The NCFOODSAFE Project

NCFEDA
North Carolina Foodborne Events Data Integration and Analysis Tool

A New Informatics Tool for Food Safety in North Carolina

May 2011

Authors
Dr. Noel P. Greis
Dr. Monica Nogueira
Center for Logistics and Digital Strategy
Kenan-Flagler Business School

Dr. Pia MacDonald
Dr. Rachel Wilfert
North Carolina Center for Public Health Preparedness
Gillings School of Global Public Health
Prepared for:
Department of Homeland Security
U.S. Department of Homeland Security
Washington, D.C. 20528
www.dhs.gov
Contract No. HSHQDC-08-C-00100

Prepared by:
RTI International–Institute of Homeland Security Solutions
Research Triangle Park, North Carolina

This document is in the public domain and may be used and reprinted without special permission. Citation of the source is appreciated.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.


This report is based on research conducted under the Institute for Homeland Security Solutions (IHSS) under contract to the Department of Homeland Security, Washington, DC. (Contract HSHQDC-08-C-00100). The findings and conclusions in this document are those of the author(s), who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Homeland Security. Therefore, no statement in this article should be construed as an official position of the Department of Homeland Security.
Abstract

A safe and secure food supply is critical to our national security. The recent string of food recalls and foodborne illnesses has created a sense of urgency in addressing gaps in the food safety system, and made safe food a high priority for the nation. Early detection and rapid response are challenges that must be met to minimize the impact of a contamination event—whether unintentional or the result of a terrorist act. Researchers from University of North Carolina’s Center for Logistics and Digital Strategy (CLDS) at Kenan-Flagler Business School and the North Carolina Center for Public Health Preparedness (NCPHP) in the Gillings School of Global Public Health have forged a collaboration to improve food safety in North Carolina.

The NCFOODSAFE project bridges existing gaps in current North Carolina food safety systems by developing a new informatics tool, the North Carolina Foodborne Events Data Integration and Analysis Tool (NCFEDA), that provides situational awareness and intelligence about an intrinsically complex and dynamic process—the detection of and response to a foodborne disease outbreak. The project is informed by an understanding of the information-sharing and communication structures among government agencies and personnel responsible for regulating and overseeing the state’s food safety system and its interplay with other jurisdictions.

NCFEDA was built to demonstrate the potential of situational awareness—created through real-time data fusion, analytics, visualization, and real-time communication—to reduce latency of response to foodborne disease outbreaks by North Carolina public health personnel. The capabilities of NCFEDA were demonstrated using a use case that involves one or more clusters of unknown illness cases presenting symptoms of gastrointestinal problems which serve as a potential indicator of an ongoing foodborne disease outbreak. This project is a first step toward the eventual integration of new capabilities within current North Carolina surveillance and response systems. NCFEDA has been designed to align with current national strategic plans for food safety and its results will serve as a model for similar efforts in other states.
# Table of Contents

Project Motivation and Goals .................................................................................................................... 1  
New Food Safety Stakeholder Model ....................................................................................................... 3  
Surveillance and Response for Foodborne Events ................................................................................... 5  
  Phase One: Surveillance .................................................................................................................... 6  
  Phase Two: Investigation .................................................................................................................... 6  
  Phase Three: Inspection ..................................................................................................................... 7  
  Phase Four: Recall ............................................................................................................................. 7  
North Carolina Food Safety Gap Analysis ................................................................................................ 7  
  Map of Food Safety System for North Carolina ................................................................................ 10  
  Collect Data about North Carolina Food Safety System’s Operational Needs ................................. 12  
  Assess Gaps in North Carolina Food Safety System ....................................................................... 17  
  Identify Opportunities for Reducing Latencies in North Carolina Food Safety System ..................... 21  
Building Situational Awareness—Closing the Gaps ............................................................................... 25  
  Data Integration ................................................................................................................................ 26  
  Visualization ...................................................................................................................................... 26  
  Analytical Tools ................................................................................................................................. 27  
  Real-Time Collaboration ................................................................................................................... 27  
  Simple Example ................................................................................................................................ 27  
Foodborne Events Data Integration and Analysis Tool (NCFEDA) ........................................................ 29  
  Overview ........................................................................................................................................... 29  
  Data Sources .................................................................................................................................... 31  
  Architecture ....................................................................................................................................... 36  
  Triggers ............................................................................................................................................. 38  
  Dashboards ....................................................................................................................................... 39  
  Analytical Tools ................................................................................................................................. 39
List of Figures

20. Scenario 2—Excerpt of Suspected Cluster Information Generated by NCFEDA Analytics Engine ............................................................................................................................................52
22. Excerpt of Resulting Solution Model Generated by NCFEDA Analytics Engine for Scenario 2 ......................................................................................................................................54
23. NCFEDA Login Page ........................................................................................................55
24. Screenshot of NC Events Initial Page of NCFEDA Use Case ........................................57
25. Screenshot of NC Events Page of NCFEDA Use Case Showing A Warning with ELI=3 ..........59
26. Screenshot of NC Events Page of NCFEDA Use Case Showing A Warning with ELI=5 ..........61
27. Screenshot of US Recalls Page of NCFEDA Use Case ................................................62
28. Screenshot of NCDA&CS Complaints Page of NCFEDA Use Case ................................63

List of Tables

1. Food Safety System Data Collection .........................................................................................13
2. North Carolina Food Safety System Gap ..................................................................................17
3. Types of Gaps Identified in NC Food Safety System ...............................................................22
4. Food Safety & Defense Data Collected Per Agency .................................................................33
5. Data Fields of Testing Datasets Obtained from Agencies ....................................................35
Project Motivation and Goals

The recent string of food recalls and foodborne illnesses—from the 2008 jalapenos *Salmonella* Saintpaul contamination to the costly 2009 Peanut Corporation of America *Salmonella* Typhimurium contamination—has made safe food a high priority in the nation and has created a sense of urgency in addressing gaps in the food safety system. In response to this urgency, the Food Safety Modernization Act (FSMA) was signed into law by President Obama on January 4, 2011. This law paves the way for the first major overhaul of U.S. food safety laws since 1938.\(^1\) The FDA previously derived its authority mainly from two current federal statutes: The Federal Food, Drug, and Cosmetic Act\(^2\) (FFDCA) enacted in 1938 and the Public Health Service Act\(^3\) (PHSA) enacted in 1944. The new law had rare bipartisan support. FSMA calls on all stakeholders of the food safety system—public sector agencies, private sector companies, and consumers alike—to assume a new posture in food safety that moves from reaction to foodborne events to prevention of foodborne illness and food contamination.

Recent high-profile contamination events such as the peanut and jalapenos examples have elevated the need to share critical intelligence related to these public health threats. However, the current food safety system is characterized by a proliferation of multiple data systems and tools whose lack of interoperability hinders effective information gathering and timely response. Further, most of the public health and food safety informatics work in the United States—from early detection of food-related outbreaks by local and state health departments to confirmation by the Centers for Disease Control and Prevention (CDC) through “fingerprinting” of pathogenic contaminants—takes place at different local, state, and federal jurisdictional levels. As a result, large gaps exist in our ability to meet the challenge of food safety in the United States under the purview of FSMA.

Bridging these gaps to ensure safe food depends on our ability to collect, interpret, and disseminate electronic and other information across organizational and jurisdictional boundaries. Our food safety system is dynamic and complex—with an array of governmental agencies at different jurisdictional levels charged with regulating and supervising the safety of millions of food products produced by thousands of companies across the globe. These agencies have different responsibilities for food safety, as will be discussed later, and maintain their own internal processes and information systems. Further, each processed food product

---

\(^1\) The FDA Food Safety Modernization Act, H.R. 2751, passed in the House on June 9, 2010, and passed in the Senate with changes by voice vote on December 19, 2010. The act was cleared for the White House on December 21, 2010, when the House of Representatives agreed with the Senate’s changes. The bill was signed into law on January 4, 2011, by President Obama.

\(^2\) 21 U.S.C. § 301 et seq.

\(^3\) 42 U.S.C. § 201 et seq.
contains many ingredients and a single production batch may have products from multiple suppliers—even for a single ingredient. Not only does the total lack of visibility in the food chain make it difficult to trace a contaminated food product to the farm or country where it was produced, but the potential for contamination exists at any point along the supply chain from producer to consumer.

Timely response is an essential factor for successfully containing an outbreak. To meet this challenge our food safety system requires robust surveillance mechanisms, rapid laboratory analysis and reporting, accurate and complete data on patients’ food histories, access to key officials’ contact information, product tracing, and mass inspection and recall capabilities. In North Carolina, food safety agencies have employed lessons learned from a massive 2007 food recall related to botulism in canned meat products produced by the Castleberry Company that resulted in intensified efforts to improve command and control, communication, operations, and planning capabilities. Recognizing the inefficiency of local food retailers to respond to that voluntary recall, state authorities mounted a fast and efficient response that mobilized more than 1,000 people to inspect more than 16,000 sites and to remove more than 35,000 cans of food from shelves at an estimated total cost of $418,000 [1]. Although CDC linked eight cases of botulism poisoning in the country to the Castleberry outbreak, no cases of illness were reported in the state of North Carolina.

Continuous surveillance for early detection of foodborne outbreaks and rapid response to reduce the scale and scope of outbreaks are essential components of timely response for safeguarding our food supply. Our current ability to detect and respond to foodborne outbreaks is hampered by a number of gaps in the food safety system that create latencies in these processes. FSMA provides increased authority and resources for FDA to address many of the existing gaps in our food safety system. The law seeks to bridge some of the biggest gaps in surveillance and response activities by mandating the implementation of new information processes and informatics tools that reduce both the scale and scope of a food contamination event—whether unintentional or intentional. However, although FSMA provides a mandate, it does not promote or specify an approach or toolset, thus leaving open the question how to best reduce latencies.

The goal of the NCFOODSAFE project is to develop a new informatics tool(s) that effectively and efficiently reduces the latencies associated with foodborne disease surveillance and response. We focus on opportunities to reduce latencies associated with tasks such as (1) identification of a possible cluster of foodborne illness; (2) determination of the type of food(s) causing the illness; (3) identification of the particular food product(s) that are the source of contamination and causing the foodborne illness; and (4) recalling all contaminated food product(s) in the food chain including those derivative products that contain the contaminated ingredients.

We used a structured approach to (1) assess current gaps in the current system; (2) develop new information concepts supporting situational awareness that bridge these gaps;
and (3) develop a new informatics tool that implements these new concepts to create situational awareness in the event of a foodborne disease outbreak. The project proceeded as follows: We began by performing a detailed gap analysis of the processes and information flows by which critical information related to the foodborne illness surveillance and response tasks are accomplished in North Carolina. This information was used to develop information requirements in support of situational awareness for food safety surveillance and response. Finally, these requirements guided the design and development a proof-of-concept informatics tool, the North Carolina Foodborne Events Data Integration and Analysis Tool or NCFEDA, which is capable of reducing particular latencies in the detection and management of food contamination and foodborne illness.

NCFEDA is envisioned as an experimental prototype that can serve as an experimental software platform to explore new processes and tools in surveillance and response, and that offers the North Carolina Department of Agriculture & Consumer Services (NCDA&CS) [2], the North Carolina Division of Public Health (NCDPH) [3], and the North Carolina Department of Environment and Natural Resources (NCDENR) [4], a prototype tool that could extend their current capabilities in responding to a food safety event and reduce latency. NCFEDA is not intended to be an operational tool, but rather a prototype for further development depending on feedback from stakeholders.

### New Food Safety Stakeholder Model

The passage of FSMA serves as a milestone in food safety law in the United States. Undergirding this law is a set of assumptions that suggest a new food safety stakeholder model in which the private and public sectors, and the consumer, must work together to meet the challenge of safe food. Although the public sector has traditionally been the guardian of food safety, increasingly private sector enterprises and the consumer will play an important role. The private sector has been given more responsibility for recording and providing (when requested by FDA) information about its processes, suppliers, and customers. Consumers will have more opportunity to provide information to regulatory agencies and private sector enterprises in the event of food contamination.

The food supply chain can be modeled by the systems of these key stakeholders whose unique capabilities and resources make them an indispensable part of the overall food safety system. It is the interoperability and synergy among these stakeholders’ systems that command, control, and sustain the overall food supply system, allowing it to function and perform beyond the capabilities of its individual elements. It is this same interoperability and synergy among stakeholders that complicates the food safety problem—especially the need for trace back and trace forward in the event of a foodborne outbreak.

The food supply chain starts at the farm and encompasses food transportation companies, processing facilities, distributors, retailers, brokers, importers, and governmental
agencies responsible for overseeing and regulating the system—and ends at the consumer’s table. Given the large-scale and distributed nature of the food system, it can be viewed as a "system of systems" whose components are complex, heterogeneous, self-organizing networks of systems that operate independently but are ultimately integrated into a dynamic, evolving “organism” that expertly manages the continuous production, distribution, and sale of food. Bringing these systems together into an efficient and effective food safety network is our country’s challenge—for which FSMA provides critical enabling authority to the FDA.

In particular, under FSMA new responsibilities fall on private sector companies. Food manufacturers will be required to register and to examine their processing systems to identify possible ways that food products can become contaminated and to develop detailed plans to keep that from happening. Companies must share those plans with the FDA, and provide the agency with records, including product test results, showing how effectively the companies can carry them out. The FDA is mandated to work with private sector companies on pilot projects to develop traceability systems that strike a balance between protecting public health and preventing any undue burden to small businesses—the new law exempts small farmers from extensive and expensive traceability and recordkeeping requirements if they primarily sell food directly to consumers or to grocery stores.

Increasingly, the consumer is a key stakeholder in the system. Unfortunately, until now the consumer has had limited direct input into the food safety system. Laboratory reports of cases of foodborne illness typically take many days, or even weeks, to find their way into the food safety system. Progressively more, however, consumer input into the surveillance and response processes is occurring through new channels. “Complaints” hotlines to food retailers and to public agencies provide real-time signals of possible foodborne illness. Consumers also “blog” information important to the surveillance and response processes using social media and other emerging technologies. Harnessing these sources of real-time consumer information is essential to reducing latencies within the new food safety stakeholder model.

Figure 1 presents the new Food Safety Stakeholder Model comprising the food safety system’s four major stakeholders. This stakeholder model is used as the foundation that informs the analysis and prototype development presented in this report:

1. A private sector that controls the production and commercialization of food products, the sale and distribution of potentially contaminated products, and participates in the recall of tainted products.
2. A public health system in charge of surveillance and management of outbreaks of disease caused by food contamination.
3. The governmental agencies pertaining to agricultural activities and the protection of the environment and natural resources that regulate the production of food for human consumption by the agricultural and food manufacturing sectors, oversee the safe use of natural resources and clean environmental and sanitary conditions of establishments offering food services, and monitor and assist in food recall efforts.
Surveillance and Response for Foodborne Events

Maintaining food safety comprises a set of processes that fall broadly into the categories of surveillance and response. How do these surveillance and response processes work in practice? The key tasks associated with these processes are represented by four phases of the food safety communication wheel shown in Figure 2. During the first phase public health officials engage in surveillance activities to determine whether individual cases of foodborne illness are part of a larger outbreak. Laboratory testing to look for common pathogens (e.g., *Salmonella*) is used to confirm that an outbreak has occurred and that a cluster of cases has a common pathogen. Once a common pathogen has been identified and an outbreak has been confirmed, epidemiologists conduct interviews to discover the offending food types (e.g., tomatoes). During the third phase, the specific food products (e.g., Red Hat Tomatoes) and facilities (e.g., Richland Produce Company) are tested and inspected to identify specific product brands or production facilities. Once a source has been located, the difficult task of recalling all contaminated products in the food chain begins. The scale and scope of a food contamination event is directly related to the speed with which these tasks can be performed. Reducing the latencies associated with these events is crucial to saving lives and reducing costs.
Figure 2. Key Tasks in Surveillance and Response

The four phases are summarized below:

Phase One: Surveillance

Determine That an Outbreak Has Occurred

During Phase One public health officials engage in surveillance activities to determine whether cases are part of a larger outbreak. Local public health departments are usually the first to pick up the signals of foodborne disease. These signals may correspond to isolated reports of illness or they may be causally linked and part of a larger outbreak. When state and local officials suspect a set of similar cases, samples are sent to CDC for DNA “fingerprinting” using a system called PulseNet to confirm that the illness is a result of the same pathogen (i.e., Salmonella Saintpaul). PulseNet is a network of national and regional laboratory networks dedicated to tracking foodborne infections worldwide. Using PulseNet, CDC is able to identify similar clusters across states, which is critical because the likelihood of cross-state contamination from a single source is quite high for ingredient-driven contaminations.

