FSIS Establishment-Specific Data Release Strategic Plan

United States Department of Agriculture

Food Safety and Inspection Service

January 2014

Table of Contents

Executive Summary

The Food Safety and Inspection Service (FSIS) within the United States Department of Agriculture (USDA) inspects meat, poultry and processed egg product establishments to ensure that the food produced in them is safe, wholesome and properly labeled. FSIS' mission is carried out by performing a variety of important activities, such as conducting inspections, then verifying and enforcing FSIS regulations and policies at establishments, and performing microbiological and chemical residue sampling at establishments. These activities, among many other key FSIS functions, produce a large volume of data. For example, FSIS collects data on regulated, domestic slaughter and processing establishments and product from equivalent foreign country inspection systems. Data from most FSIS activities are stored in a multi-dimensional data storage application, known as the FSIS Data Warehouse. FSIS produces reports for internal use, as well as reports and data to be shared publicly through the Agency's website[1] and through other public communication venues. Most of this data is shared with the public in an aggregated and/or summary format.

However, as a result of many factors, including recent policy documents calling for increased data sharing and greater transparency released by the Obama Administration and the Office of Management and Budget (OMB), the administrative burden Freedom of Information Act (FOIA) requests place on FSIS, the implementation of the Public Health Inspection System (PHIS), which allows for improved data collection and distribution, and the desire to share, in public fashion, data collected as a part of the Agency's activities, FSIS began exploring in 2010 how best to share establishment-specific data with the public. As a first step, FSIS consulted the National Advisory Committee on Meat and Poultry Inspection (NACMPI) in 2010.

Based on NACMPI review in 2010, FSIS went on to seek additional review from the National Research Council (NRC) within the National Academies to "conduct a study to examine the potential food-safety benefits and other consequences of making establishment-specific data publicly available on the Internet."[2] The NRC convened the *Committee for a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data* in 2010. In examining this question, the Committee reviewed FSIS' current data sharing activities, explored how other government agencies share data with the public, and recommended an approach for FSIS' release of establishment-specific data that considers the benefits and costs of data sharing.

In light of the NRC's recommendation to develop a data release plan, FSIS convened a new internal workgroup through the Agency's Data Coordination Committee (DCC). This workgroup included representatives from the Office of the Chief Information Officer (OCIO), the Office of Data Integration and Food Protection (ODIFP), the Office of Investigations, Enforcement and Audits (OIEA), the Office of Field Operations (OFO), the Office of Public Affairs and Consumer Education (OPACE), the Office of Public Health Science (OPHS), and the Office of Policy and Program Development (OPPD), along with special consultation with the FSIS Freedom of Information Act (FOIA) office. This work culminated in the development of this Strategic Plan. FSIS intends to share this draft of the Strategic Plan with Agency stakeholders for review and feedback, with the intention of finalizing and publicly sharing the Strategic Plan thereafter.

[1] For more information, please visit: www.fsis.usda.gov.
[2] National Research Council, Committee on a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data. The Potential Consequences of Public Release of Food Safety and Inspection Service Establishment-Specific Data. 2011. Available at: http://www.nap.edu/catalog.php?record_id=13304.

1.0 Introduction

Introduction

FSIS inspects meat, poultry and processed egg product establishments to ensure that the food produced in them is safe, wholesome and properly labeled. FSIS' mission is carried out by performing a variety of important activities, such as conducting establishment inspections, verifying and enforcing FSIS regulations and policies, and performing microbiological and chemical residue sampling. These activities, among many other key FSIS functions, produce a large volume of establishment-specific data. For example, FSIS collects and stores data on all federally-regulated, domestic slaughter and processing establishments, warehouses, transporters, and retail stores, as well as inspection findings at point-of-entry (POE) for product from equivalent foreign government inspection systems.

