obtained authorization to access the SpamDB remotely must be provided

with the appropriate software, have that software installed on the computers that are used to access the SpamDB, and have been issued valid security credentials (i.e. a user id, password, and secure “token”).

.6 State that any questions regarding the security of the system should be directed to the FTC’s Chief Information Security Officer.

Any questions regarding the security of BCP’s spam database will be directed to the FTC’s Chief Information Security Officer, Margaret Mech, at (202) 326-2609.

8 Data Retention

.1 For what period of time will data collected by this system be maintained?

Information received by the SpamDB is maintained in the database for as long as necessary for law enforcement or other authorized purposes. When incorporated into other agency records (e.g., investigatory files), the information is subject to applicable records disposition schedules and procedures established by the National Archives & Records Administration (NARA).

.2 What are the plans for destruction or disposal of the information?

Any disposal of information will be conducted in accordance with Office of Management and Budget (OMB), NIST, and NARA guidelines.

.3 Describe any privacy risks identified in the data retention and disposal of the information, and describe how these risks have been mitigated.

As previously stated, the database servers are located within a secured room in the Internet Lab facility. SpamDB administrators follow FTC policies for safeguarding sensitive personally identifiable information and sensitive health information, including the policies for the proper disposal of such information. The data maintained on the auxiliary servers and backup tapes is encrypted, and all backup tapes are stored in locked cabinets accessible only by BCP administrators.

9 Privacy Act

.1 Will the data in the system be retrieved by a personal identifier?

Email messages stored within the SpamDB are not organized by personal identifier. However, the SpamDB supports “key word” searching, which permits users to retrieve records by personal identifier.

.2 Is the system covered by an existing Privacy Act System of Records notice (SORN)?

The SpamDB is considered to be one part of the FTC's consumer complaint records system. For Privacy Act purposes, that system is called the "Consumer Information System--FTC," FTC-IV-1, and is described further in a Privacy Act SORN posted on the FTC's main web site.¹⁰ Information that is removed from the Lab and incorporated into investigatory records is considered part of a different FTC Privacy Act system, "Nonpublic Investigational and Other Nonpublic Legal Program Records," FTC-I-1, which is described in the separate SORN for that system.¹¹

10 Privacy Policy

.1 Confirm that the collection, use, and disclosure of the information in this system has been reviewed to ensure consistency with the FTC's privacy policy.

The SpamDB does not have a Web site that would be required to post a privacy policy, but the collection, use, and disclosure of the information in the SpamDB has been reviewed to ensure consistency with the FTC's privacy policy posted on the FTC's main Web site, see <http://www.ftc.gov/ftc/privacy.shtm>.

¹⁰ See <http://www.ftc.gov/foia/sysnot/iv-1.pdf>.

¹¹ See <http://www.ftc.gov/foia/sysnot/i-1.pdf>.

11 **Approval and Signature Page**

Prepared for the Business Owners of the System by:

_____ Date: _____
David Torok
Associate Director, BCP
Division of Planning Information

Review:

_____ Date: _____
Alexander C. Tang, Attorney
Office of the General Counsel

_____ Date: _____
Kellie Cosgrove Riley
Acting Chief Privacy Officer

_____ Date: _____
Margaret Mech
Chief Information Security Officer

Approved:

_____ Date: _____
Stanley Lowe
Chief Information Officer