Phase Two: Investigation

Determine the Pathogenic (or Other) Origin of the Foodborne Outbreak

Confirmation of the pathogenic source becomes the starting point for investigations by the response teams to determine the specific food types that are responsible for the illness. Once a common pathogen has been identified and an outbreak has been confirmed, epidemiologists
such as CDC’s OutbreakNet team conduct interviews to discover the offending food types (e.g., tomatoes) in Phase Two. For example, after laboratory results confirm a certain specific pathogen strain linking cases to an outbreak cluster, epidemiologists conduct detailed interviews of affected individuals (referred to as “control panels”) to determine their food history and any other relevant details that may be related to their illness. Early investigation efforts may point to certain food types or food products as the cause of the contamination.

Phase Three: Inspection

**Determine the Source of the Foodborne Disease Outbreak and Contamination**

During Phase Three, suspected food products and facilities are tested and inspected to identify specific product brands or production facilities. Results may be reported to FDA and USDA networks such as the Food Emergency Response Network (FERN) and exchanged using the online platform eLEXNET. For example, once a possible food type has been determined to be the cause (c.f. tomatoes), inspectors and other state and local officials test product samples at companies (c.f. Red Hat Tomatoes) that are suspected to have produced the contaminated product. Determining the cause and source of an outbreak can be a time-consuming iterative process during which public health and agricultural experts work together to link a suspected food type with specific products or facilities through laboratory tests for suspected pathogens.

Phase Four: Recall

**Locate and Recall All Contaminated Products**

The time-consuming difficult task of recalling all contaminated products is accomplished in Phase Four. Once a particular food product or production facility has been identified, activities are launched to identify growers, manufacturers, and others that may be linked in the food supply and distribution chain. If contamination occurs at a facility that produces an ingredient, it is important to know all downstream customers who may have used this ingredient in their products. Advisory alerts may be issued by health and consumer service officials warning the public to avoid these specific food products. At the same time, voluntary recalls may be issued for that specific product and any derivative products. Recalls may require time-consuming investigation—to both trace back to the source of the contamination and to trace forward over the product’s supply chain to propagate the recall to all its derivative products already in the food chain.

North Carolina Food Safety Gap Analysis

Numerous delays—or latencies—occur in the surveillance and response processes described in Chapter III. A gap analysis was performed to identify these delays that
compromise our ability to respond to food safety events quickly and efficiently. Broadly, a gap analysis is a tool that helps an organization to compare its actual performance with its potential performance. In this study, the gap analysis focused on the current communication structures and processes among government agencies and organizations responsible for regulating and overseeing North Carolina’s food safety system, and the state’s interplay with other local and federal jurisdictions. The purpose of the gap analysis was to pinpoint existing gaps in these communication structures and processes that could be bridged by better information acquisition, analysis, and visualization capabilities—and that could be implemented in new informatics tools.

The gap analysis was developed to answer the following questions relevant to all stakeholders in the food safety system: (1) Who are the state’s food supply key stakeholders; (2) How do governmental agencies on the state and local levels communicate and cooperate to respond to foodborne disease outbreaks—whether naturally occurring or deliberately initiated; (3) What type of surveillance, response, and outbreak management systems are in place to identify and react to such outbreaks; (4) What data are collected by the state’s main electronic surveillance and response systems; (5) What types of delays are the most disruptive to the state’s food safety system; (6) What causes such delays; (7) What and how can the private sector contribute to ensuring the safety of the food supply chain; and (8) How is the consumer currently engaged in the food safety system?

A brief description of each of the four tasks of the gap analysis follows. Figure 3 provides an overview of the methodology for the gap analysis.

The first task was to review the food safety system in North Carolina for the purpose of understanding the interconnections, interrelationships, and interdependencies among all stakeholders through a comprehensive review of publicly available official regulations, descriptive materials, and legislation relevant to food safety.

The second task was to collect relevant data about the North Carolina food safety system’s operational needs to understand the perspectives of the various stakeholders and to develop a map of stakeholder information flows during foodborne disease outbreaks and food contamination events.

The third task was to analyze these results and to identify observed gaps in the North Carolina food safety system, including gaps in (1) information flow among major partnering stakeholders; (2) data collection efforts and systems; and (3) information sharing and system interconnections across stakeholders.

The fourth task was to identify opportunities for addressing these gaps, including not only organizational and technical gaps but also gaps in existing policy and enforcement mechanisms.
Figure 3. Gap Analysis Methodology

**Task 1. Review Current North Carolina Food Safety System**
- Understand interconnections, interrelationships, and interdependencies among stakeholders
- Identify critical foodborne illness surveillance & response systems
- Identify conceptual systemic challenges and gaps to reach local/state/national goals

**Task 2. Collect Data About North Carolina Food Safety System’s Operational Needs**
- Identify and contact key players at each major food safety subsystem to engage them in the project
- Interview major stakeholders to learn from their experience and expertise on system’s operation
- With stakeholders’ assistance, develop a map of information flow during foodborne events

**Task 3. Assess Gaps in North Carolina Food Safety System**
- Analyze the results obtained from stakeholders interviews and subsequent dialogue
- Consolidate results into a cohesive information flow map for the food safety system
- Identify gaps in: 1) information flow among major partnering stakeholders; 2) data collection; 3) systems’ interconnections

**Task 4. Identify Opportunities for Addressing Gaps in North Carolina Food Safety System**
- Address system formulation gaps
- Address system operational gaps
- Address gaps in existing policy and enforcement mechanisms
A summary of the results of the four tasks follows:

**Map of Food Safety System for North Carolina**

Responsibility for surveillance and response against foodborne illness in North Carolina is assumed by a complex system of local, state, and federal agencies. Figure 4 shows a high-level map depicting the communication pathways connecting agencies overseeing surveillance and response to foodborne disease outbreaks at the state level, and its connections with systems at the national and local levels.

**Figure 4. High-level Graph of Foodborne Disease Communication Flow**
At the federal level, at least 15 different federal agencies are responsible for monitoring and regulating our food safety system. These agencies provide orientation and support for national outbreaks, tracebacks, and recalls, and assist in multistate outbreak investigations. The U.S. Department of Agriculture (USDA), and the FDA and CDC, both part of the U.S. Department of Health and Human Services (DHHS), have central roles in monitoring and regulating the system at the national level.

Among CDC’s foodborne disease surveillance systems, PulseNet is a network of public health laboratories that detect foodborne disease case clusters by performing DNA “fingerprinting” via pulsed-field gel electrophoresis (PFGE) on disease-causing bacteria isolated from humans and from suspected food [5]. The Foodborne Diseases Active Surveillance Network (FoodNet) performs active foodborne disease surveillance and research to better understand the epidemiology of foodborne diseases in the country. FoodNet is a collaborative project between CDC, 10 sites (California, Colorado, Connecticut, Georgia, Maryland, Minnesota, New Mexico, New York, Oregon, and Tennessee) participating in CDC’s Emerging Infectious Program, USDA, and FDA [6].

FDA’s Center for Food Safety and Applied Nutrition (CFSAN) has jurisdiction over all food—except for meat, poultry, and egg products which fall under the jurisdiction of USDA’s Food Safety and Inspection Service (FSIS). Some of the main tasks performed by these two organizations include coordination of the inspection programs for food manufacturing facilities, laboratory testing of food products, and issuing advisory alerts and recalls of food products that are contaminated, adulterated, unsafe, or inappropriate for human consumption. In addition to agencies that regulate food at the federal level, each state has its own administrative network with authority to control and normalize all processes related to its food safety in accordance with national standards.

In North Carolina, the Division of Public Health (DPH), the North Carolina Department of Agriculture and Consumer Services (NCDA&CS), and the Department of Environment and Natural Resources (NCDENR) are responsible for food safety surveillance and response. At DPH, communicable disease surveillance information is gathered in two separate informatics systems—the North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) and the North Carolina Electronic Disease Surveillance System (NC EDSS). These systems, and numerous other health services and programs, are coordinated by DPH to promote and protect the health of all who live in North Carolina [7].

NC DETECT provides statewide early event detection to public health officials and hospital users using data collected daily from several secondary data sources: (a) Carolinas Poison Center (CPC)—statewide and updated hourly; (b) Emergency Departments (ED)—111 civilian hospitals (all but Washington County Hospital) plus eight Department of Veterans Affairs facilities; (c) Pre-hospital Medical Information System (PreMIS)—statewide ambulance run data sent daily but with lag times up to 2 weeks; (d) Urgent Cares (UC)—available for
Mecklenburg County Area only; and (e) NCSU College of Veterinary Medicine Laboratories (CVML)—data available only for Microbiology Lab [8].

NC EDSS is the state’s primary communicable diseases data exchange and reporting system between local health departments (LHDs) and DPH. NC EDSS is currently used by DPH and 86 LHDs [9]. In the future, the system will also be used by eight HIV/STD Regional Offices and the NC Department of Environment and Natural Resources (NCDENR). The state’s communicable diseases data collected by these systems are independently reported to CDC, including those related to food.

Currently, the NC EDSS and NC DETECT systems are not integrated, but access is granted for limited data consultation to local and state public health officials with the required authority. In addition, NC DETECT data are sent to CDC’s BioSense program, and NC EDSS is also part of the national Public Health Information Network (PHIN). Once an emergency room, or laboratory, reports a case to NC EDSS, the case is assigned to LHD nurses in the area for investigation. LHD nurses are also responsible for entering data into NC EDDS related to any cases which are reported to their LHD directly by physicians.

At the local level, cities or counties may have their own organizational structures responsible for public health and safety of the food supply. There is a wide range of capability of LHDs across the state. Typically, the first indications of a food safety problem are revealed at the local level. In such a case, a typical flow of information and communication among agencies across the three levels of jurisdiction proceeds as follows. A person falls ill with foodborne disease and may visit a local healthcare provider (e.g., emergency room at a local hospital or a local physician). Depending on the severity of the illness and suspected diagnosis, the physician may or may not request a clinical laboratory test to confirm the suspected diagnosis. By state mandate, all physicians are required to report cases of foodborne illness, mostly within 24 hours, by filling out and sending Part 1 of the Communicable Disease (CD) Report Forms to the LHD in their area.

Local private and public health laboratories are mandated to report any positive tests for reportable diseases to a public health agency. In North Carolina, only the State Laboratory of Public Health (NCSLPH) and a few private laboratories (e.g., LabCorp) are equipped to send results electronically to the state’s electronic disease surveillance system causing systemic delays. Physicians are the first line of defense against foodborne disease. However, given the high percentage of nonreporting, delayed reporting, and underreporting from physicians and laboratories still reliant on paper-based reporting, local public health electronic surveillance systems can also pick up the signals of foodborne disease incidents.

Collect Data about North Carolina Food Safety System’s Operational Needs

A data gathering strategy was developed to better understand and collect information about the flow of information among stakeholders identified in Task 1 above. Critical stakeholders and information system “owners” in public health and agriculture in North
Carolina were identified and interviewed to understand the existing systems within the state that are responsible for surveillance and management of food contamination events. The gathered data represent a combination of primary and secondary data sources, and interviews with key individuals in state agencies in the North Carolina Division of Public Health (NCDPH) and the North Carolina Department of Agriculture and Consumer Services (NCDA&CS), and other stakeholders in North Carolina’s food safety system.

Formal and informal interviews with representatives from all major stakeholders in the food supply chain in North Carolina were conducted, including a convenience sample of local health departments’ communicable diseases staff. Interview questions pertained to foodborne disease surveillance and reporting systems, outbreak investigation and management, laboratory capabilities, inspections and recall processes, and inter/intra-agency communication pathways.

Findings from interviews with state-level health, agriculture, and environment officials, and private sector representatives, are summarized in a report titled *North Carolina Food Safety System Assessment and Gap Analysis Report*, prepared in February 2010 for the Institute for Homeland Security Solutions. In addition, project team members participated in a number of external activities and courses to gather information about current infrastructures, processes, systems, and protocols related to the management of foodborne disease outbreaks in North Carolina. These findings offer valuable insights into the potential for improvement and the greater needs that must be addressed the development of informatics tools.

Table 1 summarizes the food safety system data collection effort for this study.

**Table 1. Food Safety System Data Collection**

<table>
<thead>
<tr>
<th>DATA COLLECTION METHOD</th>
<th>DESCRIPTION OF DATA &amp; DATA SETS</th>
</tr>
</thead>
</table>
| **Primary Literature Review**  | • Official reports published by national and international organizations, academia, and the U.S. government, including World Health Organization (WHO), Food and Agriculture Organization of the United Nations (FAO), European Science Foundation (ESF), U.S. Congress, U.S. Government Accountability Office (GAO), DHHS, USDA, FDA, Council to Improve Foodborne Disease Outbreak Response (CIFDOR)  
  • Papers published by scientific journals in the following areas: Food Safety, Public Health, Infectious Diseases, Disaster Response, Situation Awareness, Public Health Informatics, Visualization, GIS, Health Geographics |
| **Secondary Literature Research** | • Press releases and recall alerts from federal, state, and local regulatory agencies: CDC, FDA, USDA, NCDA&CS, NCDENR, NCDPH, NCSLPH, LHDs, EHSs, etc.  
  • Newspaper articles  
  • Reports published by research centers and organizations working to improve food safety, including National Center for Food Protection and Defense (NCFPD), National Food Safety & Toxicology Center (NFSTC), FMI, NC Association of Local Health Directors (NCALHD) |
<table>
<thead>
<tr>
<th>DATA COLLECTION METHOD</th>
<th>DESCRIPTION OF DATA &amp; DATA SETS</th>
</tr>
</thead>
</table>
| Training Courses Materials | • *Food Vulnerability Assessment Training Course*: CARVER + SHOCK prepared by DHS and University of Tennessee in Asheville, NC, on January 19-20, 2010  
• *NC EDSS Training Course*, by NCDPH and NC Center for Public Health Preparedness (NCCPHP) in Winston-Salem, NC, September 15-16, 2009  
• *EpiVenture in Outbreak Investigation and Management*, online course developed by CDC, September 2009  
• *Introduction to North Carolina Epidemiology Teams*, prepared by NCCPHP, September 2009  
• *Introduction to the State Public Health Laboratory: NC NCSLPH Training Units 1 and 2*, by NCCPHP, August 2009 |
| Workshops & Webinars Materials | • Workshops:  
  o *Preventing a Recall—Protecting Your Food Business*, by NCDA&CS, September 2009  
  o *2009 Commissioner’s Food Safety Forum*, by NCDA&CS, August 2009  
• Webinars:  
  o *Rapid Recall Exchange Overview for Retailers and Wholesalers* by GS1 and FMI, October 2009  
  o *Data Visualization for Health Surveillance: Current Concepts and New Horizons*, by the International Society for Disease Surveillance (ISDS), September 2009 |
| Empirical Data Collection | • *NC Food Safety Task Force subcommittee* Priority “Wish List” for future systems development received from Cris Harrelson, Food Defense Coordinator of the Food Protection Branch of EHSS/DEH/NCDENR  
• *NC Food Safety Task Force subcommittee* generated a “List of Data Collected by Agency” to be considered for future data integration or use in new tools to improve surveillance, prevention, or emergency response. The list included data from the following regulatory agencies: NCDA&CS-FDP, NCDA&CS-MPID, NCDPH, NCDENR-DEH, LHDs. Received from Cris Harrelson of NCDENR  
• Sample data of NCDENR-regulated facilities from main database table of the NCDENR Best Environmental Technology System (BETS) received from Talytha Moore, Technology Support Analyst from NCDENR;  
• Dataset of all cases of foodborne disease in NC from 1993 to 2008 provided by Jean-Marie Maillard, Communicable Disease Branch Medical Unit Physician, NCDPH on January 2010. Dataset includes 47,325 records extracted from the National Electronic Telecommunications System for Surveillance (NETSS)  
• NC Department of Agriculture database of Turbo EIR Citations received from Brett Weed, Coordinator of the Food and Drug Protection Division of NCDA&CS  
• The NC Preparedness and Emergency Response Research Center’s (NC PERRC) Public Health Regional Surveillance Teams (PHRST) Support and Services survey data received from Anne-Marie Meyer, Research Associate, NCCPHP  
• NC PERRC Surveillance Systems survey data received from Erika Samoff, Research Associate, NC PERRC  
• PHRST Organization Capacity Survey data received from Jennifer Hegle of NCCPHP;  
• NC County Demographic Database provided by Anne-Marie Meyer, NCCPHP |
Table 1. Food Safety System Data Collection (continued)