This Strategic Plan was developed to provide FSIS with a framework under which to responsibly and effectively release data collected by the Agency. This Plan was also developed to build on FSIS' 2010 *Strategic Data Analysis Plan for Domestic Inspection*,[3] as recommended by the NRC committee. This Plan seeks to achieve the following goals:

1. Describe FSIS' current data collection and release structures and activities
2. Describe Presidential and Office of Management and Budget (OMB) policy documents related to data sharing, as well as advice from the NACMPI and analyses from the NRC and FSIS to determine if and how FSIS should release establishment-specific data
3. Present FSIS' Data Release Procedures and Limitations
4. Describe FSIS' criteria for evaluating Agency datasets for public posting
5. Present FSIS' Prioritization list for data release
6. Identify potential performance measures to determine the effectiveness of data release

To achieve these goals and develop this Strategic Data Release Plan, FSIS convened a new Agency workgroup through the Agency's Data Coordination Committee (DCC). This workgroup included representatives from the Office of the Chief Information Officer (OCIO), the Office of Data Integration and Food Protection (ODIFP), the Office of Investigations, Enforcement and Audits (OIEA), the Office of Field Operations (OFO), the Office of Public Affairs and Consumer Education (OPACE), the Office of Public Health Science (OPHS), and the Office of Policy and Program Development (OPPD), along with special consultation with the FSIS Freedom of Information Act (FOIA) office. The format for this Plan will follow the goals described above, with additional material contained within the Appendices.

[3] United States Department of Agriculture, Food Safety and Inspection Service. FSIS Strategic Data Analysis Plan for Domestic Inspection. September 2010. Available at: http://www.fsis.usda.gov/wps/wcm/connect/84fa563e-0f5c-4df5-8e04-99a04e9ce102/2010_Strategic_Data_Analysis_Plan.pdf?MOD=AJPERES.

2.0 FSIS Data Collection and Data Structures

FSIS' employees (inspectors, veterinarians, laboratorians, enforcement investigations and analysis officers, among other job titles) conduct inspections, ensure compliance with existing regulations, and collect and test microbiological and chemical residue samples, along with other sample collection programs. These employees routinely collect on behalf of the Agency non-proprietary, establishment-specific data over the course of their inspection, verification, and sampling activities. These data are collected on all federally-regulated processing or slaughter establishments. An example of FSIS inspection data can be found in Hazard Analysis and Critical Control Point (HACCP) verifications, which inspectors conduct to determine if an establishment is using proper sanitation procedures, for example, as dictated in FSIS regulations. FSIS microbiological sampling results in the collection of another type of data; namely information on the absence or presence of a pathogen such as *Salmonella*, on the regulated product, such as chicken, that was sampled. This Plan focuses primarily on two types of FSIS data; 1) Inspection and Enforcement Data and 2) Sampling and Testing Data. While FSIS collects other types of data such as Microbiological Baseline Study Data and establishment-specific molecular-typing data, such as pulsed-field gel electrophoresis (PFGE), these types of data were not included in this version of the Agency's Data Release Plan.

FSIS Data Systems

Data from most FSIS applications is stored in the FSIS data warehouse. FSIS produces reports for internal use, as well as reports and data to be shared publicly through the Agency's website[4] and through other public communication venues. Most of this data is shared in an aggregated and/or summary format. For example, FSIS posts quarterly progress reports from the Agency's *Salmonella* and *Campylobacter* verification sampling programs.[5] These reports contain aggregated *Salmonella* and *Campylobacter* sampling results, including positive pathogen testing results, for a variety of product classes. As another example, FSIS posts, on a quarterly basis, summary reports on the enforcement actions the Agency has taken to ensure that products that reach consumers are safe, wholesome, and properly labeled.[6]

In limited situations, FSIS also releases information through the Agency's website on establishment-level characteristics. For example, in the past, FSIS posted on its website young chicken (broiler) establishments that fell into Category 3 based on their *Salmonella* test results as related to the Agency's performance standards. [7] Currently, FSIS posts on its website official enforcement actions the Agency has taken against establishments that have been found in violation of the Humane Methods of Slaughter Act.[8]

FSIS does release a large volume of disaggregated, establishment-specific data to the public through the Freedom of Information Act (FOIA) requests.[9] FSIS maintains a FOIA office and has a website that provides information to requestors. This site contains information on how to submit requests, annual

[4] For more information, please visit: www.fsis.usda.gov.
[5] For more information, please visit: http://www.fsis.usda.gov/wps/portal/fsis/topics/data-collection-and-reports/microbiology/quarterly-reports-salmonella/quarterly-progress-reports.
[6] For more information, please visit: http://www.fsis.usda.gov/wps/portal/fsis/topics/regulatory-compliance/regulatory-enforcement/quarterly-enforcement-reports/qer-index.
[7] Establishments that produce young chicken and turkeys are sampled for *Salmonella* and *Campylobacter* and then categorized by FSIS based on the number of positive samples identified during a sampling set. For more information, please visit: http://www.fsis.usda.gov/wps/portal/fsis/topics/data-collection-and-reports/microbiology/salmonella-verification-testing-program/salmonella-verification-testing-program.
[8] For more information, please visit: http://www.fsis.usda.gov/wps/portal/fsis/topics/regulatory-compliance/regulatory-enforcement/humane-handling-enforcement-actions/humane-handling-enforcement-actions.
[9] The Freedom of Information Act 5 USC § 552, As Amended By Public Law 104-231, 110 Stat. 3048. Available at: http://www.justice.gov/oip/foia_updates/Vol_XVII_4/page2.htm.