<table>
<thead>
<tr>
<th>DATA COLLECTION METHOD</th>
<th>DESCRIPTION OF DATA &amp; DATA SETS</th>
</tr>
</thead>
</table>
| **Information About NC Electronic Surveillance Systems and Other Informatics Tools** | • Best Environmental Technology System (BETS)  
• Multi-Hazard Threat Database (MHTD)  
• North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT)  
• North Carolina Electronic Disease Surveillance System (NC EDSS)  
• North Carolina Health Alert Network (NC HAN)  
• General Outbreak Assessment Tool (GOAT) and its exercise package, developed by Public Health Response Surveillance Team (PHRST)  |
| **Unstructured Interviews & Communications with Key Personnel from Major Food Safety Stakeholders** | • *NC Division of Public Health (NCDPH)*. A series of meetings was held with key NCDPH personnel including Megan Davies (NC State Epidemiologist and Chief, Epidemiology Section at NCDPH), David Bergmire-Sweat (Foodborne Disease Epidemiologist, NCDPH), Del Williams (Head, Communicable Disease Surveillance), and Aaron Fleischauer (CDC Career Field Officer at NCDPH). In addition, the project team participated in a number of meetings with individuals responsible for the NC DETECT system including Anna Waller.  
• *NC Department of Agriculture and Consumer Services (NCDA&CS)*. Last year, a series of meetings was held with Brett Weed (Coordinator, NCDA&CS Food and Drug Protection Division) to discuss information flows, processes, and gaps in responding to food contamination events. The team also met with Wendy Campbell (Food Defense Coordinator, NC DACS Food & Drug Protection Division) and other technical staff to better understand the department’s technology platform and current food safety protocols. Several concepts for possible tools to reduce latencies in inspections and recalls have been identified as a result of these meetings.  
• *NC Department of Environmental and Natural Resources (NCDENR)*. In January 2010, we participated in a series of meetings organized by Bart Campbell (Section Chief of the Environmental Health Service Division of NCDENR), met with several members of his staff, including Cris Harrelson. This helped us gain a better understanding of the technological gaps and needs affecting the agency and its ability to identify food-related problems in the facilities regulated by NCDENR.  
• *Private Sector*. Stephen Tracey, Head of Food Safety for Food Lion, a retail grocery company with more than 1,200 stores in 11 Southeast and Mid-Atlantic states and one of the largest employers in NC, was interviewed on February 10, 2010, to gather an overview of Food Lion food safety programs and product recall procedures, and to learn about the challenges facing the private sector in ensuring safety of food supply chain, current industry initiatives, and systems to provide better visibility of the process.  
• *Food Safety Expert*. Pam Jenkins, former Foodborne Disease Epidemiologist, NCDPH, was interviewed on December 2, 2009, to gather information and leverage her extensive experience on management of foodborne disease outbreaks in the state in past years, and her knowledge of the relationship of all the major food safety stakeholders and how these agencies work together. |
Table 1. Food Safety System Data Collection (continued)

<table>
<thead>
<tr>
<th>DATA COLLECTION METHOD</th>
<th>DESCRIPTION OF DATA &amp; DATA SETS</th>
</tr>
</thead>
</table>
| Formal Interviews With Local Public Health Responders | • With assistance from NCDPH officials, eight LHDs were selected for formal interviews. The selection process was guided by a history of recent foodborne disease outbreaks in the counties where the LHDs are located. The LHDs’ Communicable Diseases Nurses were interviewed because of their wide knowledge of and activities in surveillance, investigation, and response tasks during a foodborne disease outbreak.  
• A formal Interview Guide instrument was created to maintain the same level of quality and flow of questions during interviews individually performed by team members over a 2-month period.  
• Interview questions targeting activities performed during foodborne disease outbreak events covered the following topics: LHD programs and plans to respond to outbreaks, interviewee job role and activities, frequency and type of such events in the county, use of existing informatics tools and electronic surveillance, frequency and type of formal and informal communications among staff during an outbreak, regulatory agencies collaborating during an event, perceived gaps in the current procedures and systems, and perceived needs.  
• Per the UNC Institutional Review Board (IRB) exemption received, the identities of all LHDs and interviewees are to remain anonymous. To comply with IRB requirements, the interviews’ transcripts have been de-identified.  
• Seven of the interviews and their transcriptions have been completed and their results analyzed.  
• One last interview remains to be conducted by the end of February because of repeated postponements prompted by complications with the H1N1 outbreak and lack of availability of the interviewees. |
| Group Meetings with Officials of Major Regulatory Agencies | • By invitation of NCDA&CS, a presentation about the NCFOODSAFE Project at the NC Food Safety & Defense Task Force Meeting occurred on January 13, 2010. Representatives from all major state agencies responsible for food safety in North Carolina were in attendance. New contacts with other local agencies were established.  
• Following the NC Food Safety & Defense Task Force Meeting, NCFOODSAFE was invited to make a presentation to a large group of officials and informatics personnel from NCDENR and NCDA&CS on January 28, 2010, and facilitate brainstorming session regarding informatics needs.  
• As a major result of this brainstorming session, the NC Food Safety Task Force Subcommittee was created by Bart Campbell, Section Chief, Environmental Health Service Division, NCDENR, to work with the NCFOODSAFE project team as they identify and develop needed improvements in systems and tools for North Carolina. The subcommittee includes representatives from NCDA&CS-MPID, NCDA&CS-FDP, NCDENR, a county LHD (Wake County Health & Human Services), and NCDPH. |

The variety of methods employed during the data gathering process provided several opportunities to meet with staff of the major stakeholders both individually and in group settings. These meetings helped the team to recognize and understand existing gaps, but also demonstrated evidence of the substantial and growing interest from all stakeholders in moving toward an integrated system that transcends individual agency boundaries and that offers a platform for coordinated and shared response to food safety events. This “somewhat” new
attitude seemed to be fueled by the impact caused by the almost continuous string of nationwide recent recalls that has (a) sickened hundreds of consumers; (b) drastically reduced sales of products suspected of or confirmed to be contaminated—with great losses not only to the implicated manufacturers but to the whole affected industry; (c) resulted in large, costly recall operations for both the regulatory agencies and the private sector; and (d) perhaps most importantly, led to the passage of FSMA.

Assess Gaps in North Carolina Food Safety System

Based on the results of Task 2, a set of systemic gaps were identified as contributing to latencies in the surveillance and response processes described in Chapter III. A description of gaps identified in the foodborne disease surveillance and response system in North Carolina is presented in Table 2.

Table 2. North Carolina Food Safety System Gap

<table>
<thead>
<tr>
<th>GAP TITLE</th>
<th>DESCRIPTION OF GAPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Underreporting by Healthcare Providers to LHD</td>
<td>There is an observed lack of compliance by physicians in completing and submitting Part 1 Communicable Disease (CD) Report Forms to the LHDs. Local area physicians seldom report a foodborne illness case via a phone call. Often physicians will fax a lab result or simply assume that a laboratory report will eventually notify the LHD. However, lab reports do not contain all the necessary data required for an adequate investigation which requires LHD nurses to (1) try to contact the physicians to obtain the missing data; (2) try to complete the forms in another way without the physicians’ help; or (3) perform the investigation without the missing data. All these alternative courses delay the investigation process and strongly compromise the chances of identifying the source of contamination.</td>
</tr>
<tr>
<td>Delays and Underreporting Caused by Breakdown/ Lack of Communication Between LHD and Physicians at Local Healthcare Providers</td>
<td>LHD nurses stated that it can be very difficult to contact physicians informally to obtain more detailed information about a case. This often requires talking to receptionists and nurses who are then supposed to “pass along the message” to the physicians. This circuitous process leads to repeated calls that delay the investigation. In addition, in many cases, the LHD sends educational materials and form letters out to the local healthcare providers to inform and provide them with tools to use. However, this is a one-way process. Once the documents are sent, LHD nurses have no way of knowing how the physicians’ offices deal with them (e.g., if information sheets are corrected as requested). There is no formal channel that allows follow-up on how their efforts are received or what is done with the materials and information sent. This communication gap results not only in cases of illness going unreported, but also creates major delays in the LHD being notified that a case even exists. Often, by the time the LHD is notified of a case, usually through lab results that come in, the case patient may have been ill a few weeks prior. This makes it difficult to even attempt to properly and thoroughly investigate the cases and try to determine whether an outbreak occurred.</td>
</tr>
<tr>
<td>GAP TITLE</td>
<td>DESCRIPTION OF GAP</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Delays Caused by Physicians’ Lack of Knowledge about LHDs’ Authority to Request Patient Information</td>
<td>It is “moderately difficult” to obtain information about a case of foodborne disease from local physicians’ offices. Many of the local providers are hesitant to share case-related information and are unaware that LHDs are authorized to request this information from them. This adds additional delays as the LHDs fax over official paperwork that ascertain their authority to request patient information for reportable diseases.</td>
</tr>
<tr>
<td>Delayed Reporting by Laboratories</td>
<td>Lab results related to foodborne illness are sometimes not relayed to the LHD for weeks. Only a few laboratories (e.g., NCSLPH, LabCorp) currently report electronically through NCEDSS, but many others do not, so delayed receipt of laboratory diagnostic test results are common. This gap also leads to many cases of reportable illnesses, including foodborne, never being communicated to the LHD for further investigation.</td>
</tr>
<tr>
<td>Underreporting by Laboratories</td>
<td>Presumptive positive lab results rarely get entered into NCEDSS, even via Electronic Laboratory Reporting (ELR). There are long delays to receive a presumptive positive lab, so usually the LHD finds out after 4+ weeks the presumptive positive lab result, when the sample is actually typed. Only final pathogen typing/strain identification labs are typically entered into NCEDSS. This is especially a problem with Salmonella, but it also happens with other foodborne and communicable diseases.</td>
</tr>
<tr>
<td>Incomplete Reporting Policy for Laboratories</td>
<td>Labs are only mandated to report positive results to a public health agency. Thus, negative results are not reported in NCEDSS. This generates uncertainty for the CD nurse investigating the case who has to track down the result directly at the lab website, or by phone, to determine if the sample is missing or the test result is negative. This adds an extra step to guarantee that the sample was submitted and tested.</td>
</tr>
<tr>
<td>Local Laboratories’ Inadequate Internal Policy/Procedure</td>
<td>Additional delays on reporting are caused by local labs that follow a batch sampling policy. Under this policy, the lab waits until a number of samples are accumulated before submitting the whole batch to the NCSLPH for testing. Any sample suspect to be linked to an outbreak and submitted as part of a batch has its result further delayed.</td>
</tr>
<tr>
<td>Underreporting by Schools/Daycares</td>
<td>Survey results show that school nurses and daycare centers fail to report to LHDs when abnormal absenteeism rates are noted or an increase in incidence of illness is observed (i.e., children going home early with diarrhea). This causes opportunities for further investigation to identify or rule out a more serious public health problem to be missed.</td>
</tr>
<tr>
<td>Need for Software Visualization Tools</td>
<td>The food supply chain and its stakeholders encompass thousands of facilities and consumers distributed in 100 counties in North Carolina. The volume of data, its dynamic nature and complexity makes it difficult for human assimilation and delays the decision making process. Agencies overseeing the food safety system lack informatics tools that allow visual representation of the data on customized dashboards—and graphical interfaces populated with easy to digest and understand graphs/charts/maps—that help humans to focus on the problem at hand (e.g., locate the focus of an outbreak).</td>
</tr>
</tbody>
</table>
Table 2. North Carolina Food Safety System Gap (continued)

<table>
<thead>
<tr>
<th>GAP TITLE</th>
<th>DESCRIPTION OF GAP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Need for Integrated Tools to Conduct Epidemiological Investigation</strong></td>
<td>Existing electronic syndromic and disease reporting systems do not have the capability to perform some basic epidemiological functions, such as manipulating outbreak-related data to calculate attack rates and odds ratios, or to generate epidemic curves; which are required for conducting epidemiological investigations and identifying foodborne outbreaks. Use of external tools (e.g., EpiInfo) is also hindered by the lack of adequate capabilities to export data from NCEDDS in any functional way. Several LHDs continue to use their own forms and processes for collecting food histories from patients because NCEDSS question packages do not allow collecting as detailed a food history as LHD staff feel is necessary to properly investigate a foodborne-related event.</td>
</tr>
<tr>
<td><strong>Need for Real-Time Communications Integrator Tool to Connect All Stakeholders</strong></td>
<td>Information exchange among all parties is important for surveillance and response to foodborne disease. Public health stakeholders consider current formal case reporting systems in North Carolina to be generally efficient and sufficient, but information flows mostly in one direction. Agencies in the agriculture and natural resources fields have different reporting needs (e.g., report on whole companies and facilities sanitation status, not on individual consumers’ health problems). Today staff on stakeholders’ organizations relies on informal contact (e.g., phone) and personal relationships to obtain information that they need to conduct an investigation. Time and effort are wasted on redundant parallel investigative procedures and not enough cross-fertilization among parties exists. There is no central common tool that allows all parties to communicate their efforts, data needs, planned actions, etc.</td>
</tr>
<tr>
<td><strong>Need for Outbreak Management Tool</strong></td>
<td>Public health officials have identified the need for an outbreak management and investigation tool. This coupled with the need for an outbreak command and control room located at a central place was also stressed by several stakeholders’ representatives. In addition, a need for better coordination between LHD and local/county Environmental Health Services (EHS) personnel and how they can share NCEDSS information while researching the same illness cases was identified. It may be advantageous to the investigation if both teams can meet or make initial contact with patients together at the same time. Otherwise, need for corroborating the information by the other team later generates redundancy and slows down the process. If information pertinent to the outbreak investigation happens to be collected by other individuals investigating from a different role then it should be shared in real time. Occasionally individuals’ complaint calls originate an investigation about an outbreak or disease occurring locally, even before any reports from physicians or labs have been received. These calls could also support various surveillance and response tasks. There are many complaint hotlines available to the general population including those provided by LHDs, DPH, EH departments, NCDA&amp;CS, NCDENR, and supermarkets. Gathering all complaint calls data from different agencies as part of a centralized “situational awareness” tool seems to be an opportunity not used yet, although its possible value was cited by several organizations interviewed.</td>
</tr>
<tr>
<td>GAP TITLE</td>
<td>DESCRIPTION OF GAP</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Delays on Recall Process Caused by Lack of Suppliers’ Distribution List</td>
<td>Once the identity of the contamination product has been determined, inspectors and agents of the NCDA&amp;CS and NCDENR rely on the suppliers’ distribution list(s) from the product’s manufacturer/distributor to identify which retailers/other distributors/facilities must be informed of the recall and visited to guarantee that the tainted product is recalled and destroyed. Although food manufacturers are mandated by law to keep records of their suppliers to allow one step back and forward tracing, lack of visibility in the supply chain continues to be a major factor that disrupts, obstructs, and delays the recall process. Implementing processes that allow track and trace of suppliers is a particular challenge for certain food industries (e.g., seafood) whereas shrimp batches are assembled based on size of the shrimp and result in mixing product from many suppliers.</td>
</tr>
<tr>
<td>Confounding Factor that Delays Investigation of the Source of Contamination</td>
<td>A large number of outbreaks are linked to tainted ingredients that are used in the production of many different food products on grocery shelves. Determining the source of contamination in these cases can be delayed because it is difficult (under current systems and regulations) to both trace the offending ingredient up the chain to locate its source and down the chain to identify other products that may have been affected by the same ingredient.</td>
</tr>
<tr>
<td>Need for Real-Time, Risk-Based, Intelligent Tool for Inspections’ Scheduling</td>
<td>An important step in protecting consumers is the inspection of facilities that produce, manufacture, store, handle, transport, prepare, serve, or sell food, to ensure that they follow existing food safety regulations and procedures. However, there are thousands of such food facilities in North Carolina, but only a very small number of food regulatory inspectors to perform routine inspections. A consequence of limited resources and the high volume of inspections needed is that it leads to fixed inspection schedules—where the selection of the facility to be inspected is dictated by the time interval between inspection visits, instead of the risk posed by a particular product or plant (i.e., likelihood of failure). This approach may contribute to weaknesses in the food safety system. As an alternative to time-based inspections, the risk-based approach gives higher priority for inspection of facilities that handle food items which are more susceptible to contamination problems (i.e., “high probability”) and more dangerous to health if processed inadequately (i.e., “high consequence”) than to those that handle food items with high probability but low consequences if improperly manipulated. Risk analysis allows for scheduling inspections based on risk and a better use of resources, but is not without flaws (e.g., lack of enough data may make the calculation of risk inaccurate or impossible). However, given adequate levels of data, the application of this strategy in real time coupled with other intelligent methods would optimize inspections’ effectiveness and the reliability of the food safety system, among other benefits.</td>
</tr>
</tbody>
</table>
Table 2. North Carolina Food Safety System Gap (continued)

<table>
<thead>
<tr>
<th>GAP TITLE</th>
<th>DESCRIPTION OF GAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for Recall Management Tool</td>
<td>Lack of a (near) real-time recall management tool is a contributing factor to delays in completing a recall. If the product recalled is an ingredient (e.g., peanut butter) and affects many different food categories (e.g., soup, ice cream, pet food) then the recall becomes a major operation which is much more extensive and time consuming. During “recall effectiveness checks” NCDA&amp;CS inspectors must verify that all products have been recalled by communicating with or visiting facilities appearing on the products’ distribution list. Data on the recall process are also collected (i.e., was the recall notice received? Was the recall completed and the product handled as instructed by the notice? Was the product distributed to other consignees and have they being notified?). However, retailers who buy from big warehouses or small convenience stores do not appear on the distribution list and must be tracked down. Efficiently managing the collection and recording of this information and optimizing inspectors’ recall assignments are essential tasks for a successful recall process.</td>
</tr>
<tr>
<td>Data Silos</td>
<td>Major food supply chain stakeholders collect and maintain data about the system that is fragmented across several databases and “owned” by different stakeholders with very limited or no mechanisms for data sharing. Lack of integration exists not only between governmental agencies and private sector, but also among stakeholders in the same sphere (e.g., different governmental agencies—NCDA&amp;CS and NCDENR), agencies at same jurisdictional level—LHD and local EHS offices, and even between systems of a same organization. For example, NCEDSS and NCDETECT are not integrated. Data silos complicate surveillance of and response to foodborne disease outbreaks by diminishing the window of opportunity to identify illness cases, the source of contamination, and tainted product locations.</td>
</tr>
</tbody>
</table>

Identify Opportunities for Reducing Latencies in North Carolina Food Safety System

To assist in narrowing down the factors to be considered in the development of a new concept for reducing latencies of surveillance and response in North Carolina’s food safety system, we classified the gaps identified by our analysis into three categories: (1) policy formulation; (2) operational systems; and (3) existing policy and enforcement mechanisms. The outcome of this classification is shown in Table 3 where each column corresponds to a gap category and each row attempts to show the correlation among gaps across the three categories.