reports on the number and type of requests received, as well as a "Reading Room" with information on frequently-requested FOIA records that are of general interest to the public.[10,11] However, the process of responding to FOIA requests is both time-consuming and labor-intensive for FSIS, and oftentimes expensive for requestors. With the exception of data released through FOIA requests, FSIS does not currently routinely share disaggregated, establishment-specific data with the public.

[10] For more information, please visit: http://www.fsis.usda.gov/wps/portal/footer/policies-and-links/freedom-of-information-act
[11] For more information, please visit: http://www.fsis.usda.gov/wps/portal/footer/policies-and-links/freedom-of-information-act/fsis-electronic-reading-room/

3.0 Presidential Actions and Advisory Committee and National Academy Reviews

Obama Administration and Office of Management and Budget Actions

In 2009, the Obama Administration released a *Memorandum on Transparency and Open Government,*[12] which sought to increase public trust in the government through "a system of transparency, public participation, and collaboration." In the same year, OMB released a *Memorandum for Heads of Executive Departments and Agencies* on the President's Memorandum on Transparency and Open Government - Interagency Collaboration[13], which included steps agencies were required to take to support a goal of creating more openness in government. Among those steps were a requirement that "agencies publish information online and in a form that can be easily retrieved, downloaded, indexed, and searched with tools available on the Internet; use modern technology to share information that can be used by the public without the need for FOIA requests; and post high-value data that have not been previously made available to the public via the Internet or in a downloadable format."[14] In 2011, the Obama Administration released a *Memorandum on Regulatory Compliance,*[15] which stated that "agencies with broad regulatory compliance and administrative enforcement responsibilities…develop a plan to make public information concerning their regulatory compliance and enforcement activities accessible, downloadable, and searchable online." Most recently, in May 2013, OMB released a memorandum requiring agencies to "collect or create information in a way that supports downstream information processing and dissemination activities. This includes using machine readable and open formats, data standards, and common core and extensible metadata for all new information creation and collection efforts. Additionally, it involves agencies building or modernizing information systems in a way that maximizes interoperability and information accessibility, maintains internal and external data asset inventories, enhances information safeguards, and clarifies information management responsibilities."[16]

FSIS National Advisory Committee on Meat and Poultry Inspection (NACMPI) Consultation

As a result of these policy documents, the administrative burden FOIA requests place on FSIS, and the desire to share, in public fashion, data collected as a part of the Agency's activities, FSIS began exploring in 2010 how best to share establishment-specific data with the public. As a first step, FSIS consulted NACMPI. FSIS requested in September 2010 that the subcommittee provide input on which stakeholders should be considered in the release of data, what the prioritization should be for released data, what criteria should be used for determining that prioritization, and what variables, at what time intervals, and at what level of aggregation should information be posted.[17]

[12] "Transparency and Open Government: Memorandum for the Heads of Executive Departments and Agencies". 74 Federal Register 15 (26 January 2009), pp. 4685-4686.
[13] "Memorandum for the Heads of Executive Departments and Agencies: President's Memorandum on Transparency and Open Government - Interagency Collaboration. Memorandum Number: M-09-12. 24 February, 2009. Available at: http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_fy2009/m09-12.pdf.
[14] National Research Council, Committee on a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data. The Potential Consequences of Public Release of Food Safety and Inspection Service Establishment-Specific Data. 2011. Available at: http://www.nap.edu/catalog.php?record_id=13304.
[15] "Memorandum on Regulatory Compliance." 76 Federal Register 14 (21 January 2011), pp. 3825 -3826.
[16] "Memorandum for the Heads of Executive Departments and Agencies: Open Data Policy-Managing Information as an Asset." Memorandum Number: M-13-13. 13 May, 2013. Available at: http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf.
[17] "The National Advisory Committee on Meat and Poultry Inspection Data Collection, Analysis, and Transparency Subcommittee." 2010. Available at: http://www.fsis.usda.gov/wps/wcm/connect/caa395aa5f88467eb20fa010e95cb4db/Data_Subcommittee_Final_Report.pdf?MOD=AJPERES.

In its final report, the NACMPI committee stated that while it generally recommended that FSIS pursue posting of data with a high public health value, it acknowledged that several of the questions raised were beyond the subcommittee's capability, given the limited time available to complete the task. Thus, the committee recommended that "FSIS obtain guidance from NAS [the National Academy of Sciences], NACMCF [the National Advisory Committee on Microbiological Committee for Foods], or other entities with recognized expertise in data management and analysis to improve data accessibility and usefulness for internal as well as external stakeholders."[18]

National Research Council (NRC) Study

As a result of the NACMPI review, FSIS asked the NRC to "conduct a study to examine the potential food-safety benefits and other consequences of making establishment-specific data publicly available on the Internet."[19] The NRC convened the *Committee for a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data*. In examining this question, the Committee reviewed FSIS' current data sharing activities, explored how other government agencies share data with the public, and laid out an approach for FSIS' release of establishment-specific data that considers the benefits and costs of data sharing.

In general, the Committee found that public release of regulatory data is motivated by two broad purposes:

1. The public's "right to know". The committee concluded that "public access to information about the activities of government is basic to democratic governance."[20]
2. "Targeted Transparency". The committee concluded that "disclosure may serve as a means of achieving specific public policy objectives."[21]

In exploring how other government agencies share data with the public, the NRC selected for review a number of regulatory and non-regulatory agencies that currently share data on their activities. The U.S. Department of Labor (DOL), U.S. Environmental Protection Agency (EPA), Food and Drug Administration (FDA), the Centers for Disease Control and Prevention (CDC), USDA's Agricultural Marketing Service (USDA-AMS), and State and local public health agencies all collect and release various forms of data to the public. Information about data sharing within and from these agencies is available in the NRC report.

Based on its review of the design and implementation of data release efforts by other agencies, the NRC committee concluded that FSIS could "benefit from consultation with [*sic]* these agencies and could build on their effective practices when designing a public data release program." Additionally, the Committee conducted a review of the available literature to determine what, if any, effects were reported by these and other agencies when they released establishment-specific data. The committee found that both important potential benefits and costs were reported. Potential benefits include:

- Allowing consumers to make more informed choices

[18] "The National Advisory Committee on Meat and Poultry Inspection Data Collection, Analysis, and Transparency Subcommittee ." 2010. Available at: http://www.fsis.usda.gov/wps/wcm/connect/caa395aa5f88467eb20fa010e95cb4db/Data_Subcommittee_Final_Report.pdf?MOD=AJPERES.

[19] National Research Council, Committee on a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data. The Potential Consequences of Public Release of Food Safety and Inspection Service Establishment-Specific Data. 2011. Available at: http://www.nap.edu/catalog.php?record_id=13304.

[20] National Research Council, Committee on a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data. The Potential Consequences of Public Release of Food Safety and Inspection Service Establishment-Specific Data. 2011. Available at: http://www.nap.edu/catalog.php?record_id=13304.

[21] National Research Council, Committee on a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data. The Potential Consequences of Public Release of Food Safety and Inspection Service Establishment-Specific Data. 2011. Available at: http://www.nap.edu/catalog.php?record_id=13304.

- Motivating firms to improve performance, which may lead to:
    - Incentives to protect brand reputation in food safety or to protect or enhance customer base and profitability
    - Economic pressure to improve food safety
    - Enhanced performance benchmarking
- Providing better insights into strengths and weaknesses of different processing practices which could lead to industry-wide improvements in food safety practices
- Improving the consistency of inspector performance
- Generating research opportunities

Potential costs or unintended consequences include:

- Potential to draw inappropriate conclusions because of data misinterpretation; especially if appropriate context is not provided
- Potential to encourage firms to improve in the reported areas, at the expense of unreported outcomes
- Additional financial commitment for the reporting agency associated with designing and maintaining a useful data disclosure system
- Potential adverse effects on:
    - Inspector performance
    - Brand reputation
    - International trade
- The unintended release of proprietary or confidential information.

On the whole, the NRC Committee concluded that there are "strong arguments supporting public release of establishment-specific FSIS data, especially data that are subject to release under FOIA." To maximize its effectiveness and minimize its potential adverse unintended consequences, FSIS' "data disclosure *should* be guided by a carefully designed information-disclosure strategy."[22]

In addition to the research conducted by the NRC committee, the FSIS DCC workgroup conducted its own in-depth review of federal data sharing procedures and resources to supplement the information provided by the NRC Committee. This review focused on both regulatory and non-regulatory agencies and identified model websites, data sharing portals, and other public data sharing resources.