For example, in the first row, we see that lack of electronic medical records (EMR) accessible/ available to LHDs may be linked to delayed and underreporting of illness cases to LHDs, which constitutes one of the biggest obstacles to early detection of outbreak events. Underreporting or lack of reporting of cases of notifiable illness is also a significant cause of delay in identifying a foodborne and communicable disease outbreak. Underreporting occurs at different stages of the healthcare process. Several stakeholders commonly fail to comply with the state’s reporting mandate: (a) physicians; (b) laboratories; and (c) schools/daycare facilities. Inadequate laboratory batch processes and policies also contribute to this problem.
### Table 3. Types of Gaps Identified in NC Food Safety System

<table>
<thead>
<tr>
<th>POLICY FORMULATION</th>
<th>OPERATIONAL SYSTEMS</th>
<th>EXISTING POLICY &amp; ENFORCEMENT MECHANISMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>Lack of electronic medical records directly linked to/accessible by LHDs</td>
<td>Delays and underreporting by physicians to LHD</td>
</tr>
<tr>
<td>Incomplete reporting policy for laboratories</td>
<td>Lack of electronic laboratory records directly linked to/accessible by LHDs</td>
<td>Underreporting by laboratories</td>
</tr>
<tr>
<td>Local laboratories’ inadequate internal policy/procedure</td>
<td>Lack of real-time tool for informal communication among stakeholders</td>
<td>Physicians’ lack of knowledge about LHDs’ authority to request patient information</td>
</tr>
<tr>
<td>NA</td>
<td>Incomplete/inaccurate consumer complaint data collection</td>
<td>Underreporting by schools/daycares</td>
</tr>
<tr>
<td>NA</td>
<td>Data silos: Lack of data integration among stakeholders’ systems</td>
<td>Delays in recall process caused by lack of suppliers’ distribution list caused by lack of visibility in the food supply chain</td>
</tr>
<tr>
<td>NA</td>
<td>Lack of analytical tools to assist in:</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>• epidemiological investigation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• outbreak management</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• recall management</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• visualization</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• risk-based inspection scheduling</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• inspectors recall assignments</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• supplier food safety management tool</td>
<td></td>
</tr>
</tbody>
</table>

NA = not applicable.

National levels of underreporting calculated by CDC experts are alarming: 40-fold for *Salmonella* and 20-fold for *E. coli* O157:H7. A much lower multiplier of 2 is arbitrarily used for pathogens that cause more severe illness (e.g., *Listeria* and *Clostridium botulinum*) [10].

Similarly, in the second row, underreporting by laboratories may be associated with incomplete reporting policy for labs and lack of electronic lab reporting (ELR), thereby delaying the identification of outbreak-related events. **Incomplete or inaccurate food history** data collection from patients linked to an outbreak hinders the investigation of potential food contaminant sources. This can be a result of the latencies introduced by laboratory sample processing and reporting and lack of reporting by healthcare providers—as time passes patients’ ability to correctly recall what food they had eaten before falling ill, or where the food was bought, may diminish. There is also a need for training/recording of the “initial complaints,” and a need to establish a formal “intake complaint system” where an operator could interview the caller and input the data directly to an electronic system, or where calls could be recorded for future retrieval.
More complete policies would require labs to report negative, and positive, results. Interviewees mentioned that lack of this information often adds an extra step for them, because sometimes they are not sure whether the test result was negative or whether the sample was not tested. Granting LHDs remote access to the EMRs and laboratory test results for all patients under their jurisdiction is a factor to consider to improving surveillance. Access to EMRs is a reality today in only a few LHDs in North Carolina with recognizable positive results. Time and effort savings in accessing data from patients seen by local physicians or healthcare providers can be even greater if demographic data are also made available through this system (e.g., nurses can make faster contact with patients to schedule interviews). ELR capability has been deployed at the NCSLPH and a few private laboratories (e.g., LabCorp). NCDPH is currently in negotiations to have it extended to other private labs. Given the lack of compliance in reporting by physicians, this is an important factor that will improve foodborne and communicable disease surveillance in North Carolina. Thus, the initial factor emerging from the first two rows is the increasing need for electronically accessible data by public health officials.

Better mechanisms for informal communication appearing in the third row is another contributing element. Our data show that additional communication among major stakeholders is viewed as “extremely necessary.” One can argue that lack of communication between local labs and LHDs may contribute to the lack of responsiveness by laboratories to LHD needs during periods of high demand for laboratory testing such as during outbreak events (because of an inability to prioritize tests). Real-time instant messaging to share updates on a needs basis seems appropriate given the need to be quickly actionable in managing foodborne and communicable disease outbreaks. A database or online directory of contact information for all local physicians per county or LHD area was mentioned as a possible helpful tool to speed up investigations. It would provide LHD nurses easy access to physicians’ contact information and allow them to contact physicians directly, as opposed to going through various intermediaries. Physicians’ lack of knowledge about LHDs’ authority to request patient information is another sign of problematic communication among stakeholders. Our study shows that LHD staff are often frustrated by the lack of response from local physicians and healthcare providers in complying with state mandates for notifiable diseases reporting, even after many efforts to educate and motivate them. This has led to a generalized attitude of “disbelief” and low expectations for finding a solution in the near future.

In the fourth row, the first two factors identified above seem to merge when one examines another important source of information about the occurrence of foodborne disease locally—consumer complaint hotlines. Both regulatory agencies and the public sector maintain separate hotlines so consumers can actively contribute to early detection of outbreaks by calling and reporting an illness and providing details about where they ate or bought a food product that they suspect is contaminated or not appropriate for human consumption. Accurate, complete collection and recording of patient, or consumer, “initial complaints” may
be needed to facilitate the investigation and to increase the probability that the source of contamination will be speedily found. Evidence shows that there may be need for better training/recording of these complaints, and a need to establish a better/formal “intake complaint system” where an operator could interview the caller and input the data directly to an electronic system or where calls could be recorded for future retrieval. Moreover, integrating school and daycare staff information about suspect cases of foodborne disease may help improve underreporting in these settings.

Integration of stakeholders’ data, appearing in the fifth row, is a factor in reducing latencies in the North Carolina food safety system because it is a structural barrier to achieving a unified, comprehensive view of the system (i.e., data integration must be implemented to reach high-level situation awareness). Lack of data integration across agencies, and between governmental organizations and the private sector, delays response to foodborne illness outbreaks. It is essential to find policy and operational mechanisms that allow “fusion” of important data elements to reduce latencies. Each agency monitors and controls a portion of information that is linked to a known current outbreak event or that may help identify points of high risk for a future outbreak. The private sector is responsible for the production, processing, distribution, and commercialization of food and, thus, “owns” both the food supply chain and the data that may allow tracing the origin of a product. Privacy concerns on one hand, and competitive advantage requirements on the other, create obstacles to data sharing between public agencies and the private sector. The need for data integration will be explored in later sections of this report, but it is important to stress that the relationship between the private sector that “owns” the food supply chain and the regulatory agencies that monitor and enforce the product recall process may contribute to delayed response during outbreak events.

Another factor for improving the overall efficiency and reliability of the food safety system is the implementation and deployment of analytical tools to continuously and efficiently examine the different datasets to uncover possible links among data elements that may indicate an underlying event warranting closer investigation by food safety authorities or private sector representatives. We have identified the possible need for such capability across multiple regulatory agencies and verified that these stakeholders are well aware of such needs and are searching for solutions. Such as analytical tool could include an “intelligent” layer that processes data to generate relevant information to users. Functions identified by this study include (a) helping epidemiologists determine the foods with the highest likelihood of being the contamination source; (b) linking seemingly unrelated consumers complaints from distinct counties to uncover a possible outbreak; (c) uncovering spatial-temporal correlations between an increase in the incidence of illnesses in a neighborhood and a food processing/services facility in the area; (d) linking new trends in inspections violations to the risk-based inspections scheduling; and (e) associating inspector skill level and training to repeated inspection violations and recall assignments.
Another desirable function is the addition of data visualization or a dashboard that, when coupled to underlying analytical tools, will filter and present the most useful and relevant information to users in a format that is easy to assimilate and will help with the decision-making process. The ability to generate epidemic curves and line listings appears to be of particular interest to LHD communicable disease nurses. Lack of analytical and visualization tools to assist food safety authorities in the analysis of potential/suspect cases and to provide situational awareness to help identify relationships linking isolated events with current or nationwide trends is another desirable feature. The quantity and scope of data involved in the food supply chain, and its dynamic nature and complexity, make it difficult for human assimilation which delays the decision-making process. Agencies’ lack informatics tools that facilitate longitudinal, spatial, and qualitative investigations and that allow visual representation of the data on customized dashboards—graphical interfaces populated with easy to digest and understand graphs/charts/maps—to help humans to focus on the problem at hand (e.g., locate the source of an outbreak).

The implementation and wide adoption of a manageable track-and-trace process for the food supply chain and new real-time, or near real-time, systems that connect suppliers and retailers can provide the ability to know what is happening—receive results of microbiological work, information about contaminated products, etc.—at the moment it happens. This capability is viewed as pivotal in improving the food safety system. Lack of visibility in the food supply chain was identified as the single most important factor that delays the recall of contaminated products. Even though the new law requires total visibility of the food supply chain, no specific technology or approach has been promoted. Recent outbreaks have revealed that many food manufacturers are unable to trace food products even one echelon forward or backward within their own supply chain. However, manufacturers are the first to be informed if one of their products tests positive for a foodborne pathogen and one of their first actions should be to inform their suppliers if a product recall is warranted. Again, the private sector is the guardian of vital information, and only with its cooperation can the recall process happen in an efficient and speedy fashion.

Building Situational Awareness—Closing the Gaps

The results of the gap analysis in Chapter IV show an overarching need for capabilities that build situational awareness across stakeholders in the food safety system. Situational awareness is defined by Endsley as the “perception of the elements in the environment within a volume of time and space, comprehension of their meaning and the projection of their status in the near future” [11][12]. Theoretical frameworks for situational awareness are based on an understanding of cognitive processes of the human mind for decision-making. The most common theoretical framework is provided by Endsley and defines three stages of situational awareness formation: perception, sensemaking or comprehension, and prediction or projection.
• **Level 1 PERCEPTION:** Perceive the status, attributes, and dynamics of relevant elements in the environment. For stakeholders in the food safety system, this refers to the process of monitoring the environment to become aware of incoming consumer complaints, lab reports of CDS, announcements of recalls, updates from state health systems such as NC DETECT, etc.

• **Level 2 SENSEMAKING:** Synthesize the “unconnected” elements through pattern recognition, correlation, interpretation, and so forth. For stakeholders in the food safety system, this refers to the process of seeing patterns across types of complaints and confirmed laboratory reports, both within and across counties/states, and linking those events to other events such as recent recalls of contaminated products or outbreaks of similar disease.

• **Level 3 PROJECTION:** Project the current situation in space and time through knowledge of the various elements in the environment and their dynamic relationships. For stakeholders in the food safety system, this means comparing the current situation with historical trends or with external related trends to determine whether the current observed event is an outbreak and what its future state might be.

We identified four essential capabilities that contribute to situational awareness in food safety: (1) **data integration**; (2) **visualization**; (3) **analytical tools**; and (4) **real-time collaboration**. These four capabilities support an operational environment necessary for understanding, evaluating, and responding to foodborne outbreak events in an effective and timely manner. In continuously and rapidly changing environments such as public health, capabilities that support situational awareness maximize results of operational procedures, improve team collaboration, and enable better informed decision-making.

Relevance of each capability to our end goal of reducing latencies in surveillance and response to foodborne disease outbreaks is described briefly below:

**Data Integration**

Fusing data from all major food safety stakeholders can offer a more complete and clear picture of ongoing (i.e., near real-time) event. To create situational awareness an informatics tool must provide a coherent representation of those data elements that are relevant to respective food safety stakeholders and that are essential to perceiving the status, attributes, and dynamics of any ongoing event. Currently, each major food safety stakeholder has only partial knowledge of what is happening based on the stakeholder’s limits of responsibility and authority. Combining relevant information across all relevant food safety stakeholders into a single shared view (i.e., common operational picture) will create a more complete representation of present conditions that may allow faster recognition of existing problems and generate new knowledge that will contribute to latency reductions.

**Visualization**

A visualization tool not only provides a visual representation of data that is more easily interpreted, but can also be used as a problem-solving technique. Trying to answer questions
by examining large numerical tables or spreadsheets is typically more difficult and time-
consuming than allowing a user to process the same data presented in graphs/maps/charts. 
Exploring different views of the same data facilitates analytical reasoning by taking advantage 
of human capabilities to process images. Benefits obtained from fusing diverse data sources 
can be augmented by adding visual analysis capabilities to the system.

Analytical Tools

Analytics are broadly defined as a set of tools based on logic and statistics that are used 
to support decision-making. In food safety, analytical tools can discover disease or exposure 
patterns that require further epidemiological investigation and that will, as a result, speed up 
the process of identifying possible sources of contamination. For example, analytical tools can 
generate clusters based on similar foods consumed, places visited, or other common elements 
among data records that may help point out the source of contamination or uncover a totally 
new, still unreported, existing problem. Such tools can also assist in reducing latencies in the 
recall process by making the recall and effectiveness checks more efficient. Analytical tools 
can also be used to assess the likelihood of the emergence of a food safety event from fused 
data that can then be used to guide response.

Real-Time Collaboration

The need for better mechanisms for informal communication among stakeholders, as 
indicated by the gap analysis, is multifold and (1) calls for a communication vehicle that 
enables exchange of information between participants; (2) offers 24/7 access; and (3) entails 
keeping a comprehensive roster of responders and public health officials at the state level 
including direct contact information and location, and an analogous roster of local healthcare 
providers’ representatives and physicians at the local level. Such capability enables anytime, 
anywhere collaboration and exchange of ideas and information.

These four capabilities create situational awareness for real-time command and control 
necessary, as previously stated, for understanding, evaluating, and responding to outbreak 
events in an effective and timely manner. The NCFEDA concept presented in this document 
incorporates the first three capabilities. Real-time collaboration was not included in the 
NCFEDA framework because it may require acquiring, implementing, or modifying existing 
electronic systems used by stakeholders, a task that goes beyond the scope and scale of the 
NCFOODSAFE project.

Simple Example

To illustrate how situational awareness can be achieved in the NCFEDA tool, and how 
different data sources may be used, consider the following hypothetical and idealized scenario 
where data integration and analysis capabilities in the identification of an emerging food safety 
event, combined with visualization tools, enable coordinated response.
In this scenario, a supermarket chain in North Carolina reports to the FDA’s Reportable Foods Registry that potentially contaminated food has been received from a supplier. Simultaneously, cases of possible food poisoning have been recorded in a few North Carolina county emergency rooms (e.g., Wake, Durham, and Mecklenburg). The incidents are not concentrated enough or defined enough to provide an automatic alert in the existing syndromic surveillance system, NC DETECT. However, in this idealized scenario, relevant data fields from the supermarket report to the FDA are shared through NCFEDA with state agencies.

For example, the list of counties in North Carolina that received affected lots is shared with NCDPH and other public health officials, which confirms possible associated cases in NC DETECT. NCFEDA will make additional pertinent data fields available to other agencies. NCDA&CS will be informed about the numbers of affected lots and the stores these lots were sold to in Wake, Durham, and Mecklenburg counties, and 10 other counties all across the state. Supplementary information regarding institutions and restaurants that have purchased the affected products may be provided by the supermarket chain directly to NCFEDA. In this case, NCFEDA will provide NCDENR with a list of all institutions/facilities involved.

Information shared simultaneously with NCDA&CS launches inspections/recalls to these stores. However, NCDA&CS has limited resources and cannot launch inspections at all locations at once. Where is the risk highest? By combining information about the volume of products sold and possible cases being reported in NC DETECT, an inspection/recall scheme can be scheduled and implemented.

If the product is typically sold to nursing homes, schools, and other establishments that serve food to larger groups of people, NCDENR is called upon to revisit/inspect these high-risk facilities to perform a recall effectiveness check, thus confirming that contaminated products have been recalled and disposed of correctly. Again, NCDENR inspector assignments are calculated based on where the risk of contamination is highest. Several factors may be taken into consideration for this computation including age of population at highest risk, facility closeness to the supermarket chain stores, facility purchase frequency from the chain in the past, etc.

In this example, simple sharing of information and active mapping can provide situational awareness after the event is triggered by the report to the Reportable Foods Registry. For example, situational awareness can reduce latency of surveillance and response as follows:

1. Reports from the Reportable Food Registry enable NCDPH to confirm a contamination event that may have gone undetected because the signal was too “diffuse” to raise an alert. Recalls are started earlier.

2. Information from the supermarket chain that other counties received lots of the suspect product enables targeted inspections. Most likely the supermarket chain stores are inspected first.
3. Information from the supermarket about food service facilities that are frequent customers of the chain will increase the probability of these establishments being inspected before others.

4. NCDPH is able to alert hospitals in affected counties. The information about affected products helps select interview questions, for example ensuring that the interviewer asks about recent consumption of the suspected product or its derivatives. It also enables warning to other family members to remove the offending product remaining in their kitchens so no one else becomes ill, and guides the hospitals/emergency rooms “intake” interview process to help to ensure appropriate treatment to individuals who may come in with similar symptoms.

---

### Foodborne Events Data Integration and Analysis Tool (NCFEDA)

#### Overview

NCFEDA presents a new concept whose goal is to provide North Carolina state or local government agencies and personnel responsible for regulating and overseeing the state’s food safety system with a new framework to increase their situational awareness. As discussed in the last chapter, NCFEDA’s capabilities include: (1) data integration; (2) visualization; (3) analytical tools; and (4) real-time collaboration.

The NCFEDA framework shown in Figure 5 provides an integrated platform for the four major food safety stakeholders in North Carolina: (1) Public Health; (2) Agriculture and Environment & Natural Resources; (3) Private Sector; and (4) Consumer. As shown in the figure, NCFEDA is a “natural” progression from our conceptual North Carolina Foodborne Diseases Communication Wheel that interconnects responding agencies through electronic systems for data exchange and other informal communication channels.

NCFEDA’s role is to bridge information silos across state agencies through software tools and systems while adding new information elements from the private sector and consumers. In addition, NCFEDA’s use of visualization and analytical tools will explore relationships across these information elements in a way that creates new knowledge that can help surveillance and response teams reduce latencies.

As shown in Figure 5, NCFEDA includes all entities represented in the new food safety stakeholder model. For the public sector stakeholders (Public Health and Agriculture, Environment and Natural Resources), we have identified the corresponding response agencies and their departments or divisions, and other organizations and electronic systems, which participate in surveillance of and response to foodborne disease outbreaks. For the private sector and consumer stakeholders, we have indicated the entities corresponding to companies and institutions that hold information about food-related events needed to ensure food safety.
Public Health. The main organizations and agents of public health include (a) North Carolina Department of Public Health (NCDPH); (b) North Carolina State Laboratory of Public Health (NCSLPH); (c) local health departments (LHDs); (d) hospitals and similar healthcare providing facilities; (e) physicians; (f) clinical laboratories; (g) North Carolina Health Alert Network (NC HAN); (h) North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT); and (i) North Carolina Electronic Disease Surveillance System (NC EDSS).
Agriculture and Environment & Natural Resources. The main organizations and agents of agriculture, environment and natural resources include (a) North Carolina Department of Agriculture and Consumer Services (NCDA&CS); (b) NCDA&CS Meat and Poultry Inspection Division; (c) NCDA&CS Food and Drug Protection Division; (d) NCDA&CS laboratories; (e) NCDA&CS Multi-Hazard Threat Database (MHTD); (f) NCDENR; (g) NCDENR Division of Environmental Health (DEH); (h) Environmental Health Services Section (EHSS); (i) NCDENR Dairy and Food Protection Division; and (j) NCDENR Best Environmental Technology System (BETS).

Private Sector. The private sector comprises companies that manufacture, process, distribute, transport, broker, or commercialize food through retail (e.g., supermarkets and restaurants, or wholesale stores).

Consumer. With respect to consumers, we consider both individuals and institutions. Institutions that serve large populations are of interest including (a) schools; (b) day cares; and (c) nursing homes. We also consider consumer complaint hotlines and new emerging social networking systems that people are increasingly using to communicate ongoing events (e.g., Twitter and Facebook), or even bulletin boards maintained by advocacy groups.

Data Sources

NCFEDA brings together data distributed across multiple systems and controlled by the state’s major food safety stakeholders. One central principle of the NCFEDA concept is use only non-proprietary data elements that can be shared across stakeholders and, thus, avoid overstepping an entity’s jurisdiction, or require access to protected information.

Figure 6 provides a view of other current and potential data sources that are available to NCFEDA as a function of timeliness.

Data available at near real time has the highest potential to alert responders to emerging situations; data updated on a weekly basis are used for confirming outbreaks and are essential to outbreak investigations and recall procedures in assessing recall effectiveness; data provided monthly assist in identifying and linking sparse outbreak events and scheduling risk-based inspections; while yearly data are most useful for routine trend analysis and new policy formulation to support prevention of new events. NCFEDA-relevant data sources are described below.

Emerging Situation Alerts. To achieve situational awareness of emerging events, NCFEDA maximizes use of nontraditional real-time information, as available, that is relevant to foodborne disease surveillance and response. In particular, NCFEDA makes use of consumer complaints received by the NCDA&CS, and other agencies such as NCDENR. In addition to consumer complaints as mentioned above, NCFEDA can leverage emergency room
information from systems like North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT). Web 2.0 data such as social networks and microblogs can also provide real-time information. Information captured from microblogs will be incorporated into NCFEDA in the future.

**Outbreak Investigation and Detection.** Currently, surveillance of and response to foodborne disease as performed by respective North Carolina agencies is based on traditional outbreak and detection data that include sources like the North Carolina Electronic Data Surveillance System (NC EDSS) as reported by the NCDPH and the BETS database as reported by the NCDENR. The lag time for these data sources averages approximately 7 days.

**Inspections and Recall.** In addition, recall and inspection data from NCDA&CS and NCDENR are available with a lag time up to 1 month. These data include not only inspections reports, but also identification data on all food manufacturing facilities and food serving facilities in North Carolina. In addition, recall data are available from previous food contamination events.

---

4 NCFEDA does not currently use NCDETECT data but rather uses a proxy database provided by NCDPH.
**Routine Trend Analysis.** For retrospective analysis, including trend analysis for policy assessment, data from the North Carolina Comprehensive Assessment for Tracking Community Health (NC CATCH) system and the North Carolina Economic Development Intelligence System (NC EDIS) can provide important information. The lag time for these last sources can be 1 year or more.

The above classification approximates data availability based on reporting time by taking into consideration the lag time between the date a food safety related “event” takes place and the date data about this event are reported to state public officials. For example, for patients with less severe symptoms, information collected from LHDs and NCDPH indicates that it may take several weeks up to a month for data about those visits to physicians to be entered into NCEDSS. NCEDSS is classified under “Monthly” reporting time because, even though data entered into NCEDSS are available to NCDPH in near real-time, physicians’ delay in reporting to LHDs postpones data input into NCEDSS, and consequently data availability to NCDPH.

Table 4 presents the resulting list of data collected per agency. Data integration is one of the main contributions of NCFEDA to the food safety community in North Carolina. With the assistance of the North Carolina Food Safety & Defense Task Force, data sources available to NCFEDA have been identified by representatives of each relevant stakeholder agency. The collected data pertains to each agency’s jurisdictional responsibilities.

For example, NCDA&CS Food and Drug Protection Division regulates food manufacturers, wholesalers, warehousers, distributors, and select retail organizations, while NCDENR Division of Environmental Health regulates retail food service, child care, lodging, institutions, schools, local confinement, public swimming pools, shellfish and crustacea, camps, and tattoo artists. NCDENR is currently working with NCDA&CS to integrate complaints received by NCDENR with those of NCDA&CS. In this first phase of development, not all data sources in Table 4 were available for use by the NCFOODSAFE project. We have identified and included the most relevant datasets for the selected use case described in Chapter VII and will include additional datasets in future phases of the project.

### Table 4. Food Safety & Defense Data Collected Per Agency

<table>
<thead>
<tr>
<th>NCDA &amp; CS Food &amp; Drug Protection Division</th>
<th>NCDA &amp; CS Meat &amp; Poultry Inspection Division</th>
<th>DHHS Division of Public Health</th>
<th>NC DENR Division of Environmental Health</th>
<th>Local Health Departments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firms which produce, hold, pack, or distribute food products (including location and contact information)</td>
<td>Daily inspection procedure results</td>
<td>Patient name and address</td>
<td>Critical food safety violations via food service establishment Inspection data</td>
<td>DENR-regulated inspection violations: demographics, durations, compliance with each item (.2600 sheet, etc.)</td>
</tr>
</tbody>
</table>
Table 4. Food Safety & Defense Data Collected Per Agency (continued)

<table>
<thead>
<tr>
<th>NCDA &amp; CS Food &amp; Drug Protection Division</th>
<th>NCDA &amp; CS Meat &amp; Poultry Inspection Division</th>
<th>DHHS Division of Public Health</th>
<th>NC DENR Division of Environmental Health</th>
<th>Local Health Departments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location and information of retail grocery stores (excluding prepared food and meat market portions)</td>
<td>Meat &amp; Poultry recalls</td>
<td>Lab results</td>
<td>Environmental investigation data &amp; blood lead levels (NC EDSS)</td>
<td>Consumer complaint data linked to DENR establishment, categorization of complaint, free notes fields</td>
</tr>
<tr>
<td>Consumer complaints regarding agricultural commodities or manufactured food</td>
<td>Pathogen Sampling, results &amp; actions</td>
<td>Symptoms</td>
<td>Inspections and sampling data on dairy farms, producers, and haulers</td>
<td>GMS-Water sample reporting system. Samples are entered into the system and the results are reported into the system</td>
</tr>
<tr>
<td>Inspection results and violative findings for food manufacturers or wholesalers</td>
<td>Staffing database</td>
<td>Some food history &amp; exposure</td>
<td>Data from samples of water and shellfish meats from coastal shellfish growing waters</td>
<td>Consumer website</td>
</tr>
<tr>
<td>Recall effectiveness check data</td>
<td>Individual plant performance percents</td>
<td>Clinical outcomes</td>
<td>Coastal pollution sources adjacent to shellfish and recreational waters are mapped</td>
<td>GIS Department data</td>
</tr>
<tr>
<td>Investigative laboratory results</td>
<td>GPS plant locations</td>
<td>NA</td>
<td>Inspections and sampling data of certified shellfish and crustacean processing plants</td>
<td>NA</td>
</tr>
<tr>
<td>Food/Ag commodity survey sample test results</td>
<td>Consumer complaints</td>
<td>NA</td>
<td>Data from samples of coastal recreational waters and notifications to the public when bacteria levels exceed swimming standards</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>Notices &amp; directives</td>
<td>NA</td>
<td>Consumer complaints re: DEH-regulated facilities</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>Food safety audit results/actions</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>Employee training database</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA = not applicable.
Table 5 lists the data fields contained in test datasets.

**Table 5. Data Fields of Testing Datasets Obtained from Agencies**

<table>
<thead>
<tr>
<th><strong>NCDA&amp;CS Complaints Data</strong></th>
<th><strong>NCDENR BETS Table</strong></th>
<th><strong>NCDPH NETSS Data</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>ComplaintID</td>
<td>FacilityIDNumber</td>
<td>Event</td>
</tr>
<tr>
<td>ComplaintYear</td>
<td>FAC_ID</td>
<td>EventName</td>
</tr>
<tr>
<td>DateEnteredDB</td>
<td>FAC_TYPE</td>
<td>EventDate</td>
</tr>
<tr>
<td>DateModified</td>
<td>FAC_TYPE_ORiginal</td>
<td>Status</td>
</tr>
<tr>
<td>AssignedToInspector</td>
<td>FAC_ID_NUMBER</td>
<td>EventType</td>
</tr>
<tr>
<td>DateReceived</td>
<td>STATUS_CODE</td>
<td>Age</td>
</tr>
<tr>
<td>LoggedByComplOfficer</td>
<td>STATUS_CODE_ORiginal</td>
<td>AgeType</td>
</tr>
<tr>
<td>ReceivedByMeans</td>
<td>CORP_ID</td>
<td>Race</td>
</tr>
<tr>
<td>ReferredType</td>
<td>EHS_ID</td>
<td>Ethnic</td>
</tr>
<tr>
<td>ComplainantLastName</td>
<td>WATER_TYPE</td>
<td>Sex</td>
</tr>
<tr>
<td>ComplainantFirstName</td>
<td>WASTE_TYPE</td>
<td>County</td>
</tr>
<tr>
<td>ComplainantCity</td>
<td>FAC_SOURCE</td>
<td>CountyCode</td>
</tr>
<tr>
<td>ComplainantState</td>
<td>FAC_NAME</td>
<td>MMWRWK</td>
</tr>
<tr>
<td>ComplainantZipCode</td>
<td>FAC_PERMITTEE</td>
<td>Reported</td>
</tr>
<tr>
<td>ComplainantCounty</td>
<td>FAC_MANAGER</td>
<td>Died</td>
</tr>
<tr>
<td>HomePhone</td>
<td>FAC_TERRITORY</td>
<td>Year</td>
</tr>
<tr>
<td>ProductBrand</td>
<td>FAC_MAP_NUM</td>
<td>NA</td>
</tr>
<tr>
<td>ProductUsed</td>
<td>FAC_SEATS</td>
<td>NA</td>
</tr>
<tr>
<td>LotCodeDateExpDate</td>
<td>FAC_OLD_ID</td>
<td>NA</td>
</tr>
<tr>
<td>LotSize</td>
<td>FAC_EMAIL</td>
<td>NA</td>
</tr>
<tr>
<td>PurchaseDate</td>
<td>FAC_LATITUDE</td>
<td>NA</td>
</tr>
<tr>
<td>ProductDescription</td>
<td>FAC_LONGITUDE</td>
<td>NA</td>
</tr>
<tr>
<td>FDAIndCode</td>
<td>FAC_PARCEL_ID_NUMBER</td>
<td>NA</td>
</tr>
<tr>
<td>FDAProductCategory</td>
<td>FAC_OP_HOURS_FROM</td>
<td>NA</td>
</tr>
<tr>
<td>ComplaintCode</td>
<td>FAC_OP_HOURS_TO</td>
<td>NA</td>
</tr>
<tr>
<td>MedicalCode</td>
<td>FAC_MAX_CAP_CHILDREN</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 5: Data Fields of Testing Datasets Obtained from Agencies (continued)

<table>
<thead>
<tr>
<th>NCDA&amp;CS Complaints Data</th>
<th>NCDENR BETS Table</th>
<th>NCDPH NETSS Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>ContainerCode</td>
<td>FAC_STATE_NOTES</td>
<td>NA</td>
</tr>
<tr>
<td>ComplaintSummary</td>
<td>FAC_NUMBER_STAFF</td>
<td>NA</td>
</tr>
<tr>
<td>ManufacturerName</td>
<td>FAC_LICENSING_AGENCY</td>
<td>NA</td>
</tr>
<tr>
<td>ManufacturerAddress</td>
<td>FAC_FOOD_SERVICE</td>
<td>NA</td>
</tr>
<tr>
<td>ManufacturerCity</td>
<td>FAC_POOL_ONSITE</td>
<td>NA</td>
</tr>
<tr>
<td>ManufacturerState</td>
<td>FAC_BUILT_PRE_1978</td>
<td>NA</td>
</tr>
<tr>
<td>ManufacturerZipCode</td>
<td>FAC_INTENT_SUSP_DT</td>
<td>NA</td>
</tr>
<tr>
<td>RetailerName</td>
<td>FAC_SUSPEND_DT</td>
<td>NA</td>
</tr>
<tr>
<td>RetailerAddress</td>
<td>FAC_LATE lettre DT</td>
<td>NA</td>
</tr>
<tr>
<td>RetailerCity</td>
<td>FAC_BAD_CHECK_DT</td>
<td>NA</td>
</tr>
<tr>
<td>RetailerZipCode</td>
<td>FAC_ADDED_DT</td>
<td>NA</td>
</tr>
<tr>
<td>RetailerCounty</td>
<td>FAC_LAST_UPDATE_DT</td>
<td>NA</td>
</tr>
<tr>
<td>ComplaintRemarks</td>
<td>FAC_LAST_UPDATE_USER</td>
<td>NA</td>
</tr>
<tr>
<td>Archieved</td>
<td>COUNTY_ID</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>COUNTY_ID_FIPS</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>FAC_CATEGORY_ID</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>FAC_CATEGORY_ID_Orig</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>FAC_CHANGE_IND</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>FAC_DAYS_OF_OPERATION</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>FacilityNum_Text</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>FacilityType_Text</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>UpdateInfo</td>
<td>NA</td>
</tr>
</tbody>
</table>

NA = not applicable.