[22] National Research Council, Committee on a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data. The Potential Consequences of Public Release of Food Safety and Inspection Service Establishment-Specific Data. 2011. Available at: http://www.nap.edu/catalog.php?record_id=13304.

4.0 Criteria for Evaluating FSIS Datasets for Public Posting

In its final report, the NRC *Committee for a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data* stated that "the criteria for choosing which datasets to make public are directly related to the potential users. The many parties that may use the data will use them in different and creative ways that agency planners themselves might not foresee. Although the committee believes that it will be difficult for FSIS to predict the full array of users and uses of the data, it also recognizes the importance of determining the utility of data for different users."

Therefore, FSIS has developed a number of criteria to evaluate potential datasets for public release. An outline of the criteria is provided below, with a more extensive discussion of each criterion provided in the sections below.

Criteria

1. Are the potential data already eligible for release through a formal FOIA request to FSIS? This includes consideration of:
    a. The frequency of past requests for this information via FOIA
2. Does the potential data contain Personally Identifiable Information (PII) for FSIS personnel?
3. What is the potential impact on industry and does the potential data contain imbedded corporate proprietary data?
4. What are the expected personnel and monetary costs to FSIS to release the data?
5. What is the estimated utility of releasing the data? This includes consideration of:
    a. The potential impact to public health
    b. The estimated utility for this information to positively contribute to scientific research and endeavors
    c. The estimated utility of this information for the regulated industry
    d. Whether releasing the data increases FSIS' ability to be transparent to stakeholders
6. What is the potential for misunderstanding/misuse of the data? This includes consideration of:
    a. The completeness of the data
    b. The quality of the data
    c. The complexity of the data
    d. The availability of context for the variables (definitions for each field plus text that describes the methods of data collection, sources of variability, and changes in procedures that affect data consistency)
7. What are the security implications to FSIS and the regulated industry in releasing the data?

FOIA Requests

As mentioned in Section 2, FSIS releases a large volume of disaggregated, establishment-specific data to the public through Freedom of Information Act (FOIA) requests and information that has been requested of the Agency multiple times is oftentimes posted on the FOIA webpage for general public use.[23] To eliminate duplication of effort, FSIS will determine if data being considered for public release has been previously released on the FOIA webpage (or through another venue). If not, the frequency of past FOIA requests for this information will help FSIS determine the relative demand for the data being considered for release.

[23] The Freedom of Information Act 5 USC § 552, As Amended By Public Law 104-231, 110 Stat. 3048. Available at: http://www.justice.gov/oip/foia_updates/Vol_XVII_4/page2.htm.

Personally Identifiable Information (PII) for FSIS Personnel

Candidate datasets for release will be thoroughly examined by multiple FSIS personnel to determine if the dataset contains any PII. Included in, but not limited to FSIS' definition of PII are the names and addresses of, and specific work assignments for, FSIS field inspectors. Any data fields that contain PII will not be released due to the high personnel costs associated with manually redacting PII. It is important to note that, simply because text fields that contain PII will not be released, that does not mean that other information about that data field will not be released. For example, while the text in an NR may not be released due to the presence of PII, information about whether or not a NR was recorded, the date it was issued, and what regulations were cited in relation to the NR will be shared. In the future, FSIS will consider requests to release manually redacted, high-value data fields that potentially contain PII on a case-by-case basis, but only after careful consideration of the potential benefits and consequences of release.

Impact on Industry and Corporate Proprietary Data

Candidate datasets for release will be thoroughly examined by multiple FSIS personnel to determine both if the dataset could have an adverse impact on the industry and if it contains any corporate proprietary information. First, and most importantly, FSIS will not release corporate proprietary data. Second, FSIS will involve the regulated industry in a pre-release of each dataset to ensure utility. FSIS will also seek to ensure that released data are properly explained to decrease the possibility that data will be misused and misunderstood, as described in Section 6.0.

FSIS Personnel and Monetary Costs

FSIS will consider both the Agency's personnel and monetary costs associated with releasing data to the public. Primary among these considerations are the burden of releasing historical FSIS data and the effort required to release certain types of data.

PHIS, which was fully-implemented by FSIS in 2012, replaced several older data collection and management systems used to record inspection data, in particular Performance-Based Inspection System (PBIS). PHIS contains completely redesigned forms for data collection, as well as completely new data structures for storing the data. Analyses that combine data collected under PHIS with data collected in older data systems requires significant contextual knowledge of the changes in both data collection and storage structure. The level of documentation required for datasets from the older data systems is significantly higher than that for PHIS datasets. For this reason, at this time, only FSIS data collected and stored since the full implementation of PHIS will be considered for release. If FSIS resources permit, historical data from older data systems stored in the FSIS Data Warehouse[24] will be considered for release at a later date.

FSIS collects establishment-level data in a variety of formats, including in free text fields. Free text fields are entries into data systems where users are allowed to type any text desired. This is in contrast to other fields where the user must select a specific value from a drop-down list of choices or must enter a date / numeric value. Examples of free text fields include certain fields in inspection records where inspection personnel record their observations, problem descriptions in noncompliance reports, and many fields in Food Safety Assessments (FSAs) where detailed descriptions are recorded. It is not possible to use software to automatically redact these free text entries to remove any personally identifiable or industry

[24] The FSIS Data Warehouse (DW) is the Agency's primary repository for data from its various information systems. Key data from systems such as PHIS, Laboratory Information Management System (LIMS), In-Commerce System (ICS), Supplier Tracking for *E. coli* Positives system (STEPS), and data from FSIS' legacy systems are stored in the DW. The DW serves as the primary source for data across FSIS for use in data analysis and reporting.

proprietary information. To redact manually would require resources beyond the value of the information. For these reasons, no free text data fields will be released.

Estimated Utility of Data Release

FSIS will evaluate the estimated utility of each potential data release to determine which datasets will be most beneficial to the Agency's stakeholders. Given FSIS' mission to protect the public's health and ensure a safe food supply, it is important that FSIS release data that will be used to benefit the public's health and reduce foodborne illness. Therefore, among the factors that FSIS will consider are the potential impact on public health, the regulated industry, and other stakeholders, the potential to positively contribute to scientific research and endeavors, and the impact the data release has on transparency.

Potential for Misunderstanding and/or Misuse of FSIS Data

Each dataset will be evaluated to determine the potential for misunderstanding or misuse of the information. If it is highly likely that the released data could be misinterpreted by the public, the dataset will be reviewed to determine if additional explanatory information or contextual information could reduce the potential for misinterpretation. Additionally, FSIS intends to publish a data dictionary and a user guide for data use, interpretation, and limitations for each data set. If FSIS determines additional information will not alleviate the potential for misunderstanding/misuse, that dataset be removed from consideration for public release.

Additionally, because there is a high likelihood that incomplete data will be misunderstood or misused; FSIS will not release any partial or preliminary datasets. Examples of data that will not be released until finalized include noncompliance reports that are under appeal and sample data for *Salmonella* performance sampling sets that are not final. The Agency intends to release only data that is at least six months old to ensure that any known errors have been corrected. FSIS will establish an internal review process to ensure that datasets are quality-checked and suitable for release.

Finally, FSIS will utilize the established Information Quality Process set forth in the USDA Quality of Information Guidelines[25] for correction of information disseminated by the Agency, if required.[26]

Security Implications

FSIS will not release data that poses significant security implications to either the Agency, FSIS staff, or the regulated industry. Therefore, FSIS will evaluate the security risk posed by of each potential data release to determine which datasets should not be released.

[25] For more information about the USDA Information Quality Guidelines, please visit the following website: http://www.ocio.usda.gov/policy-directives-records-forms/information-quality-activities.

[26] For more information about the USDA Quality Information Guidelines and the FSIS Information Quality Process, please visit the following website: http://www.fsis.usda.gov/wps/portal/footer/policies-and-links/information-quality.

5.0 FSIS Prioritization for Data Release

Using the criteria developed and considering the procedures established by FSIS, the Agency has developed a list of datasets for public release. A sample dataset has been included as Table 1 below to indicate what data fields FSIS intends to release with each dataset.