Architecture

NCFEDA is envisioned as a “meta”-level software tool that provides critical capabilities in data integration, visualization, and analytics for managing a foodborne contamination event. This meta-level software tool is a computer application that supports reification of data elements through aggregation (e.g., sharing the counts of certain data field values instead of the individual values) or encapsulation such as normalizing aggregated data values.
Data integration occurs across responding agencies—NCDPH, NCDA&CS, and NCDENR—as necessary for situational awareness, and also includes new data sources from the private sector and the consumer. For example, on the private sector side, recall alerts and enforcement reports provide information about contaminated food products as reported by manufacturing companies to FDA and USDA. On the consumer side, consumer complaints collected by agencies’ complaint hotlines are used as triggers for the NCFEDA system.

Figure 7 presents NCFEDA’s architecture integrating all major components of the system (i.e., external input data, NCFEDA analytical tools, and stakeholders’ dashboards).

A discussion of each of the components follows.

**Databases**

A *Food-related Events Database* and a *Geocode Database* store all raw data obtained from both private and public sources and are foundation resources for all information analyzed by NCFEDA and displayed on Stakeholders’ Dashboards. NCFEDA brings together—"fuses"—food event data distributed across multiple systems and controlled by the state’s major food
safety stakeholders. The original stakeholders’ data are anonymized as necessary, processed, and stored in NCFEDA databases. Whenever location information is provided (e.g., zip code of a manufacturer/retailer/consumer) the corresponding record is geocoded (i.e., latitude and longitude coordinates are generated for the given location) to allow plotting this location on a map to be displayed by NCFEDA visualization tools.

In addition to the cleansing process all data provided by public health agencies go through, NCFEDA contains its own set of tools that allow harvesting data directly from public websites. These software applications enable communication between NCFEDA and the Internet, grabbing of available data, extraction of links, raw and metadata from a webpage, and further navigation to the links obtained to harvest any additional pages. This is the method used by NCFEDA to obtain information about recalls of food products issued by the FDA and USDA. Given that these recalls are intended for human consumption, they are written in natural language which poses an additional challenge for the extraction of information. A customized text application was incorporated to NCFEDA to allow automatic parsing of the recalls text and extraction of relevant information. These data are stored in NCFEDA’s databases and processed to feed several of NCFEDA’s dashboards, as described in the next sections of this document.

Triggers

The Analysis Trigger Module works in conjunction with the Analytics Engine to analyze incoming data and determine their relevance to the current situation(s) being monitored or other data previously stored in the system. As time passes, NCFEDA databases must be maintained to keep up to date with new available information. The bulk of these data constitute the history of food safety in North Carolina and will be used by NCFEDA analytical tools to support or refute possible conclusions regarding emerging food events. The more near real-time events (i.e., those transmitted or harvested daily or in the very near past—2 to 3 days’ window—by NCFEDA) are treated as “analysis triggers” by the Relevance Engine.

Every new data arrival is analyzed by the engine and may or may not activate one or more NCFEDA logic rules which are the basis for NCFEDA’s decision processes. Currently, the following triggering events are represented in NCFEDA’s logical rules and analyzed: (a) Illness cases reported to public health officials; (b) Recall notices issued by FDA and USDA; and (c) Consumer complaints reported to NCDA&CS. Upon data availability, it is expected that at least two other triggers are added to the engine: (1) Inspection violation codes issued by NCDENR to all food services facilities the agency regulates, and those issued by NCDA&CS to food manufacturers and distributors; and (2) Microblog messages published on the Internet by individuals implicating a food service facility where they had eaten or reporting sickness caused by a food product consumed at home.
Dashboards

Fused event data identified as relevant to users by the engine are sent to the Visualization Tool for processing and generation of graphs and charts to be displayed by Stakeholder Dashboards. The dashboards are a group of dynamic graphical user interfaces configured to present each agency-user with a common operating picture and additional information that convey situational awareness targeted to the agency’s audience. Given that each responding agency has distinct goals and needs, the Dashboards may contain both elements that are shared by all responding agencies (e.g., information concerning ongoing outbreaks) and information that is of specific interest to each agency (e.g., levels of recall effectiveness for NCDA&CS). Each Dashboard provides selected functionalities to allow its users to execute queries by “slicing and dicing” the data presented.

Analytical Tools

The human decision-making process takes into account what is known by the decision maker—"facts" describing the situation at hand and preestablished procedures/regulations—"processes" that dictated how that particular situation must be handled, and process this information through an activity known as “logical reasoning”—which allows humans to identify relationships among seemingly independent elements of a problem in the search for a solution. When it is not possible to apply any known processes to the known facts, humans can resort to using logic knowledge to link apparently unrelated facts to get a better understanding of the problem and find an answer, or delay any decision until more information is available or a new method is devised.

Connecting information, or finding the relationship among isolated facts, and selecting what is relevant to the task at hand is key to enable humans to make better decisions in a timely and efficient manner. Today, representing facts and processes to fit traditional computers’ execution models is an ordinary task and makes it possible to automatically control many operations with these machines. Facts are well suited to be represented in databases and processes as sequences of instructions of computer programs.

In a simplified view of the food safety system, food-related data compiled from different sources constitute factual knowledge about foodborne illness cases, food-related consumer complaints, food products’ recalls, inspections’ violation codes imposed by regulating agencies on food service facilities with poor quality standards, etc.; physicians’ reporting of suspected foodborne disease cases to public health authorities, pulling out of commerce contaminated products implicated in a recall, and collecting samples and testing food products implicated in a consumer complaint call are examples of procedural knowledge undertaken to deal with a possible food-related threat or event.

Representing logical reasoning in such a way that it can be performed by computers, on the other hand, is not an easy endeavor because the relationships to be represented may require a complex set of rules that cannot be easily encoded in a database or a program with a
well-defined flow of control. Usually logical knowledge is encoded as “inference rules” using some computer programming language and these rules are processed, together with facts, by another software application called a “reasoning engine” to produce answers. NCFEDA’s Relevance Engine is such an engine which performs rule-based predictive analytics, or “reasons” about an existing situation as described by the known facts and encoded rules, and deduces relationships among data to generate answers—relevant information that is relayed to users to improve their situational awareness and help their decision-making process.

**Figure 8** provides an example of a set of rules used by NCFEDA to reason about what information is most relevant when the incoming triggering event is a cluster of illness cases reported to public health officials. Values of data fields from different data sources are considered by NCFEDA’s reasoning process in the search for possible relationship among these elements. In this example, the search process is illustrated by four consecutive steps that are initiated when a triggering event is entered into NCFEDA’s databases. At each step, any relevant information identified is “pushed” to users in the form of a short message that

**Figure 8. Example Rule Set Triggered by Illness Cluster Detected by Public Health**

Cluster of Reported Illnesses Identified by NC Public Health Surveillance System

**Pathogen:** P  
**Start Date:** MM/YYYY  
**Counties:** L1, L2, ...

**Step 1:** Search **Recalls** that match rule:
- IF RECALL issued one month prior to MM/YYYY date AND cites pathogen P AND recalled product was distributed in NC or nationwide, THEN return recall text to allow access to recalling company, product name and distribution state.

**DATE ISSUED:** MM/DD1/YYYY  
FDA recalls product Q manufactured by company C due to contamination with pathogen P. The product was distributed to following states: S1, S2, NC, ...

**CALL DATE:** MM/DD2/YYYY  
Consumer called to report illness confirmed as caused by pathogen P after eating product Q from company C. Consumer resides in county L2.

**VIOLATION DATE:** MM/DD3/YYYY  
Inspected facility F tested positive for pathogen P on MM/DD3/YYYY. Company F manufactures product Q and is located in NC county L2.

**DATE POSTED:** MM/DD4/YYYY  
Blogger reported illness confirmed as caused by pathogen P after eating product Q from company C at restaurant R. Blogger resides in NC and cites county L1.

**Step 2:** Search **Complaints** that match rule:
- IF COMPLAINT received one month prior to MM/YYYY date OR consumer resides in one of the counties in the cluster L1 or L2 or ... AND/OR illness caused by pathogen P AND/OR recalled product found by RECALL search rule AND/OR recalling company found by RECALL search rule, THEN return complaint call text.

**DATE POSTED:** MM/DD1/YYYY  
Blogger reported illness confirmed as caused by pathogen P after eating product Q from company C at restaurant R. Blogger resides in NC and cites county L1.

**Step 3:** Search **Violations** that match:
- IF CODE VIOLATION occurred one month prior to MM/YYYY date AND/OR facility located in one of the counties in the cluster L1 or L2 or ... AND/OR pathogen P found in facility AND/OR company manufactures recalled product found by RECALL search rule, THEN return code violation text.

**DATE POSTED:** MM/DD1/YYYY  
Blogger reported illness confirmed as caused by pathogen P after eating product Q from company C at restaurant R. Blogger resides in NC and cites county L1.

**Step 4:** Search **Microblogs** that match:
- IF MICROBLOG MESSAGE posted one month prior to MM/YYYY date AND/OR blogger resides in NC (or mentions county L1 or L2 or ...) AND/OR illness caused by pathogen P AND/OR blogger cites recalled product or recalling company found by RECALL search rule, THEN return message text.

**DATE POSTED:** MM/DD1/YYYY  
Blogger reported illness confirmed as caused by pathogen P after eating product Q from company C at restaurant R. Blogger resides in NC and cites county L1.
provides key elements to increase the users’ understanding of the present state of affairs. As it is shown by screenshots (Figures 25, 26, and 27) presented in Chapter VII, a graphical marker indicating the location of the resulting elements is also placed on a map depicted in one (or more) of the NCFEDA system dashboard screens.

In addition to identifying relevant information of possibly emerging events to be “pushed” to users, NCFEDA’s engine will also rate the level of importance of this information to users and compute a (currently) simple measure of the likelihood of the event in consideration being a real threat. The computation of the “Event Likelihood Index” (ELI) metric is roughly based on the number of relationships identified by NCFEDA as connecting the key data elements part of the description of the possibly emergent event. Figure 9 presents the seven possible levels of the ELI Ratings scale. We demonstrate the usage of the ELI measure in the software prototype that implements the first NCFEDA use case, as described in the next section.

**Analytics Engine**

NCFEDA reasoning capabilities are powered by formal logic. This means that the NCFEDA system “reasons” about food safety events by applying deductive reasoning (i.e., inference rules) to facts informed to the engine to infer (new) knowledge. Conceptually, the NCFEDA is composed of several modules which together constitute its *Analytics Engine*—the intelligent component of the system responsible for drawing conclusions about a given food safety situation.

The *Analytics Engine*’s modules execute the following main functions: (1) analysis of the incoming triggering events; (2) fusion of event data previously acquired directly or indirectly from stakeholders’ surveillance and reporting systems; and (3) processing of new trigger information against known data by using the relevance engine’s deductive mechanisms together with various sets of rules (i.e., predictive analytics). The NCFEDA *Visualization Tools* module is set apart from the *Analytics Engine* because its main function is to process (i.e., create visual representations of) the results produced by the engine’s rules for display on users’ dashboards, to help increase users’ situational awareness by presenting information in a more user-friendly interface.

More specifically, apart from external data inputs and outputs, the *Analytics Engine* implementation consists of the following:

---

**Figure 9. Event Likelihood Index Table**

<table>
<thead>
<tr>
<th>ELI RATINGS</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No Likelihood</td>
</tr>
<tr>
<td>1</td>
<td>Possible Likelihood</td>
</tr>
<tr>
<td>2</td>
<td>Low Likelihood</td>
</tr>
<tr>
<td>3</td>
<td>Some Likelihood</td>
</tr>
<tr>
<td>4</td>
<td>Moderate Likelihood</td>
</tr>
<tr>
<td>5</td>
<td>Significant Likelihood</td>
</tr>
<tr>
<td>6</td>
<td>High Likelihood</td>
</tr>
<tr>
<td>7</td>
<td>Highest Likelihood</td>
</tr>
</tbody>
</table>
1. **A Relevance Engine**—the core, domain-independent inference module plus a set of approximately 100 inference rules that we created to define the food safety domain problem;

2. **Auxiliary Data Concepts**—a set of seven factual databases, which store concepts of interest to the task of reasoning about food safety events, represented as logical knowledge for easy processing by the Relevance Engine. These concepts include four (simplified) ontologies for food, foodborne diseases, and geographical information, and three databases which contain FDA’s Food Code and the medical and consumer complaints codes used by NCDA&CS to process consumer complaints about food products.

At its present state, the NCFEDA Analytics Engine processes triggering events’ data against other food safety data already stored in its databases and generates one or more possible “models” of the situation being evaluated. By definition, a model is a consistent set of knowledge assertions that the engine infers from the given inputs and the concepts it knows. The NCFEDA Analytical Engine data usage flow is illustrated by the diagram in Figure 10 and indicates the types of results expected to be produced by the engine for two different use cases scenarios.

For example, in Scenario 1 presented in Figure 8, the set of rules given is used by the NCFEDA Analytics Engine when one or more confirmed clusters of illnesses are created and reported to the NCFEDA system by public health officials to match patterns among all available data and auxiliary data concepts. These and other rules enable the engine to potentially find one or more “models” of the present situation where:

A. One or more recall notices may be linked to the confirmed cluster(s) of illnesses provided;
B. One or more consumer complaint calls may be linked to the given outbreak cluster(s);
C. A numerical indicator—the ELI—which is a measure of how strong is the evidence that engine results (a) and (b) are in fact linked to the outbreak.

Assume a second scenario, namely Scenario 2, where the engine has no knowledge of a newly occurring cluster of illnesses corresponding to a foodborne disease outbreak confirmed by public health. The NCFEDA system, however, continuously receives records of ill patients’ visits reported by healthcare providers and the Analytics Engine seeks to find relationships among these records to form new suspected cluster(s) of illnesses, unbeknown or not yet reported by public health officials. If a new suspected cluster is identified by the engine, it proceeds to compute similar results as those generated for confirmed clusters, as described above, and generates “models” that matches this situation. If the engine can find a model—a set of consistent assertions that are valid in the present situation—then this model is fed to NCFEDA Visualization Tools which convert these statements into visual elements to be displayed on the User Dashboards.
When integrating the different food-related data sources, the application focus on what type of positive information each individual data source can provide that helps NCFEDA achieve the primary goal set for the NCFOODSAFE project of creating informatics tools to help reduce latency in surveillance and response to food safety events. The diagram in Figure 11 presents the simplified pattern matching scheme, corresponding to Scenario 1, which is used by the NCFEDA Analytics Engine to identify relationships among records of different public health agencies.
Although patient illness records, and therefore illness clusters, may indicate a definite diagnosis and foodborne pathogen (e.g., *Salmonella*) in the majority of the cases it is not known what food caused the problem. This is crucial information sought by public health officials so that measures can be taken to eliminate the source of the contamination and protect the public. If enough information is available via multiple data sources, the NCFEDA Analytics Engine process can potentially find this answer in minutes. The pattern matching process applied by the Analytics Engine is described in more detail below for both scenarios.