Datasets preliminarily identified by FSIS for public release include the following:[27]

- *Salmonella* sampling and serotype data for young chickens (HC11_BR)
- *Salmonella* sampling and serotype data for young turkeys (HC11_TU)
- *Salmonella* sampling and serotype data for comminuted poultry (HC01_GC & HC01_GT)
- *Campylobacter* sampling data for young chickens (HC11_BR)
- *Campylobacter* sampling data for young turkeys (HC11_TU)
- *Salmonella* sampling and serotype data for RGB (HC01_GB and MT43S)
- STEC sampling data for Raw Ground Beef (RGB) (MT43)
- STEC sampling data for beef trim (MT55 and MT60)
- STEC follow-up sampling data (MT44, MT52, and MT53)
- STEC sampling data for beef components (MT54)
- *Listeria monocytogenes* and *Salmonella* sampling data for Ready-To-Eat (RTE) (RTE001 and ALLRTE)
- Chemical residue sampling data—routine testing (RM12)
    - Repeat violators v. single violators
- Advanced Meat Recovery (AMR) sampling data (AMR01) and follow-up (FAMR01)

Other data sources to be considered for future release:

- Verification of compliance and noncompliance with FSIS regulations
- Import sampling data
- State sampling data

It is FSIS' intention to follow the process outlined below for releasing datasets:

- FSIS intends to release only one dataset from the Priority List at a time.
- For each dataset, FSIS intends to first release a sample data set with all associated documentation to interested stakeholders for review.
- Once a review of the sample dataset has occurred, and any necessary changes are made, FSIS will release the full dataset on Data.gov.
- Datasets will be released in a format that is compatible for analysis with existing FSIS public datasets, such as the Meat, and Poultry, and Egg Inspection Directory (MPI Directory)[28] and existing FSIS datasets on Data.gov.

[27] When applicable, FSIS will also releasing information about antibiotic resistance, particularly multi-drug resistance, and other information that involves products or samples that have a greater propensity to cause human illness.

[28] For more information about the FSIS Meat, Poultry, and Egg Inspection Directory, please visit the following website: http://www.fsis.usda.gov/wps/portal/fsis/topics/inspection/mpi-directory.

Table 1: Sample *E. coli* O157:H7 dataset for Public Release

Establishment Number	Establishment Name	FormID	Collect Date	Analysis	Sample Result
XXXXX	Company A	100051332	1/3/2012	*E. coli* O157:H7	Negative
XXXXX	Company B	100051347	1/3/2012	*E. coli* O157:H7	Negative
XXXXX	Company C	100051357	1/3/2012	*E. coli* O157:H7	Negative
XXXXX	Company D	100051493	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company E	100051971	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company F	100051643	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company G	100051982	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company H	100051704	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company I	100052187	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company J	100051985	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company K	100051525	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company L	100051426	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company M	100051631	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company N	100052483	1/4/2012	*E. coli* O157:H7	Negative
XXXXX	Company O	100052569	1/5/2012	*E. coli* O157:H7	Negative
XXXXX	Company P	100052153	1/5/2012	*E. coli* O157:H7	Negative
XXXXX	Company Q	100051930	1/5/2012	*E. coli* O157:H7	Negative
XXXXX	Company R	100052246	1/5/2012	*E. coli* O157:H7	Negative

Note: FSIS anticipates using similar variable fields for other pathogen dataset releases, with accommodations made for variables that are important to include for specific pathogens, such as SetID for *Salmonella* microbiological testing data.

6.0 FSIS Data Release Procedures

Given the complexity of releasing FSIS data and the findings and recommendations of both the NACMPI and the NRC committee, FSIS developed a number of data release procedures.

Location of Data

FSIS intends to use Data.gov as a repository and point-of-access for released data. Data.gov is a Federal government-sponsored website designed to "increase the ability of the public to easily find, download, and use datasets that are generated and held by the Federal Government."[29] Data.gov was selected as the point-of-access for FSIS establishment-specific data as it is a centralized platform for secure data storage and downloads. FSIS currently has several datasets on Data.gov, including the Meat, Poultry, and Egg Inspection (MPI) Directory by Establishment Number and Establishment Name.[30]

Data Documentation

Providing proper documentation along with the release of any dataset is critical. This documentation will provide context for the data and guide users on the proper interpretation of the values. No FSIS dataset will be released without specific documentation describing, at a minimum:

1. Dataset overview and explanation
2. The meaning of all fields (database-specific dictionaries)
3. The context under which the data was collected
4. Sources of variability and specificity of methods used
5. Relationship to other released datasets
6. Data use limitations

Availability of this background information for each dataset and the ability to quickly produce the required documentation will weigh in the prioritization of FSIS datasets for release. FSIS will share this documentation with industry stakeholders prior to the release of datasets to ensure that no proprietary information will be compromised and the information to be posted is accurate.

Data Format

It is expected that released data will be analyzed by researchers using analytical software and frequently combined with other data to elicit new insights. For this reason, it is impractical to release data in formats (such as PDF) that cannot be readily imported into data analysis programs. Therefore, all data will initially be released in machine-readable, comma-separated values (CSV) format to facilitate user analysis. In the future, FSIS will likely utilize other technologies, such as application programming interfaces (APIs) to release data.[31] Further, as stated previously, given the challenges, complexity, and the time considerations involved in reviewing data that is captured and stored in a free-text format (such as a text-box), FSIS will not release data from large free-text fields.

Level of Aggregation

To support their recommendation that FSIS release data at the most disaggregated level, the NRC committee stated that "users can always aggregate data for their analytic needs, but they cannot access disaggregate detail from aggregated data." However, when consulting with stakeholders, FSIS found that many potential users are not data analysts and may not possess the skills and technology needed to

[29] "About Data.gov." Data.gov. Available at: http://www.data.gov/about.
[30] For more information, please visit http://catalog.data.gov/dataset/meat-poultry-and-egg-inspection-directory-by-establishment-name and http://catalog.data.gov/dataset/meat-poultry-and-egg-inspection-directory-by-establishment-number.
[31] An Application Programming Interface, or API, is a set of software instructions and standards that allows machine to machine communication.

aggregate raw data. Therefore, FSIS will determine the most appropriate level(s) of aggregation for each data set as a part of the review process.

Data Security

FSIS selected Data.gov as the location for released data for several reasons; one of which was the security Data.gov provides for large datasets. Data.gov maintains a number of data policies that ensure that data posted on the website follow Federal guidelines for data sharing. Specifically, all information available through Data.gov is "in compliance with the required confidentiality, integrity, and availability controls mandated by Federal Information Processing Standard (FIPS) 199 as promulgated by the National Institute of Standards and Technology (NIST) and the associated NIST publications supporting the Certification and Accreditation (C&A) process. Submitting Agencies are required to follow NIST guidelines and OMB guidance (including C&A requirements)."[32] Additional data release consideration specific to Data.gov can be found at http://www.data.gov/data-policy.

Maintenance and Future Updates

FSIS is aware that data may need to be updated on occasion due to major errors or inconsistencies. To ensure proper maintenance of released datasets, FSIS will implement a periodic review of released datasets to ensure accuracy. In this process, FSIS will work with the stakeholder to ensure recorded data is correct and accurate. If needed, new datasets will be loaded to Data.gov to replace outdated datasets.

.

[32] "Data Policy Statements." Data.gov. Available at: http://www.data.gov/data-policy

7.0 FSIS Measurement of Effectiveness of Data Release

Measuring the impact of FSIS releasing establishment-specific data to the public is an important piece of the Agency's data release plan. However, as noted by the NRC committee, it is methodological challenging to establish a causal link or statistical association between FSIS activities and actions, such as posting establishment-specific data, and specific public health outcomes, such as a reduction in foodborne illnesses. Further, as the committee stated, it "recognizes that the United States does not have the data or intervention analysis systems in place that could directly measure the potential public-health (or other) effects of specific activities in the FSIS food-safety programs."[33] Therefore, until such an analysis system is established, FSIS intends to use indirect measures of public health impact to determine the effectiveness of this data release effort. As recommended by the NRC committee, FSIS is considering using the following quantitative metrics to measure effectiveness:

1. Number of FSIS datasets released
2. Number of visits to FSIS Web data locations, including number of web downloads of FSIS data
3. Change in the number of FOIA requests by type of data released
4. Cost-savings to FOIA office, if any, as a result of public posting of data
5. Number of reported and peer-reviewed reports generated using FSIS establishment-specific data.
6. Presentations at professional and invited meetings by Senior management and staff
7. Data shared through FSIS Constituent Updates and internal FSIS meetings with industry and consumer groups

FSIS also intends to use qualitative measures to assess the effectiveness of the data release. These measures include an assessment of how data are interpreted and used by stakeholders.

The FSIS data release workgroup will review these metrics and use them to guide future choices for data release.

[33] National Research Council, Committee on a Study of Food Safety and Other Consequences of Publishing Establishment-Specific Data. The Potential Consequences of Public Release of Food Safety and Inspection Service Establishment-Specific Data. 2011. Available at: http://www.nap.edu/catalog.php?record_id=13304.

8.0 Appendices

TBD