**Scenario 1 Triggered by Confirmed Cluster of Foodborne Illnesses.** A depiction of some of the pattern matching rules, applied by the engine when it is informed of one (or more) confirmed clusters of illnesses, is presented in Figure 12. Applying rule *Cluster2Recall* the engine may be able to match the cluster-confirmed pathogen, patients’ residing areas, and the period of illness onset, with the pathogen indicated in a recall notice, geographical area where the product has been distributed, and dates of product distribution. This rule will return the information extracted from the recall notice, such as recalled food product, recalling manufacturer, area, and other product distribution information. Rule *Cluster2Complaint* tries to match the geographical area affected by a confirmed cluster, a time window encompassing the reported illness cases part of the outbreak, and if enough information is available also match the cluster pathogen with consumer complaints originated in the same geographical area, timeframe, and reported illness. Successful application of this rule will generate a list of consumers who have called in and may be affected by the ongoing outbreak although not yet linked to it. If such a consumer is identified, information about the person’s residing location, illness condition, and the possibly adulterated food product(s) implicated, is used by rule *Cluster2Complaint2Recall* together with the information obtained with rule *Cluster2Recall* to narrow down the number of possible matches to those with stronger evidence to be connected to the reported outbreak cluster. Results of this rule are then parsed and translated from logical statements to data that are fed to NCFEDA Visualization Tools to generate graphical elements to be displayed on users’ dashboards. An example of the results obtained through these rules is presented next.
Scenario 1 Implementation and Results. To demonstrate the power and potential of the inference rules of the NCFEDA Analytics Engine let us examine a simplified example of a possible outbreak situation. By definition, a foodborne disease outbreak consists of two or more patients with the same foodborne illness linked to a common source of contamination. To comply with HIPAA law, all patient or complainant personal information protected under the law has been stripped from any records processed by NCFEDA and any example presented in this document. We assume that the date of disease onset adopted for a cluster is the earliest disease onset date among all patients associated to the cluster. This patient is then referred to as “patient #1” of that cluster.

In this example, assume that NCDPH staff has detected three ongoing clusters of foodborne disease outbreak and information regarding these clusters is provided to the NCFEDA system as follows:
A. Cluster 1 contains five patients diagnosed with salmonelosis who reside in three counties. Patient #1 of Cluster 1 has unique ID PT01035 and disease onset date of February 2011. The illness records of the five patients associated to Cluster 1 contain the following information:
- Patient #1 has ID PT01035 and resides in Orange County.
- Patient #2, ID PT01041, from Wake County.
- Patient #3, ID PT01042, from Wake County.
- Patient #4, ID PT01074, from Durham County.
- Patient #5, ID PT01075, from Durham County.

B. Cluster 2 contains two patients diagnosed with salmonelosis who reside in two counties. Patient #1 of Cluster 2 has unique ID PT01078 and disease onset date of January 2011. The illness records of the two patients associated to Cluster 2 are:
- Patient #1, ID PT01078, Orange County.
- Patient #2, ID PT01081, Wake County.

C. Cluster 3 contains three patients diagnosed with listeriosis who reside in three counties. Patient #1 of Cluster 3 has unique ID PT01022 and disease onset date of January 2011. The illness records of the patients associated to Cluster 3 are:
- Patient #1, ID PT01078, Pitt County.
- Patient #2, ID PT01049, Franklin County.
- Patient #2, ID PT01053, Wake County.

The cluster information entered into the NCFEDA system is translated into logical statements—"facts"—to be processed by the NCFEDA Analytics Engine. The facts corresponding to the above information, formatted in the logical language understood by the Analytics Engine, are shown in Figure 13.

The original records of the patients associated to these clusters, as formatted in Figure 14, include additional information, such as date and time of reporting to NCDPH syndromic surveillance systems, and are represented by the following facts in the NCFEDA Syndromic Surveillance DB:

Consider that the five hypothetical recall notices presented in Figure 15 have been issued by USDA and FDA between December 2010 and February 2011. The fact describing Recall 1, identified as “rc1,” can be read as “On February 1, 2011, at 9:00 A.M., USDA issued a recall notice for a Cargill's meat product because of Salmonella contamination. The product has been distributed to the states of New York, New Jersey, and North Carolina. No illness has been reported to be associated to this recall.” Facts associated to the other recalls listed in Figure 15 are read in a similar manner.
Figure 13. Scenario 1—Excerpt of Cluster Data Processed by NCFEDA Analytics Engine

```
% CLUSTER 1
cluster(cluster1, salmonella, pt01035, feb, 2011, 3, 5).
cluster_illrecord(cluster1, pt01035, orange).
cluster_illrecord(cluster1, pt01041, wake).
cluster_illrecord(cluster1, pt01042, wake).
cluster_illrecord(cluster1, pt01074, durham).
cluster_illrecord(cluster1, pt01075, durham).

% CLUSTER 2
cluster(cluster2, salmonella, pt01078, jan, 2011, 2, 2).
cluster_illrecord(cluster2, pt01078, orange).
cluster_illrecord(cluster2, pt01081, wake).

% CLUSTER 3
cluster(cluster3, listeria, pt01022, jan, 2011, 3, 3).
cluster_illrecord(cluster3, pt01022, pitt).
cluster_illrecord(cluster3, pt01049, franklin).
cluster_illrecord(cluster3, pt01053, wake).
```

Figure 14. Scenario 1—Excerpt of Patient Illness Records Processed by NCFEDA Analytics Engine

```
ilness_record(pt01022, 9,0,am, 20,jan,2011, pitt, ncdetect, listeria).
ilness_record(pt01035, 9,0,am, 1, feb,2011, orange, ncdetect, salmonella).
ilness_record(pt01041, 9,0,am, 2, feb,2011, wake, ncdetect, salmonella).
ilness_record(pt01042, 9,0,am, 2, feb,2011, wake, ncdetect, salmonella).
ilness_record(pt01049, 9,0,am, 29,jan,2011, franklin, ncdetect, listeria).
ilness_record(pt01053, 9,0,am, 30,jan,2011, wake, ncdetect, listeria).
ilness_record(pt01074, 9,0,am, 3, feb,2011, durham, ncdetect, salmonella).
ilness_record(pt01075, 9,0,am, 3, feb,2011, durham, ncdetect, salmonella).
ilness_record(pt01078, 9,0,am, 17,jan,2011, orange, ncedss, salmonella).
ilness_record(pt01081, 9,0,am, 16,jan,2011, wake, ncedss, salmonella).
```
Examples of facts corresponding to consumer complaint calls are shown in Figure 16. Often, information provided by consumers is incomplete; other times, consumers may prefer to remain anonymous and abstain from providing complete information about a complaint. Records with incomplete information are allowed in our system and unknown data are represented by value “unk,” as shown in Figure 16. A more comprehensive description of the theoretical foundation of the methodology employed in developing the rules for the NCFEDA Analytics Engine, and the syntactic and semantics of the language, goes beyond the scope of this report and will be submitted for publication as a technical paper.

When NCFEDA Analytics Engine’s inference rules from Figure 8 process these data in conjunction with its other rules, auxiliary data concepts, and databases, it generates a solution model—partially shown in Figure 17—which may potentially help public health officials with their outbreak investigations.
Figure 16. Scenario 1—Excerpt of Consumer Complaints Processed by NCFEDA Analytics Engine

complaint_user(c108388, 6, 26, am, 28, jan, 2011, unk, nc, unk, wake, no_illness_or_injury, unk, unk).
complaint_mft(c108388, unk, unk, unk, unk, unk, unk, unk, unk).
complaint_food(c108388, unk, retail_stores, unk, unk).
complaint_user(c108388, 5, 11, pm, 18, jan, 2011, unk, nc, unk, wake, medical_condition_unknown, unk, unk).
complaint_mft(c108388, unk, unk, unk, unk, unk, unk, unk, unk).
complaint_food(c108388, vegetable, vegetables_beans_corn_peas_tomatoes_peppers, unk, unk).
complaint_user(c100378, 12, 17, pm, 14, jan, 2011, unk, nc, unk, wake, no_illness_or_injury, unk, unk).
complaint_mft(c100378, unk, unk, unk, unk, unk, unk, unk, unk).
complaint_food(c100378, cheese, cheese_products, unk, unk).

Figure 17. Excerpt of Resulting Solution Model Generated by NCFEDA Analytics Engine for Scenario 1

outbreak(cluster1,salmonella,5)
outbreak(cluster2,salmonella,2)
outbreak(cluster3,listeria,3)

recall_connected(rc1,cluster1,meat,cargill,north_carolina)
recall_connected(rc1,cluster2,meat,cargill,north_carolina)
recall_connected(rc2,cluster2,chicken,sandia,nationwide)
recall_connected(rc4,cluster3,spinach,giant,south)
recall_connected(rc5,cluster3,lettuce,medleys,nationwide)

pref_recall_connected(rc1,cluster1,meat,cargill,north_carolina)
pref_recall_connected(rc1,cluster2,meat,cargill,north_carolina)
pref_recall_connected(rc4,cluster3,spinach,giant,south)

complaint_connected(c108388,cluster2,unk,retail_stores,wake)
complaint_connected(c108388,cluster3,unk,retail_stores,wake)
complaint_connected(c108388,cluster2,vegetable,vegetables_beans_corn_peas_tomatoes_peppers,wake)
complaint_connected(c108388,cluster3,vegetable,vegetables_beans_corn_peas_tomatoes_peppers,wake)
complaint_connected(c108378,cluster2,cheese,cheese_products,wake)
complaint_connected(c108378,cluster3,cheese,cheese_products,wake)

belong(meat,meat_poultry)
belong(spinach,vegetables_beans_corn_peas_tomatoes_peppers)

ccr(cluster3,listeria,rc4,spinach,giant,south,vegetable,vegetables_beans_corn_peas_tomatoes_peppers,wake)
Three important conclusions reached by the engine are:

1. *Giant’s* spinach product distributed to states in the U.S. South region, and being recalled under Recall Notice RC4, is the probable source of listeria contamination for Cluster 3;
2. The existence of some consumer complainants who may be linked to Cluster 2 and Cluster 3, and yet not officially associated with those outbreak clusters;
3. A preferred set of associations of clusters and recall notices based on the geographic area of the product distribution (i.e., recalls of products reportedly distributed in or closer to North Carolina are given preference by the engine when creating possible connections to the state’s clusters).

In addition, the engine computes the ELI which indicates how strong is the evidence that such outbreak event and the answers derived by the engine are relevant and should be investigated by public health staff. The ELI measures computed for Scenario 1 are presented in Figure 18. Engine results are then parsed and translated into graphical elements by the Visualization Tools which feed NCFEDA users’ dashboards.

**Scenario 2 Triggered by Analytical Engine’s Detection of Suspected Cluster of Foodborne Illnesses.** When no new confirmed cluster of illness is provided as an input, but the engine detects increased number of illness cases reported in a certain geographical area of North Carolina another set of rules is triggered. These rules would verify if there is enough evidence to link these cases—based on time of illness onset or reporting, geographical proximity, and confirmed pathogen/disease or probable cause—and group these cases together to create a “suspected” cluster. NCFEDA uses a numerical threshold to decide if any given geographical region (e.g., a county) should be evaluated for cluster formation and region’s physical proximity and adjacent condition to decide which counties to group together.

For each suspected cluster formed, the Analytics Engine will apply rules, depicted in Figure 19, similar to those of Scenario 1 above. For example, the *SuspCluster2 Recall* rule matches a suspected cluster pathogen, geographical coverage area, and timeframe to existing food recall notices, similar to the *Cluster2Recall* rule of Scenario 1, to identify contaminated food products that may be the cause of contamination for the suspected cluster formed. Rules *SuspCluster2Complaint* and *SuspCluster2Complaint2Recall* are analogous to the *Cluster2Complaint* and *Cluster2Complaint2Recall* rules, respectively.
Scenario 2 Implementation and Results. Engine rules are not triggered only by new information about ongoing clusters of foodborne disease. In the absence of new confirmed clusters, and upon continuous input of new reported patient visits to healthcare providers, the engine checks for increased reporting activity on a geographical area during a certain timeframe based on matching diagnoses. Any record whose listed diagnosis specifies a causing pathogen (e.g., *Listeria*) can only be matched to other records with the same pathogen (*Listeria*) diagnosis. Often, physicians cannot determine the patient’s condition before laboratory tests are conducted to determine which pathogen is causing the illness. In these cases, if the physician suspects foodborne disease then this is listed as a probable diagnosis. In the cases where it is only clear that the patient is suffering from gastrointestinal problems but there is not enough evidence for the physician to suspect a more serious condition linked
to foodborne disease, the probable diagnosis may be indicated only as “gastro-intestinal ulceration” (GIU). Those records presenting an initial suspected, but more generic diagnosis of either “foodborne disease” or GIU are assumed to (1) match other records with a specific pathogen diagnosis (e.g., GIU matches *Listeria*, or GIU matches *Salmonella*); (2) the same diagnosis (e.g., GIU matches GIU); or (3) be a match themselves (e.g., GIU matches foodborne disease).

In this example, the engine processed a larger set of illness records than the one shown in Figure 14, many of which had GIU as the probable diagnosis. Based on the disease’s pattern matching conditions applied by the engine, six counties, listed in Figure 20, were found to have reporting activity above the threshold stipulated for the engine; two counties (i.e., Mecklenburg and New Hanover) were found to have elevated activity for two different pathogens (i.e., *Salmonella* and *Shigella*). Only one suspected cluster, however, was formed by the engine because of the restrictions imposed by the rules that require geographical proximity for counties to be grouped together. Neighboring counties Forsyth and Guilford satisfied this condition and were clustered by the engine with a suspected *Salmonella* outbreak, as Figure 20 shows. Once this cluster was formed the engine also listed the records of all patients linked to the suspected cluster.

**Figure 20. Scenario 2—Excerpt of Suspected Cluster Information Generated by NCFEDA Analytics Engine**

```plaintext
county_th_reached(wake,campylobacter,giu)
county_th_reached(durham,salmonella,giu)
county_th_reached(mecklenburg,salmonella,giu)
county_th_reached(mecklenburg,campylobacter,giu)
county_th_reached(guilford,salmonella,giu)
county_th_reached(forsyth,salmonella,giu)
county_th_reached(new_hanover,salmonella,giu)
county_th_reached(new_hanover,shigella,giu)
suspected_cluster(forsyth,guilford,salmonella,giu)
record_susp_cluster(forsyth,guilford,pt01332,salmonella,forsyth)
record_susp_cluster(forsyth,guilford,pt01340,giu,forsyth)
record_susp_cluster(forsyth,guilford,pt01341,salmonella,guilford)
record_susp_cluster(forsyth,guilford,pt01344,giu,guilford)
record_susp_cluster(forsyth,guilford,pt01353,giu,guilford)
record_susp_cluster(forsyth,guilford,pt01354,salmonella,forsyth)
record_susp_cluster(forsyth,guilford,pt01364,giu,forsyth)
record_susp_cluster(forsyth,guilford,pt01374,giu,guilford)
record_susp_cluster(forsyth,guilford,pt01384,giu,forsyth)
record_susp_cluster(forsyth,guilford,pt01386,giu,guilford)
susp_outbreak(forsyth,guilford,salmonella)
```
Additional recall notices, some of which are shown in Figure 21, were fed to the NCFEDA system to simulate in Scenario 2 normal conditions of operation for NCFEDA, because new food recalls are issued almost every day by the national agencies. These recalls’ statements were automatically generated by harvesting the FDA webpage to collect recent official Recall Notices, and later by post-processing those data to extract the information to be translated to a format that the NCFEDA Analytics Engine can understand.

Figure 21. Scenario 2—Excerpt of New Recall Notices Processed by NCFEDA Analytics Engine

It is important to notice that the reason for a food recall may be different from contamination by a pathogen. This is illustrated by the data in Figure 21 which list some recalls involving (1) the presence of an allergen (e.g., “undeclared peanuts”); (2) chemical contamination (i.e., “contaminated with monensin”); and (3) product adulteration (i.e., “may contain foreign particles”). None of these problems would match a diagnosis listing a specific pathogen, but some may match a GIU generic diagnosis if the patient’s symptoms do not give clear indication of a different condition (e.g., food allergy or chemical poisoning).

NCFEDA Analytics Engine processing of the information on the suspected cluster it has found with the previous and new recall notices, and consumer complaints, leads to the generation of the solution model presented in Figure 22. Given that the suspected cluster is caused by Salmonella we see that the engine’s rules links this cluster to two previously known recalls also caused by Salmonella (i.e., recalls rc1 and rc2).
However, it gives preference to recall rc1 because it finds that that product was distributed to North Carolina, while the rc2 product was distributed to a much larger geographic area (i.e., nationwide). As expected, the engine did not find evidence to link the suspected cluster to any of the new recalls, because only one of those was due to *Salmonella*, but its product was distributed exclusively in Michigan, and therefore does not affect North Carolina. Two consumer complaints were, however, found to be possibly connected to the suspected cluster. But because “meat” is not the product listed by these complaints, the engine rules cannot find a tighter connection among the suspect cluster, recall notices, and local complaints. Thus, the ELI index computed by the engine for the suspected cluster has a value of 5, indicating “significant likelihood” for this cluster.

### NCFEDA Use Case

A typical situation faced by public health officials tasked with detection of and response to foodborne contamination events involves one or more clusters of unknown illness cases presenting symptoms of gastrointestinal problems which may be an indication of an ongoing foodborne disease outbreak. Recognizing that this is a common scenario our stakeholders must handle, the first NCFEDA use case chosen for implementation as a working software prototype focuses on a foodborne event that is triggered by such a cluster. The use case develops over a 3-day period with daily, new incoming information from various sources being provided to NCFEDA’s relevance engine that evaluates it against knowledge previously acquired and determines what information is relevant to users, and measures the likelihood that a foodborne event or threat is occurring.

The data sources provided to NCFEDA for the use case are (1) Illness case records reported to NCDPH containing among other fields the office visit date, probable diagnosis, and a patient’s county of residence; (2) Recall notices of food products issued by FDA stored as records containing the recall issuing date, the product recalled, the company recalling the
product, the cause for the recall (i.e., pathogen causing the contamination when available) and areas (states) where the product has been distributed; and (3) Consumer complaint records of individuals’ calls to NCDA&CS implicating a food product including all of the following fields if available: date of the call, complainant county of residence, product implicated, retailer/manufacturer/food service provider implicated, complainant medical status (i.e., illness, hospitalization), diagnosis, and description of the complaint.

The prototype was developed under the premise that at an initial phase of use, access to the NCFEDA software web-based tool was to be restricted to those representatives of North Carolina regulatory agencies. This implies that each agency using the system must register those individuals in their staff who should be granted access to the tool. Users would first connect to the system by accessing a NC login page, shown in Figure 23, where they must enter the name of the agency they work for, their user login ID, and a personal password to be verified by the system before any further access can be granted to the proprietary views designed for each agency and those views common to all participants. We envision that the login page can also provide users with links to sites hosting most relevant official and unofficial news related to food safety. A snapshot of the prototype login page depicted in Figure 23 informs the title and provides a direct link to the latest recall issued by FDA, and similarly to the latest recall issued by USDA, and an additional link to a site hosting food safety top news.

Figure 23. NCFEDA Login Page
In what follows, we briefly describe the 3-day use case implemented in this demonstration and provide snapshots of the prototype screens at each step. We assume that the prototype has been receiving daily reports from all its data sources and operating as expected. Thus, when the NCFEDA prototype application is started by a user submitting its credentials via NCFEDA login page a sample of recent data is loaded into the prototype tables corresponding to data received on those days previous to the use case scenario. The initial NCFEDA screen, shown as Figure 24, corresponds to a Common Operating View of food-related events occurring in North Carolina and is intended to be used by all agencies. Its goal is to increase user Situational Awareness of ongoing events. This page is composed of six main areas: (1) Navigation Tabs at the top left allow a user to travel between different screens; (2) The Demo Clock and Start Button on the top right are central to the synchronization of events for the prototype demonstration which compresses events occurring during a 3-day period into a 3-minute presentation; the clock informs the date and time as the demonstration progresses once the user clicks the start button; (3) The North Carolina Map depicting all the state’s counties presents a graphical rendition of the locations where events are occurring to help users assimilate the spatial distribution of possible threats; (4) The Emerging Events Table keeps a continuous record of any possible emerging event identified by the NCFEDA engine, and “pushed” to users via separate pop-up windows, by listing a short description of the result of its analysis and the corresponding ELI rating it attributed to the event; (5) New Incoming Reports/ Information Relevant To Food Safety In NC area in the middle of the screen displays three tables containing key fields of data records from three sources (Consumer Complaints Received by NCDA&CS; Illness Cases Reported to NCDPH; and Food Recalls issued by USFDA); and (6) NCFEDA Searchable Database of Food Safety Reports table at the bottom of the screen provides an easy mechanism for users to query NCFEDA databases by typing words of interest on dedicated search areas attached to each field. Each table on this screen also allows users to do executing selective ordering of the table’s records by pressing on up/down arrows associated to each field.

Once the Start Button is pressed, the use case demonstration script starts playing and new information is dynamically added to tables while simultaneously analyzed by NCFEDA. The demo can be paused at any time by pressing the same button which now reads “Pause.” When the demo is in “Pause” mode the button changes back to “Start” and can be clicked at any time to restart the demo. A summary of the each day scripted for this use case is presented next.
Figure 24. Screenshot of NC Events Initial Page of NCFEDA Use Case

More detailed information about food safety reports involved in emerging events analysis are available in the tables below.

<table>
<thead>
<tr>
<th>New Incoming Reports/Information Relevant to Food Safety in NC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Consumer Complaints Received by NCDACS</strong></td>
</tr>
<tr>
<td>CALL ID</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>1002010</td>
</tr>
<tr>
<td>1002010</td>
</tr>
<tr>
<td>1002010</td>
</tr>
<tr>
<td>1002010</td>
</tr>
<tr>
<td>1002010</td>
</tr>
</tbody>
</table>

Search NCFEDA Database for accessing each complete food safety report received in the last three months by NC Public Health Agencies.

<table>
<thead>
<tr>
<th>NCFEDA Searchable Database of Food Safety Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>REPORT TYPE</strong></td>
</tr>
<tr>
<td>Illness Case</td>
</tr>
<tr>
<td>Illness Case</td>
</tr>
<tr>
<td>Illness Case</td>
</tr>
<tr>
<td>Illness Case</td>
</tr>
<tr>
<td>Illness Case</td>
</tr>
</tbody>
</table>

WARNING: This is an experimental site being developed by the UNC-CH NCFEDA Project in collaboration with the NC Food Safety and Defense Task Force. The NCFEDA Project is funded by the Department of Homeland Security through the Institute of Homeland Security Solutions.
**Demo Day 1.** A cluster of a small number of illness cases with general symptoms of GIU is detected by public health surveillance systems and without additional laboratory tests’ results or a more precise diagnosis, no pathogen is identified. These records are displayed in both the NCDPH records’ table and the Searchable Database table. NCFEDA’s other data sources—food products’ recalls and consumer complaint calls, are analyzed against this triggering event but do not provide any relevant information that can be linked to these cluster cases. NCFEDA’s engine computes the ELI rating for this situation and displays a short message in the Emerging Events table to inform the user of its findings. The locations of offices’ visits (marked by an iconic red cross) and complainant consumers’ counties of residence (marked by the icon of a green telephone) are plotted in NCNCFEDA’s North Carolina map as shown in Figure 25.

**Demo Day 2.** A new small cluster of Salmonellosis cases is detected by public health. Given that now a pathogen has been positively identified, NCFEDA searches among both incoming and previously active recall notices to verify if any of those are a result of contamination by *Salmonella* and among those recalls if the recalled product has been shipped to North Carolina. But no results are found. NCFEDA also searches among incoming consumer complaint calls, and those complaints under investigation by NCDA&CS, for any illnesses identified as caused by *Salmonella* or implicating food products susceptible to this pathogen. NCFEDA narrows the search to a call reporting hospitalization because of possible consumption of contaminated fruit. Because fruit is susceptible to *Salmonella*—as documented by existing recall data, the NCFEDA relevance engine deduces that there is a possible emerging contamination event because of *Salmonella* and issues a warning to users. An emerging events map pops up on a separate window, as shown in Figure 25, displaying the location of all relevant events linked to this threat. When the user hovers the computer mouse over the map icons, detailed information about each reported case/complaint is displayed. In addition, the Event Likelihood Index for the event is computed (ELI=3) by the relevance engine and reported on the corner of this pop-up window.
Figure 25. Screenshot of NC Events Page of NCFEDA Use Case Showing A Warning with ELI=3
**Demo Day 3.** An increasing number of illness cases are reported to public health and new clusters are detected, but lacking personalized information about patients because of privacy concerns NCFEDA cannot deduce relationships among patients’ cases beyond same county of residence. The arrival of a new recall notice expanding the area of distribution of the recalled product to the state of North Carolina and previously restricted to three other states in the West Coast of the United States is processed and analyzed by the relevance engine. Cantaloupe is the recalled product because of contamination by *Salmonella*. The relevance engine recognizes that cantaloupe is a fruit and that the pathogen causing this recall is also the same pathogen causing a reported illness and hospitalization as reported by a consumer complaint call. NCFEDA issues a new warning, as shown in **Figure 26**. The emerging events map appears displaying all events linked to this threat which now includes a recall notice (shown in the yellow box on the bottom left corner of the pop-up window viewed in **Figure 26**). The new ELI rating computed is higher (ELI=5) than before since more connections among data records were discovered and confirmed by the relevance engine. The message “pushed” to users is also more precise including the suspect food product (cantaloupe) and the pathogen (*Salmonella*).

The scripted portion of the use case demonstration completes at the end of the *Demo Day 3*. The prototype contains two other working screens that can be visited and demonstrated to users independently of the use case script as described next.

Another functionality available/common to all users is the *US Recalls* screen, depicted in **Figure 27**, and accessible by clicking on the “US Recalls” tab located in the Navigation Tabs area of the application. The goal of this screen is to help users put into the national context all active recalls across the United States. It contains three main areas: (1) *Navigation Tabs* (common to all NCFEDA screens); (2) A *Search Engine*, presented on the left side of the screen, enables users to find in real-time more detailed information about any existing event by searching the web; and (3) A *Food Recalls Map* shows a graphical representation of the states/areas affected by currently active food recalls.

Typing one or more words in the search area of the *Search Engine* and clicking the *Search Button* will provide users with four different types of search results related to the word of interest: (1) Links to LOCAL resources (those located in North Carolina); (2) Links to general WEB resources (i.e., articles, papers); (3) Links to VIDEO resources available on the web that match the search; and (4) Links to BLOGS with content matching the searched information.

At the top of the *Food Recalls Map* area, users can select from a drop-down menu, a specific food product—as classified by the FDA Food Classification Code—to request the display of the distribution area covered by currently active recalls of such a product. The states affected by the recall are then colored on the map. Hovering over the colored area of the map forces the display of a box containing the detailed information contained in the recall selected. NCFEDA uses data and text mining techniques to extract this information from the free-text Recall Notice(s) issued by the FDA/USDA it harvests directly from these agencies’ websites.
Figure 26. Screenshot of NC Events Page of NCFEDA Use Case Showing A Warning with ELI=5
An example of a proprietary and customized dashboard for the NCDA&CS is provided by clicking tab **NCDA&CS Complaints** (Figure 28). It is intended that only representatives of this agency will be granted access to this dashboard, unless the agency authorizes others on an individual basis. This dashboard contains tools to assist NCDA&CS users in performing retrospective analysis on 12 years of NC consumer complaints data. There are four main areas in this screen: (1) **Navigation Tabs**; (2) **Complaints Calls Graph** which presents the annual distribution of calls per food product implicated; (3) **US Map**, and associated selection controls, marked by icons depicting the geolocation of consumers/retailers/manufacturers associated with the calls; and (4) **Events Heatmap** showing a graphical representation of food safety data (i.e., illness cases, lab results, complaint calls, and FDA/USDA recalls) where the magnitude of monthly values are represented by a color gradient. Hovering over any of the graphic areas of the dashboard enables users to view information details and slice-and-dice the data.
Conclusions

Our nation’s food safety system depends on the coordinated and speedy completion of a set of tasks related to the detection of and response to foodborne disease outbreaks and food contamination events. Completion of these tasks is hampered by gaps in the information acquisition, sharing, and analysis processes. These gaps are the source of major latencies between the first case of confirmed foodborne disease and the removal of all offending products from retail, institutional, and residential shelves across the country. In the case of the widely reported Peanut Corporation of America Salmonella contamination, the latency between first reported illness and the recall of the last contaminated product from retail shelves was nearly 10 months.

A gap analysis of food safety system in North Carolina revealed that latencies in the surveillance and response processes could be reduced by the development of a new informatics tool that provides situational awareness across the food safety stakeholder
community including public sector regulatory agencies, private sector companies, and consumers. In response to these results, the North Carolina Foodborne Data Integration and Analysis Tool (NCFEDA) was developed. NCFEDA is envisioned as a “meta”-level software tool that provides critical capabilities in data integration, visualization, and analytics for managing a foodborne contamination event.

Conceptually, the NCFEDA is composed of several modules which together constitute its Analytics Engine—the intelligent component of the system responsible for drawing conclusions about a given food safety situation. NCFEDA reasoning capabilities are powered by formal logic. This means that the NCFEDA system “reasons” about food safety events by applying deductive reasoning (i.e., inference rules) to facts informed to the engine to infer (new) knowledge. A use case was developed to demonstrate the latency reducing capabilities of NCFEDA that involved one or more clusters of unknown illness cases presenting symptoms of gastrointestinal problems.

Noel P. Greis, PhD, is director of the Kenan Institute’s Center for Logistics and Digital Strategy and professor of Operations, Technology and Innovation Management at the Kenan-Flagler Business School at the University of North Carolina at Chapel Hill. Dr. Greis is the co-director of the recently established UNC-Tsinghua Center for Logistics and Enterprise Development in Beijing, China, a joint center of Tsinghua University’s Department of Industrial Engineering and the Kenan-Flagler Business School at the University of North Carolina at Chapel Hill. Dr. Greis’s research is transforming the way we predict, evaluate, and respond to complex and critical events in domains such as medicine and public health, food safety, supply chain management and logistics, defense and security, as well as energy and the environment. Dr. Greis is an expert in the area of intelligent systems design and development and works with organizations to develop knowledge-based systems and predictive analytics that support decision-making in complex, disruptive, and dynamic environments. Dr. Greis is also an expert in the use of intelligent agent-based modeling and simulation to predict the behavior of complex systems—thus improving decision-making capability.

Monica Nogueira, PhD, is director of the Intelligent Systems Laboratory (ISL) of the Center for Logistics and Digital Strategy at Kenan-Flagler Business School (KFBS) at The University of North Carolina at Chapel Hill. In this capacity, she is responsible for overseeing the projects developed by the ISL for its corporate and institutional clients and works closely with UNC professors and students and the KFBS staff. Dr. Nogueira is an expert in data modeling and analysis, which she applies to design and build decision support tools that mine and correlate information from large and diverse datasets to extract relevant knowledge that enables users to act on those problems that are most significant to them. She uses her computer science expertise to design and implement new software applications that utilize state-of-the-art technologies to solve practical problems for ISL external and internal customers. Her primary research interests include new technologies and their practical uses to create new methodologies that support “intelligent” tools and their application to everyday
problems in logistics and supply chains and data and text mining methodologies and tools for knowledge discovery and extraction. Dr. Nogueira has developed a number of projects and tools that demonstrate the use of radio frequency identification (RFID) technology for controlling the safety of perishable products (i.e., cold chain for food and medical drugs). In these projects, RFID is used as the integrator element that allows tracking and tracing of the perishables and total visibility throughout the supply chain, guaranteeing their safety.

References


Acknowledgements

The NCFOODSAFE project team wants to thank the many stakeholders of the food safety system in North Carolina for their generous support of this project, especially the contribution of their ideas, experiences, and data. The team particularly wants to thank Brett Weed, Food Defense Coordinator at the North Carolina Department of Agriculture and Consumer Services, for his championship of this project and for providing key data sets that drive the NCFEDA software application. The members of the North Carolina Food Safety and Defense Task Force, in particular Cris Harrelson, David Bergmire-Sweat, and Brett Weed, provided critical feedback during the course of the project. The team is grateful to all the anonymous staff from local health departments who were interviewed and provided valuable information to the project. The team also wants to thank the Institute for Homeland Security Solutions for providing the funding that made this project possible.